The Costs of Payment Uncertainty in Healthcare Markets

Abe Dunn
Bureau of Economic Analysis

Joshua D. Gottlieb
University of Chicago and NBER

Adam Shapiro
Federal Reserve Bank of San Francisco

Pietro Tebaldi
University of Chicago and NBER

April 2020

Working Paper 2020-13

The Costs of Payment Uncertainty in Healthcare Markets[*]

Abe Dunn   Joshua D. Gottlieb   Adam Shapiro   Pietro Tebaldi

April 24, 2020

Abstract

What does it cost healthcare providers to collect payment in the complex U.S. health insurance system? We study this question using rich data on repeated interactions between a large sample of physicians and many different payers, and investigate the consequences when these costs are high. Payment uncertainty is high and variable, with 19% of Medicaid visits not reimbursed after the first claim submission. In such cases, physicians either forgo substantial revenue or incur costs to collect payment. Using physician movers and practices that span state boundaries, we find that providers respond to these costs by refusing to accept Medicaid patients in states with more severe billing hurdles. This supply margin is even more responsive to these costs than to reimbursement rates. Using these supply estimates, we calculate that the costs of billing Medicaid consume one-quarter of the average revenue from a Medicaid visit. We estimate a model of the billing process, and find that the variable costs of billing each visit account for 21 percentage points of this total cost. Analyzing healthcare prices without accounting for billing costs and payment uncertainty may substantially misrepresent differences between private payers and Medicaid.

The complexity of the U.S. health care system is legendary (Cutler and Ly, 2011). Along with taxes, medical bills are among the most complicated to produce and process correctly. With every patient visit, there is a chance that collecting payment turns into a costly bargaining process between payer and healthcare providers (such as those considered in other industries by Backus et al., 2018; Larsen, 2014; Jindal and Newberry, 2015). This is a form of contractual incompleteness: even after care has been provided, the physician is still unsure of how much revenue she should expect, and how costly will it be to collect it.

Using rich administrative data on interactions between physicians and insurers, we offer the first detailed look at the costs of the medical billing process as faced by healthcare providers. We estimate the direct costs of the billing process itself, and a major indirect cost that results: doctors' unwillingness to treat publicly insured patients caused by billing hurdles.

Medicaid is the insurance category with the highest billing complexity (Gottlieb, Shapiro and Dunn, 2018), and we estimate that doctors' direct costs of billing Medicaid consume one-quarter of the revenue from Medicaid patients. The indirect costs arise because increased difficulty of doing business makes it less attractive to engage in economic activity (Djankov et al., 2003, 2008; Besley, 2015). We find that the the difficulty of collecting payment influences physicians' willingness to supply Medicaid visits even more than the payment rates themselves. A 6% improvement in the physician's ability to collect revenue from Medicaid has as much of an effect as a 14% increase in Medicaid's payment rate itself.

We base our empirical findings on a novel type of health care data, called "remittance data", which Gottlieb, Shapiro and Dunn (2018) introduced and summarized. The remittance data allow us to observe multiple rounds of interactions between payers and physicians, along with detailed information about the medical provider, the patient, the visit, and the reasons for why payments are denied or reduced. This data provide far more detail about the billing and collection process than the claims data that have become widely used to study

healthcare markets.[1]

The paper consists of three main parts. First, in section 1 we present the institutional background necessary to understand our context and data. Section 2 then presents descriptive statistics and stylized facts that underpin and guide our estimates and empirical conclusions. Most notably, we find that Medicaid bills experience very high denial rates and physicians often resubmit denied Medicaid claims.

In the second part of the paper, we estimate the causal effect of Medicaid billing complexity on the probability of a healthcare provider accepting Medicaid patients. We also estimate how changes in prices ("reimbursement rates" in medical jargon) affect patient acceptance. This price parameter has been analysts' main focus when studying physicians' reluctance to accept Medicaid (Polsky et al., 2015; Candon et al., 2018; Oostrom et al., 2017; Alexander and Schnell, 2019), but we find that billing hurdles are at least as significant. Section 3 introduces a conceptual framework that tells us how to think about these two factors simultaneously. It shows, quite naturally, that supply ought to respond positively to both fees and the recovery rate for those fees. But it also shows us how the magnitudes of these effects can differ and how we can interpret these magnitudes economically. In particular, we can use these estimates to make inferences about the share of Medicaid revenue spent on the billing process.

Our empirical strategies, discussed in sections 4 and 5, exploit systematic differences in the administration of Medicaid across states (Layton, Maestas, Prinz and Vabson, 2019). These differences lead to state-level variation in the reimbursement fees, and in the probability that a claim is not reimbursed after the first submission and is instead subject to *ex post* provider-payer negotiation. The remittance data allows us to construct reduced-

---

[1]Recent uses of *claims* data—distinct from our *remittance* data—include the widely used Medicare and Medicaid claims data, MarketScan (Dunn, Liebman, Pack and Shapiro, 2013; Clemens and Gottlieb, 2017), the Health Care Cost Institute (Cooper, Craig, Gaynor and Van Reenen, 2018), individual firm data (Einav, Finkelstein, Ryan, Schrimpf and Cullen, 2013; Brot-Goldberg, Chandra, Handel and Kolstad, 2017) and even many so-called All-Payer Claims Databases (Ericson and Starc, 2016; Liebman, 2018; Panhans, 2019); the latter were kneecapped when the Supreme Court in *Gobeille v. Liberty Mutual* (2016) prevented states from requiring self-insured firms to include their claims.

form indices of billing complexity by state, which we use as new explanatory variables for a key policy-relevant outcome: the probability that a provider accepts patients covered by Medicaid.

We use two alternative identification strategies, which generate very similar results. First, we consider physicians who move across states, and compare their probability of accepting Medicaid patient before and after the move (Abowd et al., 1999; Finkelstein et al., 2016, 2019; Molitor, 2018; Hull, 2018). Second, we compare the Medicaid acceptance probability across clinic locations that operate in different states but but are managed by the same physician group or health system. The first strategy controls for any differences in individual physicians' specialization or preferences, such as the level of altruism towards Medicaid patients. The second strategy addresses the concern that Medicaid supply decisions may depend on a group's managerial competence or organizational structure.

Both strategies generate robust evidence that physicians' decisions respond to payment uncertainty along with actual fees. We scale our indices such that an increase of 0.1 corresponds to a 10 percent increase in fees, or to a 10 percent increase in the recovery rate (share of revenue ultimately collected). We find that a 10 percent increase in fees leads to a 0.6 percentage point increase in physicians' willingness to treat Medicaid patients. A 10 percent increase in recovery rate leads to a 1.4 percentage point increase in accepting Medicaid.

The larger coefficient on recovery rate likely reflects the costs of payment uncertainty. To understand this, observe that a 10 percent decline in fees translates to a 10 percent decline in revenue, conditional on the quantity and composition of services provided. But a 10 percent decline in the observed recovery rate likely reflects an even steeper initial decline due to increases in Medicaid's billing complexity. But some effort from the physician may offset this hassle. In addition, a 10 percent decline in the recovery rate likely implies additional uncertainty in the ultimate payment. When deciding whether to accept Medicaid, the physician accounts for the effort cost and the cost of payment uncertainty. So it makes sense for doctors to respond to recovery rates more elastically than to fees, just as we observe.

3

Based on this logic, we use a simple framework to infer the implied costs of payment uncertainty and billing effort. Our estimates imply that these costs are around one-quarter of Medicaid revenue. This is consistent with the upper end of estimates using time-driven activity-based costing in a single large academic health care system (Tseng et al., 2018).

In the third part of the paper, in section 6, we develop an empirical model to obtain direct estimates of the costs of billing hurdles relying on more granular decisions. The model considers the dynamic decision-making process of claim preparation, submissions, and resubmissions generating the type of administrative costs we are interested in. Rather than variation across states, here we exploit the fact that providers (which we assume behaving optimally with rational expectations) make different billing decisions as they face varying values of each visit, and different probabilities of claims being paid without any negotiations. The model allows us to recover two sets of structural parameters. The first set of parameters describes the cost of "initial" billing effort which takes place after the visit but before claim submission. This effort can limit the probability of entering a costly and uncertain back-and-forth between payer and provider. The second set of parameters captures instead the costs of these negotiations, and precisely the cost of resubmitting a claim that the payer refused to pay.

The estimates of the model are remarkably consistent with our previous findings. The total cost of effort before submitting the initial claim, plus costs of any subsequent resubmissions, amount to 21 percent of revenue. After accounting for the initial billing effort, the probability of payments, and the resubmission costs, the per-visit net revenue for the average Medicaid visit is 56% of the first-claim billed amount. Our preliminary estimates of the model using data on private payers (not shown) suggest that, in the commercial segment of the market, the net expected revenues accounting for billing costs and payment uncertainty is 64% of the initial expected payment. This suggests that analyzing differences in healthcare prices ignoring billing costs and revenue uncertainty may understate the differences between private payers and Medicaid.

This paper contributes to a broader literature in public economics that considers administrative barriers to program participation as an ordeal, which may improve program targeting (as in Nichols et al., 1971; Nichols and Zeckhauser, 1982; Besley and Coate, 1992; Finkelstein and Notowidigdo, 2019). These ordeals generally attach to consumers, rather than suppliers, so program complexity deters beneficiaries' participation in SSI (Bound and Burkhauser, 1999), food stamps (Currie et al., 2001), and student aid (Dynarski and Scott-Clayton, 2006). In other contexts, Benzarti (2017) uses a policy change to elegantly infer the costs of claiming tax deductions.

But to obtain benefits from health insurance programs, *providers* must submit bills and collect payment. So the ordeals Medicaid imposes, in the form of administrative hassle, directly impact the supply side. This is one of the most complicated settings of any public program, but we are unaware of any previous evidence on how administrative hassle restricts Medicaid spending through the supply side. We open this literature in the costly, and policy relevant setting of physician participation in health care programs.

Abundant evidence shows that physicians' health care supply responds to their payment rates (Gruber, Kim and Mayzlin, 1999; Clemens and Gottlieb, 2014; Dunn and Shapiro, 2018; Alexander and Schnell, 2019). But the impact of billing hassle on supply has only been explored in small descriptive surveys (Cunningham and O'Malley, 2008; Long, 2013; Ly and Glied, 2014).[2] We study both dimensions of supply decisions together, and do so using nationally comprehensive data on Medicaid acceptance and novel estimates of billing hassle. This is the first empirical evidence of how billing costs affect the supply of patient care.

---

[2]In the hospital inpatient context, Gowrisankaran, Joiner and Lin (2019) show that electronic health records and Medicare payment policies interact in subtle ways to drive how hospitals code and bill for the care they provide. In a very different setting (corporate taxation), Zwick (2018) makes a similar point: accountants' sophistication influences the tax deductions that firms claim.

1 Institutional Background and Data

1.1 Billing in the U.S. Healthcare System

Institutional details of the United States health care system are critical to understanding our data and our analysis, so we begin by providing an overview of this billing process. When patients covered by health insurance visit physicians, they rarely make up-front payments. Instead, the medical practice submits a bill to the patient's insurer after the visit. This process is similar for private insurers—such as insurance plans purchased through employers or a health insurance exchange—and public insurers, such as Medicare for the elderly and Medicaid for the indigent.

The physician can submit the bill herself, or her practice can hire staff in a billing office, or contract with a third party intermediary called a clearinghouse. The first step in billing is to determine exactly what care the physician provided. This care is described in detail using the "Healthcare Common Procedure Coding System" (HCPCS), which contains approximately 13,000 potential treatments, along with supplementary codes that modify the basic code ("modifiers"). A claim may contain one or more such procedure codes. The physician or biller must also classify the patient's diagnosis using International Classification of Diseases (ICD) codes. She also needs to collect and report the patient's personal details and insurance coverage.

Once the information is prepared, the biller submits a claim to the patient's insurer. In the initial stage of billing, the information required and method of submission is standardized.[3] Using a specific format established by the federal government, the physician provides the insurer with identifying information for the patient and his insurance plan, the treatment provided (using HCPCS codes), the diagnosis (ICD) codes that justify that treatment, and the amount she would like to be paid.[4]

---

[3]Indeed, the process is formally standardized through CMS Form 1500, its electronic version EDI 837 (established by HIPPAA), and the Electronic Remittance Advice EDI 835 (whose glorious minutae the reader will experience in great detail below).

[4]These billed amounts are infamously outrageous, but may sometimes provide a baseline for the negotiated rates we use. We discuss our imputation methods in Appendix A.1. In the hospital payment context,

The insurer receives the claim from the biller, analyzes, and processes it. At the initial stage, this processing and decision may be handled by a third-party contractor acting on behalf of the insurer, primarily using an automated system containing payment and audit rules. This system determines whether the patient has eligible insurance, whether the insurance covers the service provided, and whether the medical care was appropriate. The insurer can also use this opportunity to look for any fraudulent claims.[5] When this evaluation is complete, the insurer can pay the physician, deny the claim, or pay for some parts while refusing to pay other parts. If paying the claim, the insurer's system must determine the payment amount. This amount should follow from an existing regulation or contract: for public insurance, the (state or federal) government establishes the rates by legislation and regulation. For private insurance, the insurer and physician will have agreed on a set of payment rules in advance.[6]

The insurer transmits its decision to the physician using a standardized electronic format, called Electronic Data Interchange 835, "Electronic Remittance Advice," which we refer to as simply a "remittance." These remittances tell the physician whether the insurer has approved the claim, how much money to expect from the insurer, and how much they are authorized to collect from the patient—the patient owes any deductible, copayment, or coinsurance directly to the physician. Depending on the physician's exact billing arrangement, the remittances may be sent straight to the physician's office or to a clearinghouse—an intermediary who the physician has engaged to process her claims.

If the process goes smoothly, the only remaining step is to collect payment. The insurer should transmit its part of the payment directly to the practice. Based on the information in the remittance, the physician can send a bill to the patient for their part.

But the process is not always this smooth. Instead of approving the claim, the insurer

---

Reinhardt (2006) describes these list charges and Cooper et al. (2018) find that they still form an important part of many hospitals' payment contracts.

[5]Although there are questions about how thoughtfully they do this (Allen, 2019) and whether they even have incentives to do so (Cicala, Lieber and Marone, 2019).

[6]See Clemens and Gottlieb (2017) and Clemens, Gottlieb and Molnár (2017) for more details on these contracts.

may deny it. The insurer may question the validity of the patient's insurance coverage, the medical justification for the treatment, or whether the insurance contract covers the care provided. It may question whether the physician has submitted the correct codes, or authorize less payment than the doctor was expecting under the payment contract. In fact, the organization that manages the Electronic Data Interchange standards maintains a list of around 350 codes for different reasons claims may be adjusted or denied.[7]

When a claim is denied, not paid, or the physician thinks the payment was too stingy, the process can continue in a few different ways. The physician can give up on the claim and write off the lost revenue. If she has not signed a payment contract with the insurer (i.e., she is "out-of-network") she may be able to bill the patient directly for any missing revenue. But in the more common situation where the physician has a contract with the insurer ("in-network"), that contract likely forbids her from asking the patient to pay for amounts the insurer has not authorized. So the physician's only option in most cases is to deal with the insurer directly.

The biller's next steps depend on why the claim was not approved. If the insurer questions the medical necessity of the treatment, the physician may have to provide additional documentation about the patient's condition, either through an online submission form or by fax. If there is an administrative error, such as a typo in the patient's name or insurance details, the practice may need to submit a corrected claim. If the physician thinks that the claim adjudication does not comply with her contract, she may have to submit a formal appeal to the insurer, requiring manual intervention and a decision by someone higher in the insurer's hierarchy. Each time the insurer processes the claim, it generates a remittance to convey the decision to the physician or her agent.

All of the processing and adjudication in this system absorb considerable resources. It also generates massive amounts of rich data, which allow us to investigate the interactions between physicians and insurers, and the implications of this process for physicians' decisions.

---

[7]http://www.x12.org/codes/claim%2Dadjustment%2Dreason%2Dcodes/

1.2 Data

Our primary data source is a convenience sample of remittance data, taken from the Electronic Remittance Advice forms (EDI 835).[8] Our data source is IQVIA Real World Data—Remittance Claims. IQVIA obtains these data from clearinghouses that receive the remittances on physicians' behalf.[9] Since the physician practice chooses which clearinghouse to work, we are effectively sampling at the physician level. For the physicians covered, we observe their interactions with the full set of insurers. This is a key difference from claims datasets used in much other research, which generally have exhaustive data from an insurer or set of insurers, but don't see the full breadth of any physician's business and often can't link the same physician across insurers.[10] The even more critical difference between our data and claims datasets is that we don't just see final processed and paid claims. We see the remittances generated each time the insurer repsonds to a physician's submission or resubmission—including those remittances indicating claim denial, nonpayment, or other complexity.

For each remittance, the data tell us the providing physician (including the national identification number), the practice submitting the bill, its zip code, the insurer providing the remittance, and the type of insurance. We see the detailed procedure (HCPCS) codes indicating what care was provided, ICD diagnosis codes, and key dates: when the service

---

[8]Since the data provider includes remittance data from whichever clearinghouses it contracts with, rather than a systematic random sample, we naturally worry about the sample's representativeness. To gauge this, and for several parts of our analysis, we merge the data with physician characteristics from the MD-PPAS and SK&A data (see below). Along observable dimensions, the physicians included in our data appear very representative of the covered specialties nationwide. This gives us a higher level of confidence on the nationwide represen4veness of our results.

[9]According to the data provider, "A Remittance Claim is an Electronic Data Interchange (EDI) 835 transaction set called Health Care Claim Payment and Remittance Advice. It provides detailed payment information relative to a health care claim and, if applicable, describes why the total original charges have not been paid in full. IQVIA collects these claims through practice management software vendors and switch clearinghouses."

[10]This includes the MarketScan data (Dunn, Liebman, Pack and Shapiro, 2013; Clemens and Gottlieb, 2017), Health Care Cost Institute data (Cooper, Craig, Gaynor and Van Reenen, 2018), individual firm data (Einav, Finkelstein, Ryan, Schrimpf and Cullen, 2013; Brot-Goldberg, Chandra, Handel and Kolstad, 2017) and even many so-called All-Payer Claims Databases (Ericson and Starc, 2016; Liebman, 2018; Panhans, 2019); the latter were kneecapped when the Supreme Court excluded data from self-insured firms in *Gobeille v. Liberty Mutual* (2016).

was provided, when the claim was submitted, and when the insurer made its decision. We see how the insurer handled the claim, including the summary of its decision (paid, denied, etc.), justification for any claim line adjustments, and how much it is paying. At the patient level, we do have an encrypted code that allows us to link the same patient across remittances, plus the patient's age.

An inherent challenge in data of this form is that we naturally do not observe the payment amounts for claims that are denied or only partly paid. We resolve this problem using two methods to impute the amount that would have been paid, had the claim been approved and processed smoothly, according to the insurer's contract with the physician. The first imputation method uses comparable claims from the same physician, and same insurer, which were fully paid. We use the average payment amount for a given service to impute the contractual payment amount when it is missing. The second method takes advantage of other services paid by the insurer to that physician on other claims. Using the full sample of other claims, we compute the average discount relative to the "billed amount" submitted by the provider. We then apply this average discount to determine the contractual payment amount when it is missing. Clemens, Gottlieb and Molnár (2017) find that three-quarters of private insurance payments to physicians are based on constant markups over Medicare charges. Cooper et al. (2018) find that, for hospitals, the other common payment method is discounts from billed amounts. The two methods yield very similar results.

We complement our data with two additional sources. The Centers for Medicare and Medicaid Services provides a dataset that it regularly updates with information on physicians' specialty, location, and practices. We use this file, called Medicare Data on Provider Practice and Specialty (MD-PPAS), to identify where physicians are located and when they move. We also use it as part of our approach to identifying those who work in the same practice.[11] Since MD-PPAS has the same physician identification number as the remittance data, merging the two to obtain physician characteristics is straightforward.

---

[11] This does not follow immediately from their tax IDs, since the overlapping organizations of health care providers are quite complex (Brot-Goldberg and de Vaan, 2018).

Finally, we augment the administrative physician characteristics from MD-PPAS with SK&A survey data also purchased from IQVIA. These data, primarily collected by the firm for marketing purposes, come from administrative records and a major manual phone survey of most practicing U.S. physicians. Among the key questions for our purposes, SK&A asks if each physician accepts Medicare patients and whether she accepts Medicaid patients. The data contain additional detail about organizations that characterize the physician's practice structure. This allows us to go beyond the tax ID number from MD-PPAS to determine which physicians are part of the same narrow practice or the same broad system. SK&A also provides the national provider identification number that allows us to merge all of these sources.

2 Descriptive Evidence on Payment Uncertainty

2.1 Summary Evidence of Payment Uncertainty

Table 1 shows key summary statistics of these data for each category of insurance. We show the average value of each line item in the claims we observe and the total value of the visit, which can include multiple line items. The third row shows the line-level probability of being denied, followed by the probability that the entire claim is denied. For claims that are denied, the final line shows the probability that the physician submits it again.

Looking across the insurers, we see that private claims are more valuable than Medicare (both Traditional Medicare and Medicare Advantage) and even moreso compared with Medicaid. Yet Medicaid has a higher probability of claim or procedure-level denial. Traditional Medicaid has a 19 percent probability of claim denial and 32 percent probability of line-level denial. In contrast, Medicare's claim denial rate is 9 percent and line-level denial is 19 percent. Both Medicare Advantage and private Medicaid plans (Managed care Organizations) have rates similar to Traditional Medicare (Medicare fee-for-service).

While these descriptive facts are striking, they do not immediately translate into real economic costs. For instance, suppose that physicians expect to have a certain share of

claims denied, and don't take any action to avoid or challenge these denials. In this case, the denials just mean that the physician faces a lower effective price for her services, and the transfers from insurers to providers are lower than we might otherwise measure. But if physicians fight back against denials, or avoid treating patients with more difficult insurance, then payment uncertainty has real economic costs.

The final row of Table 1 shows the first suggestive evidence about these costs. We see that resubmission rates are highest for Medicaid, at 25 percent for traditional Medicaid and 17 percent for MCOs. These contrast with 16 percent for Traditional Medicare, 9 percent for Medicare Advantage, and 7 percent for private insurance. This suggests that physicians expend substantial ongoing effort to obtain payment.

In order to estimate the costs of this process, we rely on two key aspects of the data. First, the dataset has broad national coverage, and different states' Medicaid programs operate indepenently. In section 2.2 we show the differences across states in fees and billing hassle. We use these differences in sections 4 and 5 to estimate doctors' billing costs based on their extensive margin decisions of whether to treat Medicaid patients. Next, section 6 develops a model of the dynamics of claim submissions to infer the costs of each step in the billing process.

2.2 Fees and Billing Hassles Across States

The fee measure is conceptually simple: we would like to know how much more one state's Medicaid program would pay for identical care compared with another state's. Because care is so heterogeneous, we cannot simply compare average prices for all treatments. Other research on Medicaid fees, such as Alexander and Schnell (2019), has had to hand-collect data from each state. This has limited most studies to considering a few specific services, such as primary care. In order to account for the broader set of care included in our sample, we estimate the following regression to compute price indices that account for the plethora

12

of treatments included:

$$\ln c_i = \xi_{a,j} \cdot \mathbb{1}_a \times \mathbb{1}_j + \chi_k \cdot \mathbb{1}_k + \omega_h \cdot \mathbb{1}_h + \rho_1 \text{patient age} + \rho_2 \text{comorbidities} + \rho_3 + u_i. \quad (1)$$

Each observation in this regression is one service line, i.e. one treatment, and c_i is the amount paid for service i. The key coefficients in the regression are the insurer-by-state fixed effects $\mathbb{1}_a \times \mathbb{1}_j$. These fixed effects represent the contribution of the state and insurer to explaining the variation in payment level. The corresponding $\hat{\xi}_{a,j}$ that we estimate serve as our state-insurer fee index. Since the dependent variable is in logs, we can interpret a 0.01 change in $\xi_{a,j}$ as approximately a 1 percent change in the insurer/state's fee.

In computing this index, the regression adjusts the raw value, c_i, for the characteristics of that service and its claim. Most significantly, we control for fixed effects for the specific procedure code $\mathbb{1}_k$ and for the physician $\mathbb{1}_h$. These controls ensure that our indices reflect differences between comparable medical care, and don't reflect differences in physician composition. In order to identify the state-by-insurer indices with these controls, our data must have physicians who practice across multiple insurers, as well as some physicians who practice in multiple states. The data satisfy these conditions, so we are able to estimate the indices conditional on physician and on procedure code. We show comparable results for indices without physician effects in the Appendix. We also control for patient characteristics, such as age and other diseases they have, in case these influence the cost of the service.

We estimate a similar index for billing difficulty. We follow the same logic as in equation (1), but replace the dependent variable with the payment probability $\ln p_i$, as motivated by section 3. This measure is based on the difference between the amount of the payment ultimately authorized by the insurer, $\tilde{c}_i$ and the original estimated value of the claim, c_i, imputed as described in section 1.2. Specifically, the payment probability is defined as $p_i = \tilde{c}_i/c_i$. This is equal to zero for claims that are not paid, so we replace the dependent variable called for by the model, $\ln p_i$, with the inverse hyperbolic sine (IHS) of p_i. IHS

is an approximation to the log transformation which is defined at zero,[12] specifically as: $\sinh^{-1}(p_i) = \ln\left(p_i + \sqrt{p_i^2 + 1}\right)$. We compute this and then estimate:

$$\sinh^{-1}(p_i) = \psi_{a,j} \cdot \mathbb{1}_a \times \mathbb{1}_j + \chi_k \cdot \mathbb{1}_k + \omega_h \cdot \mathbb{1}_h + \rho_1 \text{patient age} + \rho_2 \text{comorbidities} + \rho_3 + u_i,$$

$$(2)$$

The estimated $\hat{\psi}_{a,j}$ coefficients serve as our index of state-by-insurer billing recovery rate. Just like with the fee index, we can interpret a 0.01 change in $\hat{\psi}_{a,j}$ as approximately a 1 percent change in the insurer/state's recovery rate.

Figure 2 shows the estimated Medicaid indices across U.S. states, and Table 2 summarizes these distributions. We see higher fees in the upper Midwest and northern Great Plains states, and lower fees in the Rust Belt and south. The recovery rates, shown in Panel B, are not particularly correlated with the fees. Among high-fee states, such as Minnesota and North Dakota, the former has one of the highest recovery rates while the latter has one of the lowest. Among low-fee states, such as the lower Midwest, Illinois has a relatively high recovery rates while Michigan and Pennsylvania are low. These reflect each state's significant leeway in how to administer Medicaid. Helpfully for our purposes, it means that we have meaningful independent variation in both variables.

The indices shown here use as much of our data as possible, and reflect certain detailed choices about how to handle various data problems. When we show our empirical results in section 4.2, we will show robustness to numerous other choices about data and index construction. Appendix Figure B.1 shows, although states' values do change from year to year, the indices are overall quite stable.

Figure 3 shows a scatterplot relating the recovery rate index $\hat{\psi}_{a,j}$ and fee index $\hat{\xi}_{a,j}$ across states and across insurers. The pattern across insurers is striking: Medicaid has the lowest

---

[12]This transformation is implemented as `asinh` in statistical software and is one approach to dealing with zeros in p_i, which occur in our context the insurer does not pay so $\tilde{c}_i = 0$. It has recently been used by Card and Dellavigna (forthcoming), Chen (2013), Bastos, Silva and Verhoogen (2018), and Blouin and Mukand (2019). Bellemare and Wichman (2019) discuss how to compute elasticities when using IHS.

fees and lowest recovery rates, while private insurers have higher fees and generally less billing difficulty. We will use these indices to estimate physician supply curves with respect to both fees and billing difficulty, and we begin by developing a framework to guide that analysis.

3 Conceptual Framework

Physicians have no obligation to treat patients if they are not happy with the insurer's fees and hassle. The first question we ask is whether physicians' decide whether to treat Medicaid patients based on not just the regulated prices of different procedures, but also on the hassle costs of dealing with uncertain payments and claim resubmissions highlighted in our data.

To analyze these choices, we begin by sketching a conceptual framework to understand physicians' decisions. The framework accounts for the payment rates physicians expect, the costs of collecting that payment, and the share of revenue that they ultimately don't collect. This framework motivates the empirical approach we take in section 4.1 and guides our interpretation of the results.

We consider the doctor's decision about treating one insurer's patients in isolation. In practice, this insurer will be Medicaid. These costs can include both actual variable costs of care, as well as opportunity costs of spending time on one type of patient when another could be treated instead. Physicians may also have an altruistic motive, especially when it comes to treating Medicaid patients. We model these motives as a shifter of the log costs, so the doctor will treat patients if the net revenue exceeds $\ln k_i - \ln a_i$, where $\ln a_i$ captures doctor i's altruism.

The net revenue depends on the insurer's fee and payment propensity. The insurer pays a fee of c_s when the physician treats an insured patient; we can think of both c_s and k_i as an average across the relevant patients, or per visit from these patients. In addition to choosing c_s, the insurer chooses a payment probability p_s. So the physician can ultimately collect $c_s p_s$, while a share $1 - p_s$ of claims will be denied or have payments reduced. In order to

collect this revenue, the physician has to spend a share α of potential revenue on billing and collection. So the physician's net revenue will be $p_s c_s - \alpha c_s = (p_s - \alpha) c_s$.

Physicians compare the net revenue with costs and their altruisim motive. So a rational physician will treat the insurer's patients when

$$\ln k_i - \ln a_i < \ln(p_s - \alpha) + \ln c_s. \tag{3}$$

We assume that physicians' net costs $(\ln k_i - \ln a_i)$ follows an arbitrary distribution $F(\cdot)$, and that this distribution is the same across states. Given this assumption, equation (3) tells us that the share σ of doctors who will treat Medicaid patients is:

$$\sigma = F\big(\ln(p_s - \alpha) + \ln c_s \big). \tag{4}$$

This share is naturally increasing in fees c and payment probability p. It is decreasing in the costs of collecting revenue, α. The data, described in the following section, allow us to observe σ, p_s, and c_s. But we can't observe the collection costs α or underlying distribution of physician costs, $F(\cdot)$. But equation (4) shows how we can nevertheless estimate α. To see this, consider the following regression of a doctor's acceptance decision on $\ln p$ and $\ln c$:

$$\text{accept}_i = \beta \ln c_s + \gamma \ln p_s + \varepsilon_i \tag{5}$$

Equation (4) tells us that the estimates of $\hat{\beta}$ and $\hat{\gamma}$ will be:

$$\hat{\beta} = \frac{\partial \sigma}{\partial \ln c_s} = f(\cdot) \qquad\qquad \hat{\gamma} = \frac{\partial \sigma}{\partial \ln p_s} = \frac{\overline{p}}{\overline{p} - \alpha} f(\cdot) \tag{6}$$

where $f(\cdot)$ is evaluated at the argument shown in (4) and $\overline{p}$ is the average payment probability.

Combining these two estimates yields $\frac{\hat{\gamma}}{\hat{\beta}} = \frac{\bar{p}}{\bar{p} - \alpha}$. We can solve this for α:

$$\alpha = \left[1 - \left(\frac{\hat{\gamma}}{\hat{\beta}} \right)^{-1} \right] \bar{p}. \tag{7}$$

So an estimate of regression (5), combined with summary statistics, yields the parameters that we need to back out collection costs α.

The variation that underpins β and γ comes from doctors' decisions of whether to treat Medicaid patients. These decisions should reflect both variable costs of submitting Medicaid bills and managing payment uncertainty, and fixed costs of dealing with Medicaid. So our estimated $\hat{\alpha}$ should provide a comprehensive measure of the costs of billing Medicaid, in contrast with the model estimates from section 6, which will only capture variable costs.

We use this framework to estimate and understand the supply of care to Medicaid patients. We focus on Medicaid for three reasons. First, Medicaid has dramatically higher billing hassle than other insurers, as shown above and in Gottlieb, Shapiro and Dunn (2018). Second, Medicaid has the fewest physicians willing to treat its beneficiaries. This makes it a very relevant setting for policy analysis, and implies the most empirical variation for us to study. Finally, Medicaid is managed by state governments, who make very different decisions in different states and often do so for political reasons (Layton et al., 2019). This provides us with ample variation that physicians can't control. In commercial markets, in contrast, it would be natural to think of administrative hassle as an equilibrium outcome of patient demand and physician supply.

The key assumption implicit in equation (6) is that $F(\cdot)$ is the same across states. If physicians have sysetmatically different cost/altruism distributions in states with different fees or payment probabilities, the estimates in (6) would each have an extra term involving the covariance of states' density $f(\cdot)$ with c_s or p_s. To address this possibility, we introduce two empirical strategies in the following two sections. One controls for individual physician fixed effects, which should eliminate any differences in altruism. The other controls for the

physician group in question, which should eliminate any productivity differences.

4   Does Payment Uncertainty Matter? Evidence from Movers

We first provide a descriptive look at physicians' supply decisions through purely obser-
vational regressions. This first analysis uses the full catalog of physicians in the SK&A
data—nearly every practicing physician in the U.S. We look at the relationship between
each physician's reported willingness to treat Medicaid patients, on the left hand side, and
her state's Medicaid billing hassle and reimbursement rates, on the right hand side. For
numerous reasons, the resulting estimates need not be causal; for example, physicians who
want to treat Medicaid patients may differ from others, or they may select into states with
different Medicaid policies.

We use two empirical strategies to address these concerns. Our first strategy uses a
physician movers design to address concerns about physician-level characteristics, such as
unobservable desire to treat Medicaid patients. In section 5 we introduce a second strat-
egy, using physicians in groups that span state boundaries. By controlling for group fixed
effects, we eliminate variation due to practice characteristics, such as investment in billing
technology, other aspects of billing skill, the group's experience with a particular part of the
market, or altruism.

4.1   Empirical Strategy: Movers

Following Molitor (2018), who uses physician movers, and other mover designs in labor and
health economics (Abowd et al., 1999; Finkelstein et al., 2016, 2019; Hull, 2018), we examine
the impact of a physician's move between states with different payment rates and billing
difficulty. Under the usual assumption that the timing of a physician's cross-state move is
independent of other shocks affecting her willingness to treat Medicaid patients, this strategy
estimates the supply curve while controling for time-invariant physician unobservables.

For each mover, we use data starting up to 4 years prior to the move, through 4 years

after the move, and estimate the following regression at the physician-year level:

$$\text{Medicaid acceptance}_{i,t} = \alpha + \beta \Delta \text{Fee}_i \times \text{Post-Move}_t$$

$$+ \gamma \Delta \text{Recovery Rate}_i \times \text{Post-Move}_t$$

$$+ \phi_i \cdot \mathbb{1}_i + \vartheta \text{ Various controls}_{i,t} + \nu_{i,t} \qquad (8)$$

The dependent variable is the same as in regression (10), and the critical controls here are individual physician fixed effects , ϕ_i. This strategy identifies the supply parameters β and γ exclusively based on physicians who move. The key moment is the difference in those physicians' pre-move Medicaid acceptance and their post-move Medicaid acceptance, and how that difference varies with differences in the states' policies.

Both of these methods address unobservable differences on the supply side of the market, i.e. among physicians and their practices. Other differences among states could confound the estimates of β and γ if they are correlated with Medicaid fees and recovery rate. To address this concern, we study selection on observables using Oster's (2019) assumption of proportional selection. We report the variants of equations (10) and (8) starting from raw correlations without controls, and then adding the fixed effects, other controls, and additional local characteristics intended to capture other demand factors. We use the software provided by Oster (2019)[13] to evaluate the changes in coefficients and predictive power observed as we add controls, and determine the likely extent of bias remaining in our estimates.

To visualize the time trends in these results, we begin by estimating a dynamic event

---

[13]Available online at https://emilyoster.net/s/psacalc_0.zip

study version of equation (8), namely:

$$\text{Medicaid acceptance}_{i,t} = \alpha + \sum_{\tau \neq -1} \beta_\tau \Delta \text{Fee}_i \times \text{Year}_\tau$$

$$+ \sum_{\tau \neq -1} \gamma_\tau \Delta \text{Recovery Rate}_i \times \text{Year}_\tau$$

$$+ \phi_i \cdot \mathbb{1}_i + \vartheta \text{ Various controls}_{i,t} + \nu_{i,t} \qquad (9)$$

where τ denotes the year relative to that in which the physician moved.

4.2   Results: Movers

Figure 4 shows the results from the movers strategy. The coefficients show the effect of a one log point increase in Medicaid fees (first panel panel) or recovery rates (right panel) on the probability that a physician accepts Medicaid patients. Both panels show no evidence of any change in behavior prior to the moves. The top panel shows that, immediately following the move, Medicaid acceptance increases with an elasticity of around 0.07 with respect to fees. This effect increases over the following years, stabilizing at around 0.15 in years 3 and 4. The bottom panel shows analogous results for recovery rate. We again see no pre-trends, and after the move we estimate elasticities of 0.12 to 0.2 with respect to recovery rates.

Figure 5 shows the raw data behind this event study. We group physicians based on terciles of changes in the key independent variables—state-level Medicaid fees and recovery rates—from before to after the move. The graph simply plots the difference in Medicaid acceptance rates for physician movers—the share who treat Medicaid patients after moving minus the share doing so before moving—depending on which terciles the physician is in. It shows that our key results show up starkly in the raw data: physicians who move to better paying states, and to states with smoother reimbursement, become more likely to treat Medicaid patients.

Table 3 summarizes these results in a regression that combines the post-move years into a

single dummy variable, using the specification shown in equation (8). The first column omits the physician fixed effects, so captures the raw relationship between Medicaid acceptance, on the left hand side, and state Medicaid characteristics on the right hand side. The second column adds the physician fixed effects and provides baseline results. It shows that a 10 percent increase in the recovery rate leads to approximately a 1.67 percentage point increase in physicians' willingness to accept Medicaid patients. A 10 percent increase in fees leads to a 1.08 percentage point increase in willingness to accept Medicaid patients.

To interpret these results, we return to the framework from section 3. Equation (7) shows that the cost of collecting payment, as a share of the claim value, is $\alpha = \left[1 - \left(\hat{\gamma}/\hat{\beta}\right)^{-1}\right]\overline{p}$. Table 3 shows that $\hat{\gamma}/\hat{\beta} = 1.72$ in our baseline estimate. With an average recovery rate of $\overline{p} = 0.63$, we obtain $\alpha = 0.26$, implying that billing and collections absorb 26 percent of Medicaid fees.

This number encompasses a broad range of costs. We have estimated it based on the extensive margin decision of whether to accept Medicaid patients, not an intensive margin decision of how much care to provide per patient, or even how many patients to treat. So this cost estimate includes the full range of administrative costs that come along with treating Medicaid patients, to the extent of their correlation with our recovery rate index. This can include costs of registering with the state Medicaid office, learning Medicaid's billing rules, and the actual billing itself. Our billing model in section 6 below captures the narrower costs of billing and follow-up, while the estimate here is broader.

The remaining estimates in Table 3 add a wide range of controls to gauge the stability of the results. Column 3 controls for the log number of physicians in the physician's practice, and in the physician's practice in the same county as the index physician. Column 4 controls for local demographics, including race, age, sex, and education. Column 5 adds controls for the local insurance market description, including the share with Medicare, Medicaid eligible, uninsured, VA insurance, dual insured, and a Medicare measure of local population health risk. Column 6 adds economic characteristics, including household median income, poverty,

unemployment, and rent. Column 7 adds population and population density, and the final column controls for physicians per capita. Despite this plethora of controls, the coefficients on state Medicaid fee and recovery rate indices remain stable and highly significant. This strongly suggests that these characteristics of a state, rather than something else that happens to be correlated with them, drive the changes we see in physicians' decisions of whether to treat Medicaid patients.

Oster's (2019) extension of Altonji et al. (2005) allows us to quantify the extent of unobservable bias that would be necessary to explain away our results. The rows shown under the R^2 in the table, labeled as δ following Oster's notation, show values of 1.9 and 4.3 once we have incorporated the full set of controls. This means that unobservable bias would need to be 1.9 times as large as the effect of observed controls (for the fee index), and 4.3 times as large (for the recovery rate index) in order to eliminate the estimated effect. These are well above Oster's suggested threshold of $\delta = 1$, suggesting that our results are likely robust to observed and unobserved controls.

We next show the robustness of these results' along other dimensions. In particular, we vary the construction of our state indices to account for a range of potential concerns. First, we may be concerned that patients in different states have different demographic characteristics which make the billing process more or less difficult. Figure 6 Panels A and B show that the results are similar when we construct the indices conditional on local income.

One might worry that hassle is especially high when a physician first sees a new patient. There may be particular difficulties in determing the patient's insurance status, eligibility for Medicaid, or ensuring that the basic details are consistent with the insurer's database. We address this by constructing versions of the indices that exclude the first visit we observe between each patient and any given physician. The results using these indices are shown in Panels C and D of Figure 6, and look virtually identical to the baseline results from Figure 4. We use the complementary subsample in Panels E and F—only the first visit for each physician and patient. In this case, the results for Medicaid recovery rate become less precise,

presumably because of the smaller sample when we restrict to one visit per physician-patient.

Figure 7 shows one final aspect of robustness. We use three different approaches to construct the counterfactual payment c_i used in the denomiantor of p_i (defined on page 13). These approaches are explained in Appendix A.1, but ultimately make little difference to the results. As Figure 7 shows, the results are qualitatively and quantitatively similar regardless of how exactly we construct c_i. Appendix Tables B.1 and B.2 show the full robustness table and Oster (2019) calculation for these alternative indices. Appendix Tables B.3 through B.5 further show robustness to different indices, when also varying the method for imputing c_i. We include the same adjustments considered in Figure 6, an index incorporating a shrinkage adjustment to account for estimation error, one limiting the sample to pregnant patients, since that could generate a more similar population across states, and indices estimated without physician fixed effects. In all cases, the quantitative results remain very similar to the baseline results from Table 3.

We next examine the roles of traditional state-run Medicaid and Medicaid Managed Care programs. In recent decades, states have outsourced increasing shares of their Medicaid programs to private insurers to cover beneficiaries through private insurance plans (Congressional Budget Office, 2018). These plans, called managed care organizations (MCOs), generally receive a lump-sum payment from the state and then are responsible for managing patients' care and paying for most or all treatments. Since they are managed separately, and could be run quite differently from states' direct fee-for-service (FFS) plans, we estimate separate recovery indices for MMC and FFS, and investigate how they both impact physicians' Medicaid acceptance decisions.[14]

Our data do not separately report willingness to treat Medicaid MCO and FFS patients, but only a summary willingness to treat Medicaid patients, however physicians interpret the question. To examine how both indices impact physicians' decisions, we include changes in both MMC and FFS indices in the regression, interacting both with the state's share of

---

[14]We do not estimate separate fee indices since MCO and FFS plans pay very similar rates; the same is true for traditional Medicare and privately run Medicare Advantage plans (Trish et al., 2017).

Medicaid patients in that type of insurance (so the MCO share plus the FFS share add up to 1). We standardize these shares so the direct effects of the MCO and FFS indices can be interpreted as effects for the mean state.

Table 4 shows the results. The coefficients on Medicaid fees remain very similar to those in Table 3. The coefficients on FFS and MCO recovery rates are both positive and significant, as is the interaction between the MCO recovery rate and MCO share. The FFS recover rate coefficient is similar to the Medicaid fee index coefficient, while the MCO recovery rate coefficient is about 50 percent larger. As we add additional controls, the MCO coefficient at the mean decreases, but is offset by an increase in the interaction between MCO recovery rate and MCO share. In summary, the FFS recovery rate clearly matters for the mean state, while the MCO recovery rate becomes increasingly important as the share of MCO patients grows.

Together, these results tell a clear story that physician supply curves are multi-dimensional: physicians respond to the fees, but they also react to payment difficulty. In states where Medicaid pays higher fees, physicians are more willing to treat Medicaid patients. In states where they have more difficulty collecting revenue, they are less willing to treat Medicaid patients. This latter force is larger than the direct effect of fees, most likely because of the extra costs associated with trying to collect revenue when Medicaid is reluctant to pay. This implies that insurers and policymakers have two instruments that affect physicians' supply decisions. As states increase the share of Medicaid patients whose care is managed by private firms, these firms' decisions become increasingly important relative to the state's own administrative and payment rate choices.

5 Does Payment Uncertainty Matter? Cross-State Groups

5.1 Empirical Strategy: Cross-State Groups

The movers strategy we used in the previous section is appropriate if the individual physician mover chooses whether to see Medicaid patients and bears the costs of those choices. But

physicians increasingly operate in larger practices or major medical systems, which may have centralized billing and could make centralized dercisions about which insurance to accept. To account for differences in groups' decision-making or billing efficiency, we introduce a second identification strategy to estimate states' impacts of Medicaid acceptance.

This strategy uses physician groups that span state boundaries. The idea is to identify the impact of state policies based on differences in acceptance decisions among physicians within the same practice, but facing different state policies. This eliminates any differences due to the practice, such as its skill at dealing with billing complexity, its history with a particular part of the market, or its altruism.

Using this sample, we introduce practice group fixed effects into a physician-level regression of Medicaid acceptance on state characteristics:

$$
\begin{aligned}
\text{Medicaid acceptance}_{i,t} = \alpha &+ \beta \, \text{State Medicaid fee}_{s(i),t} \\
&+ \gamma \, \text{State Medicaid difficulty}_{s(i),t} \\
&+ \theta_g \cdot \mathbb{1}_{g(i)} + \vartheta \, \text{Various controls}_{i,t} + \epsilon_{i,t}
\end{aligned}
\tag{10}
$$

The dependent variable is a binary indicator for whether the physician reports accepting Medicaid patients. The state-level fee and difficulty indices are described in section 2.2 above. The key controls are fixed effects for each physician group or system, $\mathbb{1}_{g(i)}$. Given these fixed effects, we identify β and γ off of differences in Medicaid acceptance among physicians within the same practice.

This strategy's limitation is that it controls for unobservables at the group level but not for the individual physician. Even within a group, physicians with a stronger preference for treating Medicaid patients could sort across states in ways correlated with their Medicaid policies.

We produce estimates using three different versions of the fixed effects, $\mathbb{1}_{g(i)}$, which Appendix A.2 describes in detail. The first two definitions rely on the SK&A data described

in section 1.2. SK&A provides a physician identifier and information to group these physicians at 2 levels. The first one links physician practice locations that are "part of a medical group." We refer to this as the physician's group. The second is the "health system... owning or managing the site." This is broader than the specific group and captures the fact that groups may be linked through a broader system. The third definition is the tax identification number reported in Medicare's MD-PPAS data.

5.2 Results: Cross-State Groups

Figure 8 shows graphical estimates from this strategy. In each panel, the y axis value represents the residual from regressing each physician's Medicaid acceptance decision on a group-level fixed effects. The x axis value shows the residual from regressing the relevant state index (fee or recovery rate) on the same fixed effects. In other words, both the x and y axes show the physicians' within-group difference, relative to a group-level mean, in the variables listed. We then make a binned scatterplot by grouping these residuals into 15 bins based on their x axis values, and ploting the x and y means within each bin. In all cases we see clear positive slopes, indicating that within-group Medicaid acceptance decisions respond positively to the specific state's fee and recovery rates.

Table 5 shows the numerical estimates of equation (10). We estimate a semi-elasticity of Medicaid acceptance with respect to fees of around 0.1, meaning that ten percent higher fees lead to a 1 percentage point increase in the share of physicians accepting Medicaid patients. The semi-elasticity with respect to the recovery rate is around 0.2, so a ten percent increase in recovery rate increases Medicaid acceptance by about 2 percentage points. The magnitudes of these estimates are very similar to those from the movers strategy. Appendix Tables B.6 and B.7 shows similar results using an alternative definition of physician groups, and with indices constructed on slightly different samples.

The ratio of coefficients on recovery rate relative to fee index $\hat{\gamma}/\hat{\beta}$ is around 2, also similar to that from the movers strategy. So the two approaches tell a consistent story, both

qualitatively and quantitatively. Billing hassle is a major deterrent to doctors' willingness to treat Medicaid patients. It is quantitatively at least as important as fees.

6 Empirical Model of the Billing Process

Our estimates of physicians' supply responses to billing hassle reveal that total costs of dealing with Medicaid billing amount to around one-quarter of Medicaid revenue. But the analysis so far cannot distinguish between billing costs that result from payment uncertainty at the level of individual visits, as opposed to fixed costs or other overhead that is independent of any specific visit. The final part of this paper aims to distinguish between these types of costs.

To do this, we model the costs of engaging in the Medicaid billing process at the level of each visit. We use the richness of our data to describe the process of claim submission, denial, and potential resubmission. The model captures the physicians' costs and effort with each claim, to determine how much it costs to deal with payment uncertainty.

The basic idea of the model is that physicians' behavior in the billing process allows us to infer the costs of that process. We assume that the physician treats a patient, and the practice subsequently makes (privately) optimal decisions about how much effort to put into the billing process. This effort includes an initial claim submission and, if the claim is denied, subsequent resubmissions. By observing the biller's decisions in this process, we can infer how much these steps cost.[15]

Concretely, consider a claim that a practice submitted and the insurer denied. The practice can choose to resubmit the claim, but doing so would entail various costs. It will take staff time to review the remittance and determine the reason for the denial. Depending on the insurer's explanation, the practice may need to collect additional documentation about the patient's medical situation, the treating physician may need to update and expand

---

[15]The model presented here focuses only on the provider's decisions. While private insurers may behave strategically in their decisions, it is less clear that we should expect Medicaid to do so. Given the centrality of these issues for Medicaid billing, we focus on that market and analyze only the provider's strategic decisions.

her medidcal notes, or the back office may need to contact the patient to verify billing information. All of this takes real resources.

When deciding whether to resubmit, an optimizing practice would weigh these resources against the expected increase in payment—equal to the probability of a successful resubmission times the value of the claim if successful. Our data allow us to measure and estimate these objects. Since we see many claims that a particular physician submits to the same insurer for similar patients, with similar treatments, we can estimate the probability that a resubmission will be fruitful. We also know the claim's potential value, as explained in section 2.2. Since we know the expected net revenues from resubmitting a denied claim, the physician's decision of whether to resubmit allows us to infer the costs of doing so.

A similar logic allows us to infer the cost of effort put into the initial claim submission. We model the initial submission as a continuous effort choice by the physician practice, where more costly effort increases the probability that the first claim is paid. The value of that initial payment is the difference between receiving payment initially and the net gain from the option to resubmit denied claims, accounting for resubmission costs. Having already estimated the resubmission costs, backwards induction allows us to compute the expected value of a denied claim, accounting for the option value of resubmission. The difference between the value of a paid claim and the expected value of denied claims is the net benefit of having a claim paid in the initial submission. Optimizing physicians will put more effort into the initial submission for those claims with a higher net benefit from payment success. By comparing our estimates of this net benefit with the observed probability of successful payment, we can infer the costs of physicians' effort in the initial submission.

The reminder of this section enriches and formalizes this model. We add some additional features to this basic idea in order to match patterns that we see in the data. Most notably, a claim can contain multiple line items, some of which are approved while others are denied. Our formal model incorporates this possibility. The next subsections formalize the model and estimation procedure, and section 6.6 presents the estimates.

6.1 Observables

We observe visits, denoted by i, and each visit can have multiple line items (individual medical services) denoted by ℓ. Observable visit characteristics are collected in the vector $x_i = (diag_i, state_i, pay_i)$. In this vector, $diag_i$ is the diagnostic code, and the rest of the characteristics describe state and payer. For every line item ℓ of visit i we observe $z_{i\ell} = (proc_{i\ell}, C_{i\ell})$, describing the procedure code and the contractual value of the service.

The main dependent variables of interest are, for every pair $i\ell$, a sequence of outcomes

$$O_{i\ell} = \{Y_{i\ell}^0, R_{i\ell}^1, Y_{i\ell}^1, R_{i\ell}^2, ...\}. \tag{11}$$

If the line item ℓ for visit i is paid after the first claim is submitted, we observe $Y_{i\ell}^0 = 1$, otherwise $Y_{i\ell}^0 = 0$. If $Y_{i\ell}^0 = 0$ and the provider does not resubmit a second claim containing line ℓ for visit i, $R_{i\ell}^1 = 0$. If the provider does choose to resubmit, then $R_{i\ell}^1 = 1$. If $R_{i\ell}^1 = 1$, and the line item ℓ for visit i is paid after the submission of the second claim, $Y_{i\ell}^1 = 1$. If the insurer again denies payment, then $Y_{i\ell}^1 = 0$, and so on recursively. For visit i, the collection of all payment and resubmission decisions is collected in $\mathcal{O}_i = \{O_{i\ell}\}$.

6.2 Cost Primitives and Optimal Physician Behavior

The model begins after the patient's visit is complete. When submitting the first claim the provider chooses the billing effort level optimally, to maximize expected payoff. Higher effort implies that it is less likely to enter the costly process of claim denials and resubmissions.

Let p be the probability of the claim being paid for all line items without any issue. The provider can choose an effort level $\mathcal{E}_i(p)$, expressed in monetary terms, to achieve a particular p. The expected payoff for visit i, expressed as a function of p is then

$$U_i(p) = -\mathcal{E}_i(p) + p \sum_\ell C_{i\ell} + (1-p)\,\mathbb{E}\mathbb{D}_i. \tag{12}$$

In the case where the effort does not pay off with a successful initial claim, the provider and

29

insurer enter the process of denials and possible resubmissions. The term $\mathbb{ED}_i$ captures the expected net payoff from visit i if this occurs. This is defined as follows.

The cost of resubmitting a claim for visit i is κ_i, and each additional line item requires an additional cost of γ_i. If there are problems with the initial claim, the future payoff is then

$$\mathbb{ED}_i = \mathbb{E}\left[\underbrace{-\sum_{t>1}\left(\gamma_i\sum_\ell R_{i\ell}^t + \kappa_i\max_\ell R_{i\ell}^t\right)}_{\text{future resubmission costs}} + \underbrace{\sum_{t>1}\sum_\ell Y_{i\ell}^t C_{i\ell}}_{\text{future revenues}}\right]. \tag{13}$$

The first term inside the expectation corresponds to the costs from resubmitting future claims; this depends on the parameters (γ_i, κ_i). The second term is the expected revenue. The expectation is taken with respect to the random variables $\{R_{i\ell}^t, Y_{i\ell}^t\}_{i\ell}^t$. Importantly, some of these variables are endogenous, and determined by the provider's optimal behavior in the "continuation" problem faced after the claim incurs one or more denials.

We analyze provider behavior by backward induction, starting from the resubmission decisions. When receiving a claim with denials (partial or full), the provider can decide to resubmit any subset of the denied line items. Let D_i be the subset of line items of visit i for which payment is not granted after the first claim submission. Formally, $D_i = \{\ell : Y_{i\ell}^0 = 0\}$. The provider chooses any subset $A \subset D_i$, including possibly the empty set, to maximize

$$\mathbb{E}\left[-\sum_{t>1}\left(\gamma_i\sum_{\ell\in D_i} R_{i\ell}^t + \kappa_i\max_\ell R_{i\ell}^t\right) + \sum_{t>1}\sum_{\ell\in D_i} Y_{i\ell}^t C_{i\ell}\,\middle|\, R_{i\ell}^1 = 1\ \forall\ell\in A, R_{i\ell}^1 = 0\,\forall\ \ell\notin A\right].$$

$$\tag{14}$$

The term $\mathbb{ED}_i$ is defined by the optimal behavior in this dynamic-discrete-choice problem faced by the provider after the claim is denied, in full or partially. We assume that the provider expects to behave optimally in the future under rational expectations about the data generating process.

While we explicitly model provider behavior, we treat the payer's behavior as an ex-

ogenous component of the data-generating process. That is, providers take the insurer's behavior as given, conditional on the physician's effort and resubmission decisions. This is a reasonable approximation since the insurer processes a large number of claims from many providers, and has likely adopted standard procedures and strategies. These lead to a probability that payment for any particular line item is declined for a given visit which is likely indepedent of the provider's strategy. That said, in the future we aim to relax this simplification and augment the model to analyze the insurer's behavior.

In the first period, after the visit takes place, the provider behavior is described by the first-order optimality condition for the maximization of (12). This is

$$-\frac{d\mathcal{E}_i}{dp} + \sum_\ell C_{i\ell} - \mathbb{E}\mathbb{D}_i = 0. \tag{15}$$

If the expected payoff of the continuation game ($\mathbb{E}\mathbb{D}_i$) were known, this expression leads to a simple differential equation

$$\frac{d\mathcal{E}_i}{dp} = \sum_\ell C_{i\ell} - \mathbb{E}\mathbb{D}_i. \tag{16}$$

If the right side of (16) is known over an interval $[p_1, p_2]$, one can learn the incremental cost to increase the probability of full, prompt claim payment from p_1 to p_2 as

$$\mathcal{E}_i(p_2) - \mathcal{E}_i(p_1) = \int_{p_1}^{p_2} \frac{d\mathcal{E}_i(s)}{dp} ds. \tag{17}$$

In what follows we parametrize the resubmission and initial effort costs and estimate the model primitives.

6.3 Parametrization

For each visit i, we let effort cost functions vary with diagnosis, state, and total potential value of the procedures (abusing notation slightly, let $C_i = \sum_\ell C_{i\ell}$):

$$\mathcal{E}_i(p) = \mathcal{E}(p; diag_i, state_i, C_i). \tag{18}$$

More specifically, we set the derivative as

$$\frac{d\mathcal{E}_i(p)}{dp} = \phi(p, diag_i, state_i, C_i) \tag{19}$$

for some unknown function $\phi(\cdot)$.

In terms of the resubmission costs, we let

$$\kappa_i = \kappa(state_i) + \varepsilon_i^\kappa \tag{20}$$

$$\gamma_i = \gamma + \varepsilon_{i\ell}^\gamma. \tag{21}$$

The errors terms ε_i^κ and $\varepsilon_{i\ell}^\gamma$ are iid mean zero draws. For estimation, we let these follow a Type 1 extreme value distribution.[16]

6.4 Estimation

The first step of our estimation procedure focuses on the resubmission decisions for a given visit i after one or more line items are denied. We apply the conditional-choice-probabilities (CCP) approach introduced by Hotz and Miller (1993). Importantly, we do not need to solve for the value function numerically, because one of the available options is a "stopping option" that ends the dynamic process and leads to a deterministic outcome. The provider can decide to not resubmit any claim for the visit, and give up on recovering revenues.

---

[16]The parametrization of the resubmission costs allows for much less flexibility than the initial effort functions. This can be relaxed in the future.

In our notation this would be the event

$$R_{i\ell}^1 = 0 \ \forall \ell \in D_i. \tag{22}$$

For each procedure *proc*, we estimate the following quantity: *probability that the procedure is not paid, conditional on state, payer, and diagnostic code.* Formally, we estimate:

$$g(proc, diag, state, pay) = \Pr[Y^t = 0 | Y^{t-1} = 0, R^t = 1, diag, state, pay]. \tag{23}$$

Assuming that, after conditioning on diagnosis, payer, state, and initial billing effort, the noise components in payments of different procedures are statistically independent, for any set of procedures A and subset $B \subset A$ we compute

$$G(B, diag, state, pay | A) = \prod_{proc \in B} g(proc | diag, state, pay) \prod_{proc \in A \cap B^c} (1 - g(proc | diag, state, pay)). \tag{24}$$

Under our assumption, this is the probability that all procedures in B are not paid if the set of procedures A is resubmitted.

For every set of procedures A observed in the data for visits with characteristics $(diag, state, pay)$, we compute the probability that after (re)submitting the procedures in A, and given that a subset $B \subset A$ of these line items is denied, there is no further claim for the visit.

To reduce the dimensionality, we only keep track of two summary statistics of the sets A and B. First, $n(A) = \dim(A)$ and $n(B) = \dim(B)$ represent, respectively, the number of line items in the A and B, with $n(B) \leq n(A)$. Second, $\mathcal{C}(A) = \sum_{proc_{i\ell} \in A} C_{i\ell}$, and similarly $\mathcal{C}(B) = \sum_{proc_{i\ell} \in B} C_{i\ell}$. In words, $\mathcal{C}(A)$ is the total value for the procedures in A.[17]

---

[17]We discretize the space of amounts considering bins width width of \$25. In what follows, all references to amounts $\mathcal{C}(A)$ for any set of procedure (line items) A are intended as the midpoint amount of the corresponding bin.

We compute the following quantity:

$$Q(n', C', n, C, diag, state, pay) = \Pr\left[R_{i\ell}^t = 0 \ \forall proc_{i\ell} \in B \big| \Omega(n', C', n, C, diag, state, pay)\right],$$

(25)

where the event $\Omega(n', C', n, C, diag, state, pay)$ is defined as follows (omitting $diag$, $state$, and pay in the definition):

$$\Omega(n', C', n, C, diag, state, pay) =$$
$$\left\{ n(A) = n, \mathcal{C}(A) = C, n(B) = n', \mathcal{C}(B) = C', R_{i\ell}^{t-1} = 1 \right.$$
$$\left. \forall proc_{i\ell} \in A, \text{ and } Y_{i\ell}^{t-1} = 0 \ \forall proc_{i\ell} \in B \right\}$$

(26)

In words: $Q(n', C', n, C, diag, state, pay)$ is the probability that the provider does not submit any further claim if, after resubmitting a set of $n(A) = n$ procedures for a total amount in bin $\mathcal{C}(A) = C$, a subset of $n(B) = n'$ procedures for a total amount in bin $\mathcal{C}(B) = C'$ is denied.

The dataset we use to set up and maximize the likelihood has for observation a unique combination:

$$[\text{visit-ID, subset of line items for which payment is denied in the first claim}],$$

and for simplicity we denote an observation with

$$[visitID_i, A_i],$$

where for every visit with at least one denial, we have one observation for every possible subset A_i of line items that have been denied payment in the first claim.

For each observation, indexed by iA we need the following variables:

- $Choice_{iA} \in \{0, 1\}$: this is equal to 1 if the subset A_i is resubmitted in the second claim;

- $n(A_i)$ as defined above;

- $V(A_i)$ which, following Hotz and Miller (1993), is defined as:

$$V(A_i) =$$

$$\sum_{B \subset A_i} G(B, diag_i, state_i, pay_i | A) \cdot$$

$$(\mathcal{C}(A) - \mathcal{C}(B) - \ln (Q(n(B), \mathcal{C}(B), n(A_i), \mathcal{C}(A_i), diag_i, state_i, pay_i))) ,$$

where the sum is taken over all possible subsets of A_i;

- $(diag_i, state_i, pay_i)$, as well as additional covariates about the provider.

We also include the empty set as possible choice, for which we have $n(A_i) = 0$ and $V(A_i) = 0$. For this case, $Choice_{iA} = 1$ if no second claim for visit i is observed.

With the quantities defined above, we write the following likelihood:

$$\mathcal{L}(Choice_{iA} = 1 | \mathcal{O}_i) = \frac{\exp \left[-\kappa(state_i) - \gamma n(A) + \beta V(A) + 0.5772\beta \right]}{1 + \sum_B \exp \left[-\kappa(state_i) - \gamma n(B) + \beta V(B) + 0.5772\beta \right]}. \tag{27}$$

We set $\beta = 0.99$, and estimate $\kappa(state_i)$ and γ.

With the estimated parameters we can construct a realized version of $\mathbb{ED}_i$ for every observation in our visit-level dataset. We can therefore construct an empirical counterpart

$$\widehat{\phi}(p; diag_i, state_i, C_i) = \mathbb{E} \left[\sum_\ell C_{i\ell} - \mathbb{ED}_i \Big| diag_i, state_i, C_i, \mathbb{E}[\min_\ell Y_{i\ell}^1] = p \right] . \tag{28}$$

And obtain estimates of $\mathcal{E}_i(p_2) - \mathcal{E}_i(p_1)$ as

$$\widehat{\mathcal{E}}_i(p_2) - \widehat{\mathcal{E}}_i(p_1) = \int_{p_1}^{p_2} \widehat{\phi}(p; diag_i, state_i, C_i) dp. \tag{29}$$

35

6.5 Identification

Just as with estimation, we now discuss how we can identify the model's parameters, in two separate steps. The first step is the identification of the resubmission costs.

To estimate the resubmission costs, we need to know the expected benefits of resubmitting a given claim—that is, the probability of eventual payment times the amount the provider would receive if paid. We assume the provider makes optimal decisions after any denial. Thus the difference between the distributions of these amounts for claims that are resubmitted, and for claims that are not resubmitted, tells us how much resubmission would have cost.

Because of the richness of our data, following the analytical derivation in (Hotz and Miller, 1993) we can construct these emiprical distributions. Intuitively, we see a distribution of outcomes for claims that are resubmitted, conditional on detailed observable characteristics, that allows us to proxy for the future value of every possible action; each action corresponds to the resubmission of a subset of denied procedures, or to the end of the collection process. The difference between the value of feasible actions that are not taken by providers and actions that are instead taken identifies the parameter governing the resubmission costs.

Figure 9 and Table 6 illustrate the variation that drives the resubmission cost estimates. Figure 9 shows the empirical value functions (computed following Hotz and Miller, 1993) for the actions taken by providers and the actions not taken by providers. The difference between these distributions shows that providers choose actions with higher continuation values. Intuitively, the ranking of these different alternatives in the data is key to identifying resubmission costs.

Figure 9 does not control for any of these observables, and could mask unobserved heterogeneity affecting both continuation values and providers' decisions. To address this omission, Table 6 illustrates the relationship between continuation value and probability of making a particular choice, progressively adding controls. Column 1 does not have any controls, so reports a linear summary of Figure 9. Columns 2, 3, and 4 progressively control for diagnosis, state, and even medical group. The relationship remains very stable and even increases in

magnitude as we add controls. This monotone relationship (in column 4, a $10 increase in the continuation value is associated with a 4 percentage point increase in the probability of making that choice) identifies the resubmission costs.

Identification of the remaining cost parameters, governing the initial billing effort after the visit, takes the resubmission costs as given. This implies that we know the expected payoff of the process following an initial denial for the visit. The provider knows these payoffs as well and thus, we assume, makes effort choices accordingly. Under our assumptions, the provider exerts higher effort when billing for visits for which a denial is more costly. This higher effort in turn should lead to a higher probability of successful payment. So we should see higher payment probabilities for more valuable claims.

A main concern is that we are not observing distinct characteristics of a visit that determine together effort and the probability of denial. To mitigate this concern, we use our data to document the identifying variation leading to our parameter estimates, and to verify whether adding a large set of controls weakens the relationship between claim amount and probability of payment in the data.

Table 7 reports the estimated linear relationship between claim amount and probability of successful payment. Even when controlling for medical group, state, specialty, diagnosis, and primary procedure, the estimate coefficient is positive and significant, and does not decrease in magnitude. This relationship, after controlling for a rich set of covariates, identifies initial the effort cost function, assuming providers behave optimally.

6.6 Results

We illustrate three key results from this estimation. The first one corresponds to the resubmission costs conditional on the characteristics of the visit and the decision to resubmit. The second set of results describes the initial effort cost to avoid the payment uncertainty and the denial-resubmission back-and-forth. Lastly we quantify billing costs altogether as the expected per-visit amount, both in absolute terms and as a share of the visit initial

expected payment amount. This is the same parameter we estimated through the strategy of the previous sections.

Table 8 and Figure 10 show these results. We estimate that if the first claim is denied, the provider expects to spend $31.26 on average *if* she continues the game by submitting additional claims. In comparison, the expected revenues from resubmissions have a mean of $41.11, with a median of $77.

Given these continuation values, the model computes the return to pre-submission billing effort. We estimate that the provider has to spend $10.51 on average to increase the probability of full payment of the first claim by 0.1 (from the observed benchmark level).

Aggregating these numbers, the expected billing costs that we can attribute to the payment uncertainty and back-and-forth resubmissions processes following the first claim amount to, on average, almost $24 per visit. Compared to the average revenue of approximately $111, the estimated billing costs are approximately 21% of the initially expected revenue.

Our estimates allow us to compare the expected profit from a Medicaid patient with and without accounting for billing uncertainty and billing costs. The expected net revenues from a visit, accounting for billing effort, denials, and billing costs, can be described as:

$$\text{Expected net revenues} = \underbrace{-\$21}_{\text{initial billing effort}} + \underbrace{0.68 \times 111}_{\text{first submission}} + \underbrace{(1 - 0.68) \times 5.9}_{\text{continuation}} = 56. \tag{30}$$

This suggests that, according to our model, the expected revenue accounting for payment uncertainty and billing cost is 50.5% of the amount that one would estimate by relying on claims alone, and disregarding payment uncertainty and billing costs.

Our preliminary estimates of the model using data on private payers (not shown) suggest that, in the commercial segment of the market, the net expected revenues accounting for billing costs and payment uncertainty is 64% of the initial expected payment. This suggests that analyzing differences in healthcare prices ignoring billing costs and revenue uncertainty

may understate the differences between private payers and Medicaid.

7 Conclusion

This paper examines the economics of how physicians and insurers haggle over payments for care that physicians provide. We find that payment uncertainty is dramatic, especially when submitting bills to Medicaid. This fact motivates us to consider the effect of payment uncertainty on doctors' supply of care to Medicaid patients. We find that their willingness to provide care is at least as responsive to billing difficulty as to the reimbursement rate. This result is robust to two identification strategies.

To quantify the costs of this uncertainty, we use both the supply decisions and our detailed data on the claim submission and subsequent bargaining process. We estimate a dynamic model of this bargaining process and find that doctors spend a substantial share of their revenue trying to collect payment from Medicaid. The model finds that variable costs of billing are 21 percent of Medicaid revenue, while the broad set of costs that influence physician supply decisions imply costs of one-quarter of Medicaid revenue.

Our findings demonstrate the importance of well functioning business operations in the health care setting. Difficulty in payment collection has meaningful impacts on firms' willingness to engage in markets. In the case of a major government health care program, this hassle compounds the effect of low payment rates to push physicians away from providing care for publicly insured patients.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333. 3, 18

Alexander, Diane and Molly Schnell, "The Impacts of Physician Payments on Patient Access, Use, and Health," Working Paper WP-19-23, Northwestern Institute for Policy Research October 2019. 2, 5, 12

Allen, Marshall, "We Asked Prosecutors if Health Insurance Companies Care About Fraud. They Laughed at Us.," *ProPublica*, September 10 2019. Available online at https://www.propublica.org/article/we-asked-prosecutors-if-health-insurance-companies-care-about-fraud-they-laughed-at-us (accessed October 19, 2019). 7

Altonji, Joseph G, Todd E Elder, and Christopher R Taber, "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of political economy*, 2005, *113* (1), 151–184. 22

Backus, Matthew, Thomas Blake, Bradley Larsen, and Steven Tadelis, "Sequential bargaining in the field: Evidence from millions of online bargaining interactions," Working Paper No. 24306, National Bureau of Economic Research 2018. 1

Benzarti, Youssef, "How Taxing Is Tax Filing? Using Revealed Preferences to Estimate Compliance Costs," Working Paper 23903, National Bureau of Economic Research 2017. 5

Besley, Timothy, "Law, Regulation, and the Business Climate: The Nature and Influence of the World Bank Doing Business Project," *Journal of Economic Perspectives*, September 2015, *29* (3), 99–120. 1

_ **and Stephen Coate**, "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs," *American Economic Review*, 1992, *82* (1), 249–261. 5

Bound, John and Richard V. Burkhauser, "Economic analysis of transfer programs targeted on people with disabilities," in "Handbook of Labor Economics," Vol. 3 of *Handbook of Labor Economics*, Elsevier, 1999, chapter 51, pp. 3417 – 3528. 5

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad, "What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics," *The Quarterly Journal of Economics*, 04 2017, *132* (3), 1261–1318. 2, 9

_ **and Mathijs de Vaan**, "Intermediation and Vertical Integration in the Market for Surgeons," Working Paper 2018. Yale University, mimeo. Available online at https://www.ftc.gov/system/files/documents/public_events/1494697/brot-goldbergdevaan.pdf (accessed October 19, 2019). 10

Candon, Molly, Stephen Zuckerman, Douglas Wissoker, Brendan Saloner, Genevieve M. Kenney, Karin Rhodes, and Daniel Polsky, "Declining Medicaid Fees and Primary Care Appointment Availability for New Medicaid Patients," *JAMA Internal Medicine*, 01 2018, *178* (1), 145–146. 2

Cicala, Steve, Ethan MJ Lieber, and Victoria Marone, "Regulating Markups in US Health Insurance," *American Economic Journal: Applied Economics*, 2019, *11* (4), 71–104. 7

Clemens, Jeffrey and Joshua D. Gottlieb, "Do Physicians' Financial Incentives Affect Treatment Patterns and Patient Health?," *American Economic Review*, April 2014, *104* (4), 1320–1349. 5

_ **and** _ , "In the Shadow of a Giant: Medicare's Influence on Private Payment Systems," *Journal of Political Economy*, February 2017, *125* (1), 1–39. 2, 7, 9

_ , _ , **and Tímea Laura Molnár**, "Do Health Insurers Innovate? Evidence From the Anatomy of Physician Payments," *Journal of Health Economics*, September 2017, *55*, 153–167. 7, 10

Congressional Budget Office, "Exploring the Growth of Medicaid Managed Care," 2018. Available online at https://www.cbo.gov/publication/54235 (accessed March 27, 2020). 23

Cooper, Zack, Stuart V Craig, Martin Gaynor, and John Van Reenen, "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured," *The Quarterly Journal of Economics*, 09 2018, *134* (1), 51–107. 2, 7, 9, 10

Cunningham, Peter J. and Ann S. O'Malley, "Do Reimbursement Delays Discourage Medicaid Participation By Physicians?," *Health Affairs*, 2008, *27* (Supplement 1), w17–w28. 5

Currie, Janet, Jeffrey Grogger, Gary Burtless, and Robert F. Schoeni, "Explaining Recent Declines in Food Stamp Program Participation," *Brookings-Wharton Papers on Urban Affairs*, 2001, pp. 203–244. 5

Cutler, David M. and Dan P. Ly, "The (Paper)Work of Medicine: Understanding International Medical Costs," *Journal of Economic Perspectives*, 2011, *25* (2), 3–25. 1

Djankov, Simeon, Edward Glaeser, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "The new comparative economics," *Journal of Comparative Economics*, 2003, *31* (4), 595 – 619. 1

_ , **Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer**, "The law and economics of self-dealing," *Journal of Financial Economics*, 2008, *88* (3), 430 – 465. 1

Dunn, Abe and Adam Hale Shapiro, "Physician competition and the provision of care: evidence from heart attacks," *American Journal of Health Economics*, 2018, *4* (2), 226–261. 5

＿ , **Eli Liebman, Sarah Pack, and Adam Hale Shapiro**, "Medical Care Price Indexes for Patients with Employer-Provided Insurance: Nationally Representative Estimates from MarketScan Data," *Health Services Research*, 2013, *48* (3), 1173–1190. 2, 9

Dynarski, Susan and Judith E. Scott-Clayton, "The Cost of Complexity in Federal Student Aid: Lessons From Optimal Tax Theory and Behavioral Economics," *National Tax Journal*, 2006, *59* (2), 319–56. 5

Einav, Liran, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen, "Selection on Moral Hazard in Health Insurance," *American Economic Review*, February 2013, *103* (1), 178–219. 2, 9

Ericson, Keith M. Marzilli and Amanda Starc, "How product standardization affects choice: Evidence from the Massachusetts Health Insurance Exchange," *Journal of Health Economics*, 2016, *50*, 71 – 85. 2, 9

Finkelstein, Amy and Matthew J Notowidigdo, "Take-Up and Targeting: Experimental Evidence from SNAP," *The Quarterly Journal of Economics*, 05 2019, *134* (3), 1505–1556. 5

＿ , **Matthew Gentzkow, and Heidi L. Williams**, "Place-Based Drivers of Mortality: Evidence from Migration," Working Paper 25975, National Bureau of Economic Research June 2019. 3, 18

＿ , ＿ , **and Heidi Williams**, "Sources of Geographic Variation in Health Care: Evidence From PatientMigration," *The Quarterly Journal of Economics*, 07 2016, *131* (4), 1681–1726. 3, 18

Gottlieb, Joshua D., Adam Hale Shapiro, and Abe Dunn, "The Complexity Of Billing And Paying For Physician Care," *Health Affairs*, April 2018, *37* (4), 619–626. 1, 17

Gowrisankaran, Gautam, Keith A Joiner, and Jianjing Lin, "How Do Hospitals Respond to Payment Incentives?," November 2019. University of Arizona, mimeo. Available online at http://www.u.arizona.edu/~gowrisan/pdf_papers/hospital_payment_incentives.pdf (accssed December 16, 2019). 5

Gruber, Jon, John Kim, and Dina Mayzlin, "Physician Fees and Procedure Intensity: the Case of Cesarean Delivery," *Journal of Health Economics*, August 1999, *18* (4), 473–490. 5

Hotz, V Joseph and Robert A Miller, "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 1993, *60* (3), 497–529. 32, 35, 36, 53

Hull, Peter, "Estimating Treatment Effects in Mover Designs," Technical Report 1804.06721, arXiv.org April 2018. 3, 18

Jindal, Pranav and Peter Newberry, "Price negotiations and bargaining costs," Technical Report, Mimeo, Pennsylvania State University. Available online at https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.709.2075&rep=rep1&type=pdf (accessed Feberuary 24, 2020). April 2015. 1

Larsen, Bradley, "The efficiency of real-world bargaining: Evidence from wholesale used-auto auctions," Working Paper No. 20431, National Bureau of Economic Research 2014. 1

Layton, Timothy J, Nicole Maestas, Daniel Prinz, and Boris Vabson, "Private vs. Public Provision of Social Insurance: Evidence from Medicaid," Working Paper 26042, National Bureau of Economic Research July 2019. 2, 17

Liebman, Eli, "Bargaining in Markets with Exclusion: An Analysis of Health Insurance Networks," Technical Report, University of Georgia 2018. University of Georgia, mimeo. Available online at https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxlbGlibGllYm1hbnxneDoxMzdlNjg2ZWRhMWVlNzM3 (accessed October 19, 2019). 2, 9

Long, Sharon K., "Physicians May Need More Than Higher Reimbursements To Expand Medicaid Participation: Findings From Washington State," *Health Affairs*, 2013, *32* (9), 1560–1567. PMID: 24019360. 5

Ly, Dan P and Sherry A Glied, "The impact of managed care contracting on physicians," *Journal of General Internal Medicine*, 2014, *29* (1), 237–242. 5

Molitor, David, "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration," *American Economic Journal: Economic Policy*, February 2018, *10* (1), 326–56. 3, 18

Nichols, Albert L. and Richard J. Zeckhauser, "Targeting Transfers through Restrictions on Recipients," *American Economic Review*, 1982, *72* (2), 372–377. 5

Nichols, Donald A., Eugene Smolensky, and Thorwald N. Tideman, "Discrimination by Waiting Time in Merit Goods," *American Economic Review*, 1971, *61* (3), 312–323. 5

Oostrom, Tamar, Liran Einav, and Amy Finkelstein, "Outpatient office wait times and quality of care for Medicaid patients," *Health Affairs*, 2017, *36* (5), 826–832. 2

Oster, Emily, "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204. 19

Panhans, Matthew, "Adverse Selection in ACA Exchange Markets: Evidence from Colorado," *American Economic Journal: Applied Economics*, April 2019, *11* (2), 1–36. 2, 9

Polsky, Daniel, Michael Richards, Simon Basseyn, Douglas Wissoker, Genevieve M. Kenney, Stephen Zuckerman, and Karin V. Rhodes, "Appointment Availability after Increases in Medicaid Payments for Primary Care," *New England Journal of Medicine*, 2015, *372* (6), 537–545. 2

Reinhardt, Uwe E., "The Pricing Of U.S. Hospital Services: Chaos Behind A Veil Of Secrecy," *Health Affairs*, 2006, *25* (1), 57–69. PMID: 16403745. 7

Trish, Erin, Paul Ginsburg, Laura Gascue, and Geoffrey Joyce, "Physician reimbursement in Medicare Advantage compared with traditional Medicare and commercial health insurance," *JAMA internal medicine*, 2017, *177* (9), 1287–1295. 23

Tseng, Phillip, Robert S. Kaplan, Barak D. Richman, Mahek A. Shah, and Kevin A. Schulman, "Administrative Costs Associated With Physician Billing and Insurance-Related Activities at an Academic Health Care System," *JAMA*, 02 2018, *319* (7), 691–697. 4

Zwick, Eric, "The Costs of Corporate Tax Complexity," Working Paper 24382, National Bureau of Economic Research 2018. 5

Figure 1: State-Level Medicaid Fees and Recovery Rate

Panel A: Medicaid fee index by state

Legend:
- 3.80 – 3.82
- 3.76 – 3.80
- 3.75 – 3.76
- 3.71 – 3.75
- 3.69 – 3.71
- 3.67 – 3.69
- 3.62 – 3.67
- 3.58 – 3.62
- 3.51 – 3.58
- 3.44 – 3.51
- 3.32 – 3.44
- 3.06 – 3.32
- No data
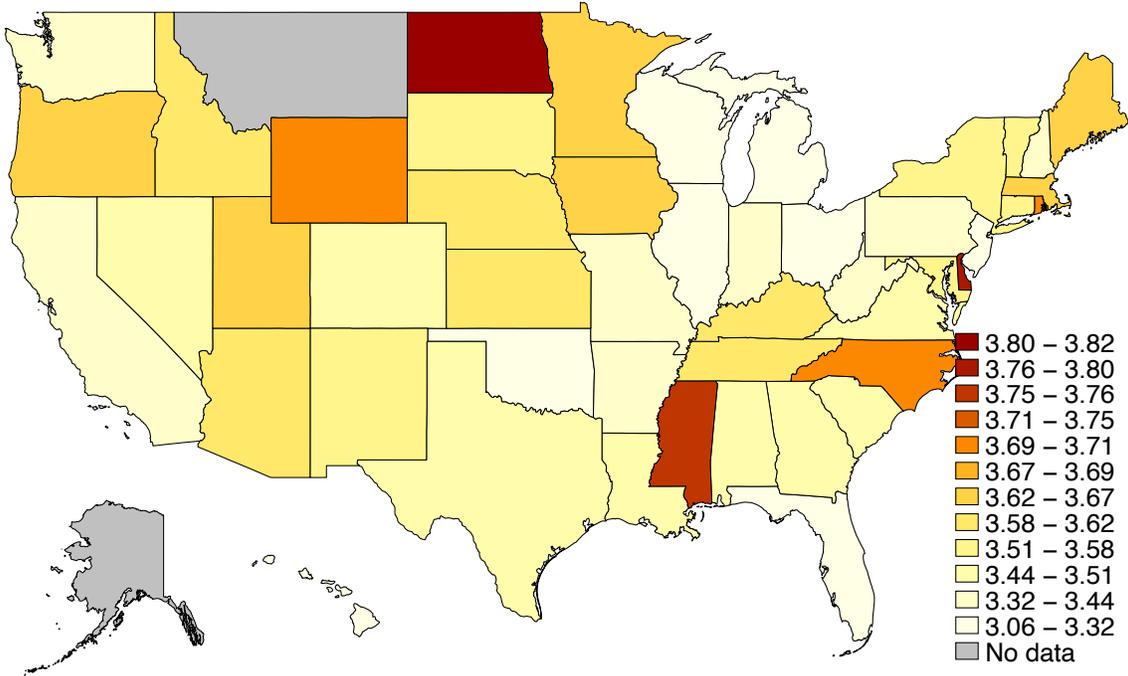
Panel B: Medicaid recovery rate index by state
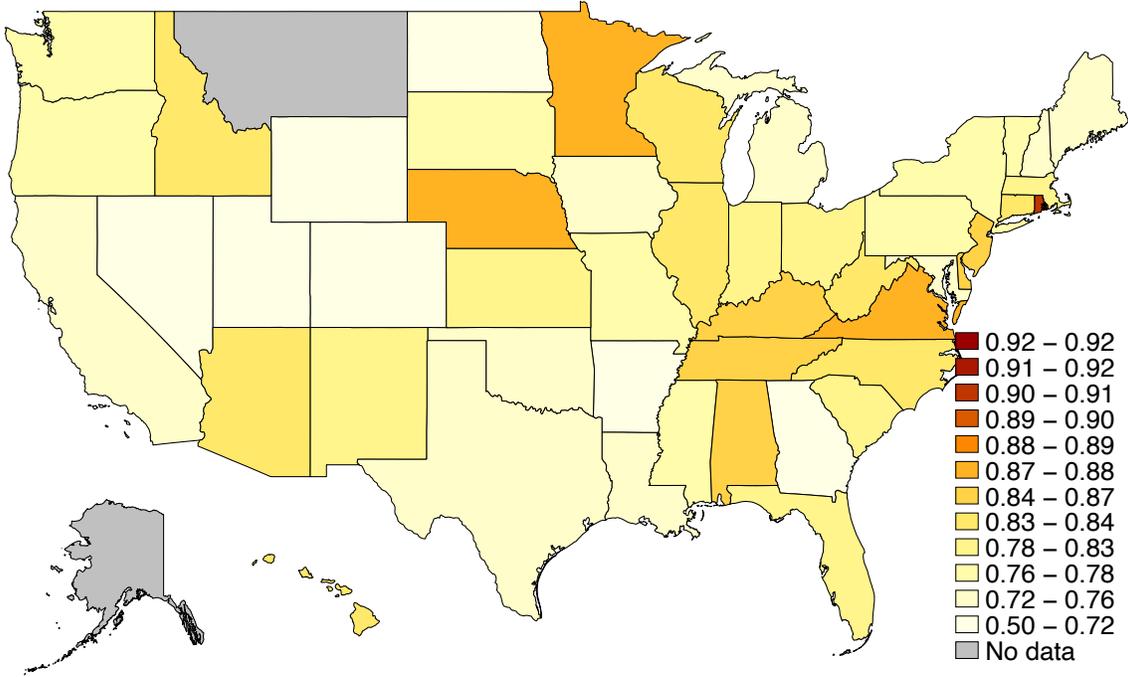
Legend:
- 0.92 – 0.92
- 0.91 – 0.92
- 0.90 – 0.91
- 0.89 – 0.90
- 0.88 – 0.89
- 0.87 – 0.88
- 0.84 – 0.87
- 0.83 – 0.84
- 0.78 – 0.83
- 0.76 – 0.78
- 0.72 – 0.76
- 0.50 – 0.72
- No data

These maps shows state-by-insurer fee and recovery rate indices estimated for each state's Medicaid program. Darker colors indicate higher values. In both panels, diffferences in the index levels can be interpreted as approximately percentage changes (divided by 100), so a difference of 0.1 between two states means a 10 percent higher fee or 10 percent (not percentage point) higher recovery rate.

Figure 2: State-Level Medicare Fees and Recovery Rate

Panel A: Medicare fee index by state

■	3.80 − 3.91
■	3.76 − 3.80
■	3.75 − 3.76
■	3.71 − 3.75
■	3.69 − 3.71
■	3.67 − 3.69
■	3.62 − 3.67
■	3.58 − 3.62
■	3.51 − 3.58
■	3.44 − 3.51
■	3.32 − 3.44
□	3.32 − 3.32
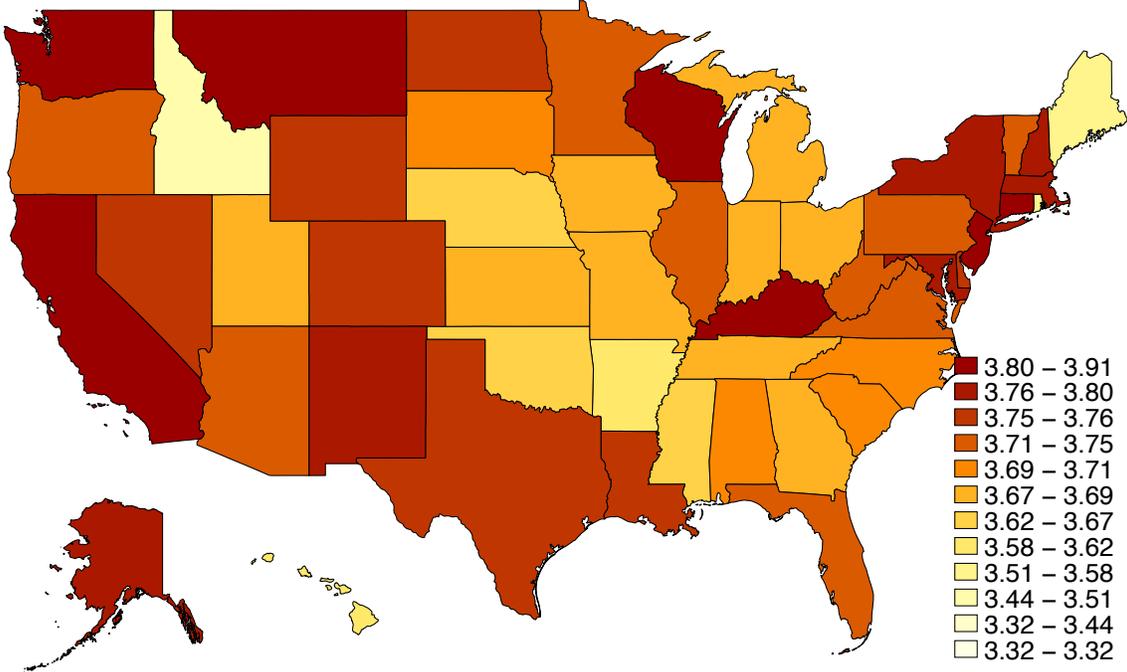
Panel B: Medicare recovery rate index by state
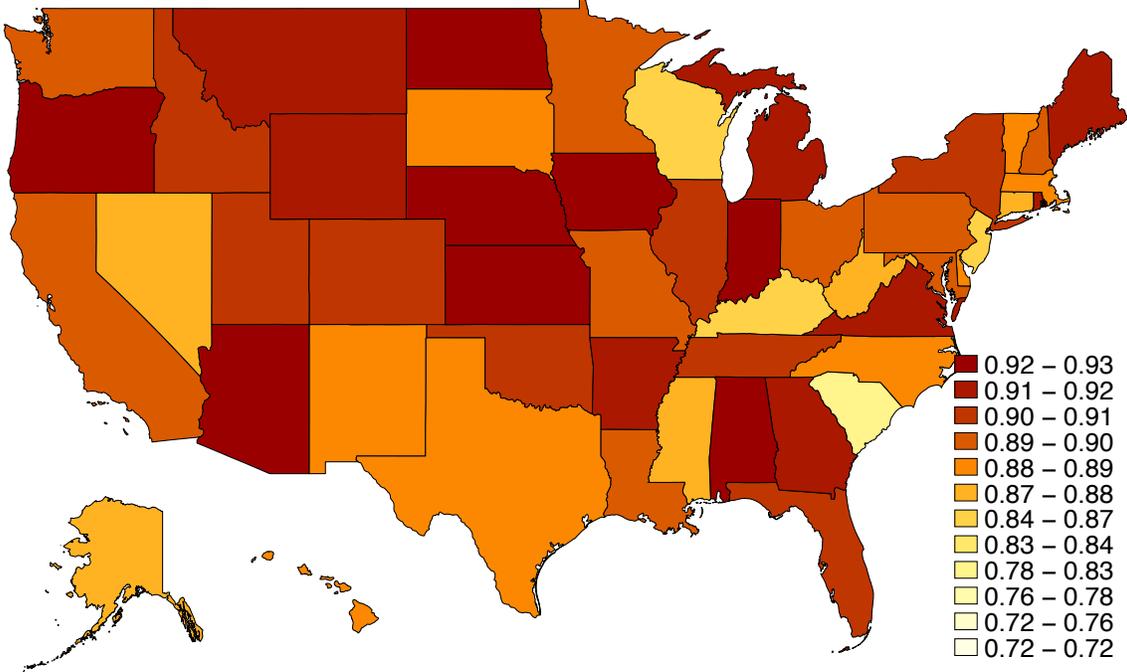
■	0.92 − 0.93
■	0.91 − 0.92
■	0.90 − 0.91
■	0.89 − 0.90
■	0.88 − 0.89
■	0.87 − 0.88
■	0.84 − 0.87
■	0.83 − 0.84
■	0.78 − 0.83
■	0.76 − 0.78
■	0.72 − 0.76
□	0.72 − 0.72

These maps shows state-by-insurer fee and recovery rate indices estimated for each state's Medicaid program. Darker colors indicate higher values. In both panels, diffferences in the index levels can be interpreted as approximately percentage changes (divided by 100), so a difference of 0.1 between two states means a 10 percent higher fee or 10 percent (not percentage point) higher recovery rate.
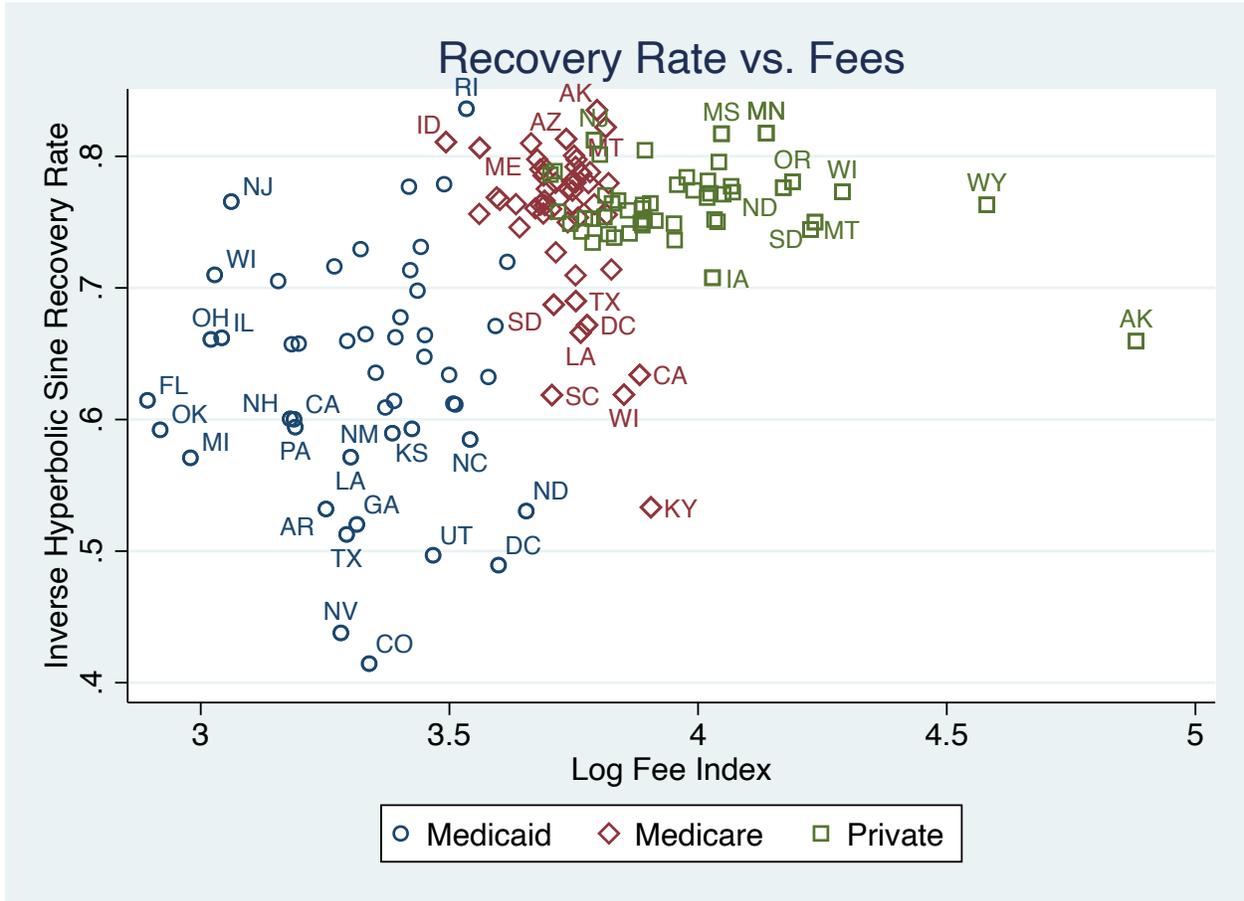
This figure shows the estimates of state-by-insurer fees and billing difficulty, the latter measured as the log share of revenue lost. The open circles show values for Medicaid, the red diamonds show Medicare, and the green squares show private insurance.

Figure 4: Supply Curve Estimates from Movers

Panel A: Movers' responses to changes in Medicaid fees
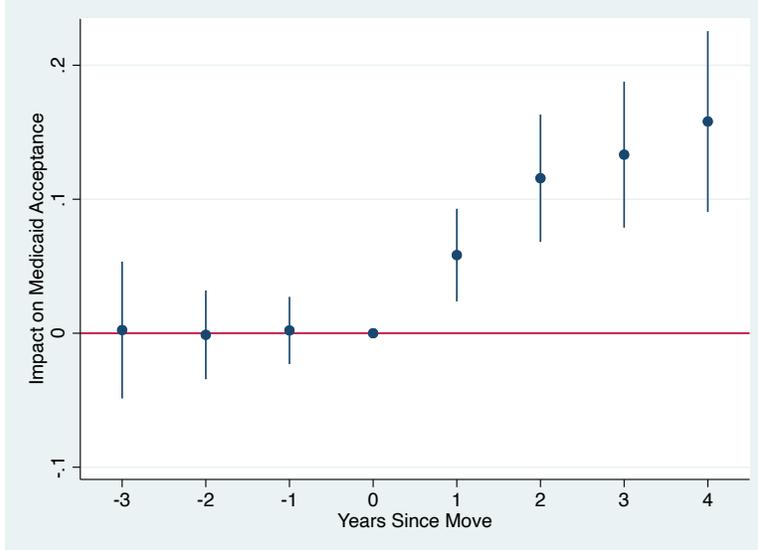
Panel B: Movers' responses to changes in Medicaid recovery rates
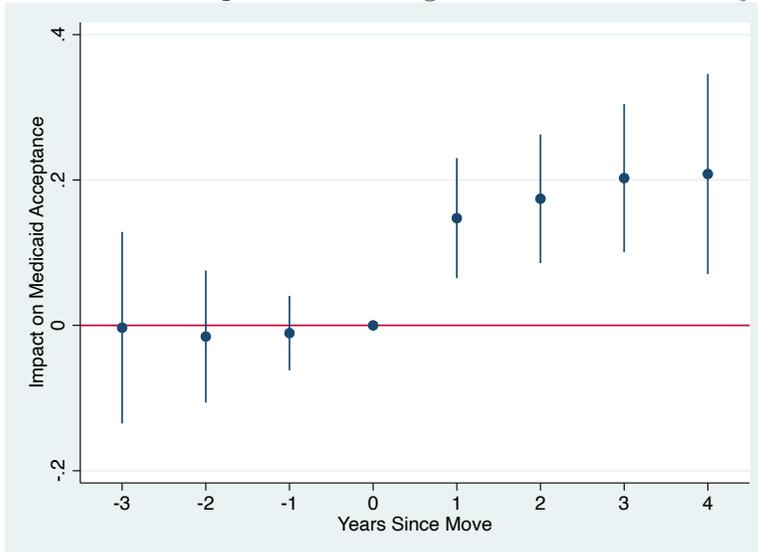
This figure shows event study estimates of how Medicaid acceptance responds to fees (Panel A) and billing difficulty (Panel B) before and after a physician moves across states. Each coefficient shows the impact of of moving to a state with a 1-log point higher index value on the probability a physician reports accepting Medicaid.

Figure 5: Raw Data: Changes in Movers' Willingness to Treat Medicaid Patients
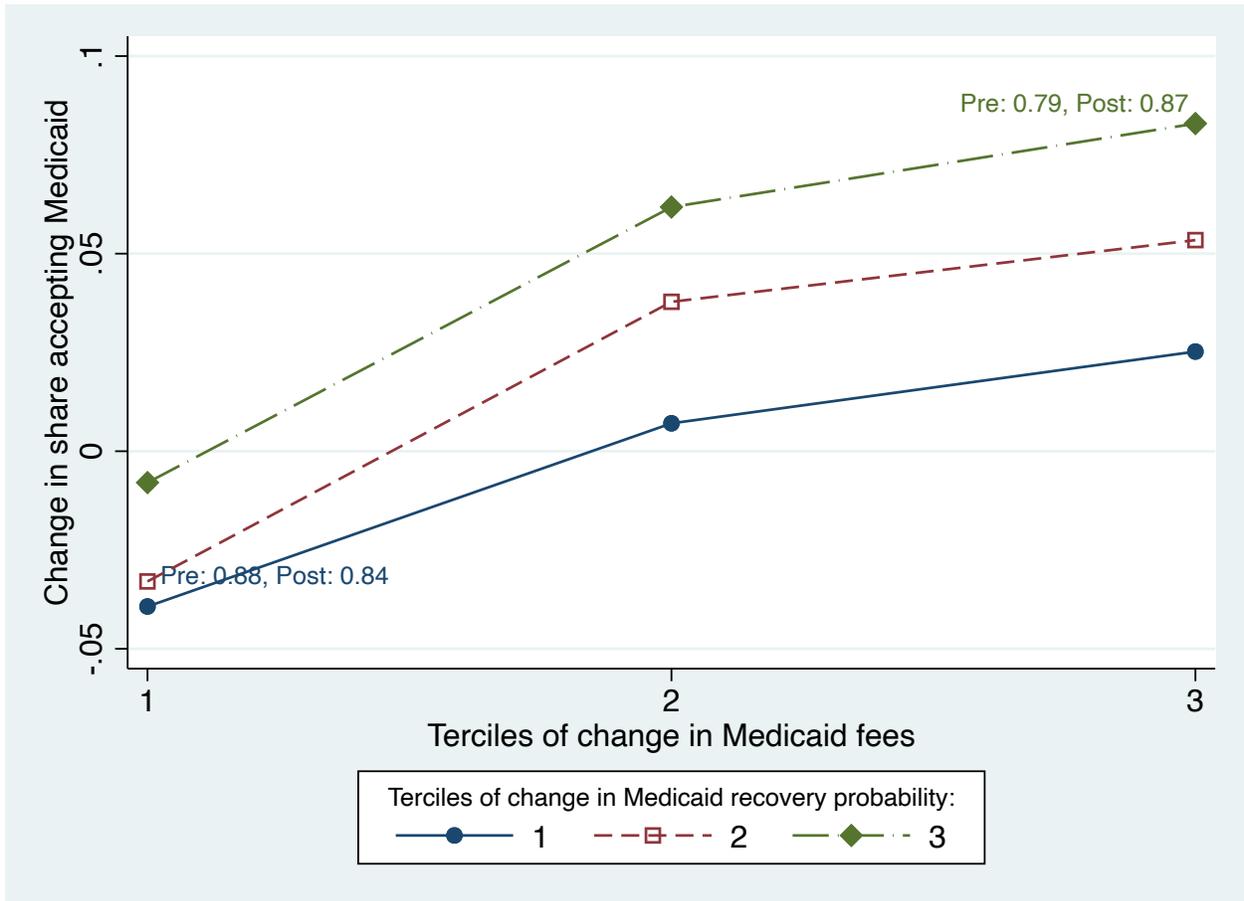
This figure shows the raw data from the movers design. The graph consists of nine points, constructed by splitting the sample of physician movers into terciles along two dimensions. The first dimension, shown in the three separate lines, is the difference in the Medicaid recovery rate index from the pre-move state to the post-move state. The blue line labled "1" is the lowest tercile, corresponding to those who move to a state with a lower Medicaid recovery rate (increased billing difficulty). The dashed line labeled "2" is the middle tercile, corresponding to those who move to a state with similar a Medicaid recovery rate to the original state. The green line with diamonds, labled "3", correspond to the top tercile of physicians who move to state with a higher recovery rate. Along each of those lines, moving from left to right shifts among terciles of the Medicaid fee index. So those physicians on the left side of the graph move to a state with lower reimbursement rates, while those on the right move to states with higher rates. For each point, defined by terciles of these two dimensions, the vertical value shows the raw change in probability of accepting Medicaid for physicians in that group. We see that higher fees and higher recovery probabilities both increase physicians' willingness to accept Medicaid.

Figure 6: Supply Curve Estimates from Movers: Index Robustness

Left column:
Impact of Medicaid fee

Right column:
Impact of Medicaid recovery rate

Panels A/B: Indices constructed conditional on zip code median income
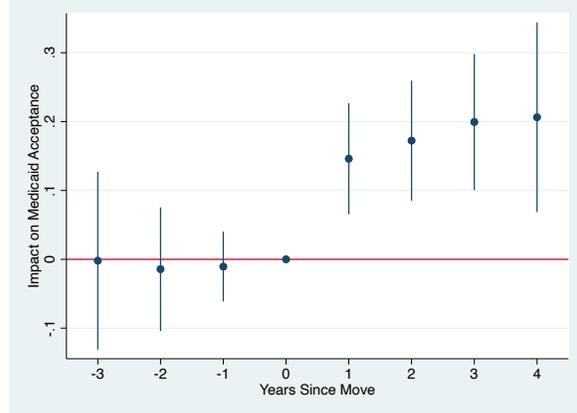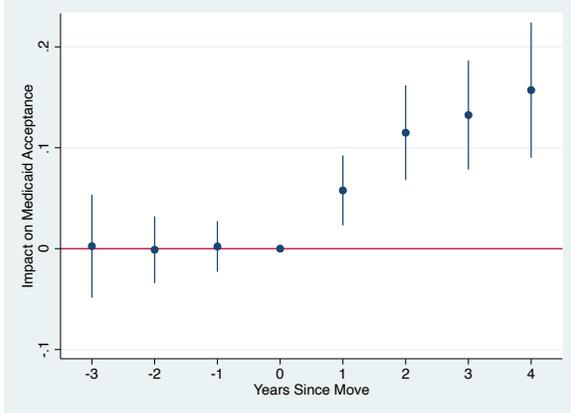
Panels C/D: Indices excluding first visit by patient/physician pair
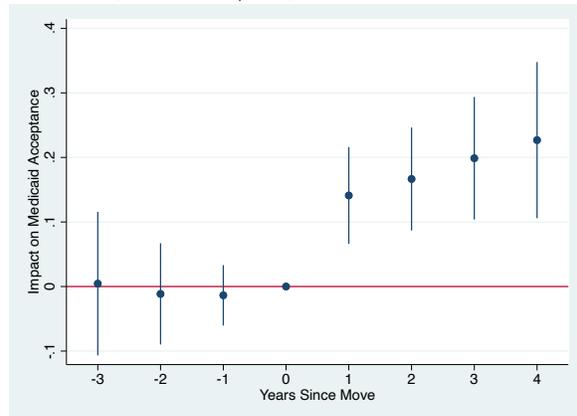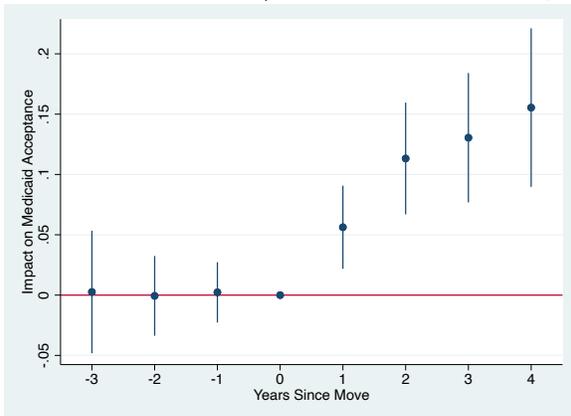
Panels E/F: Indices using only first visit by patient/physician pair
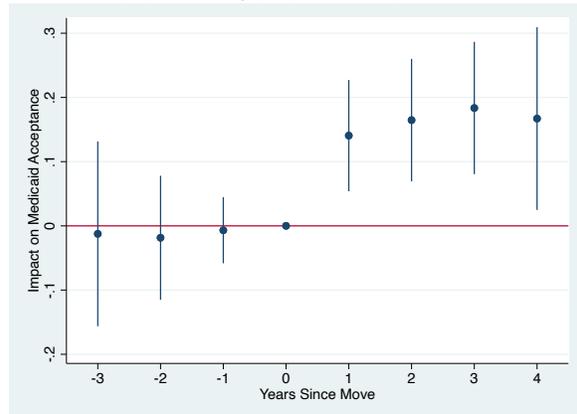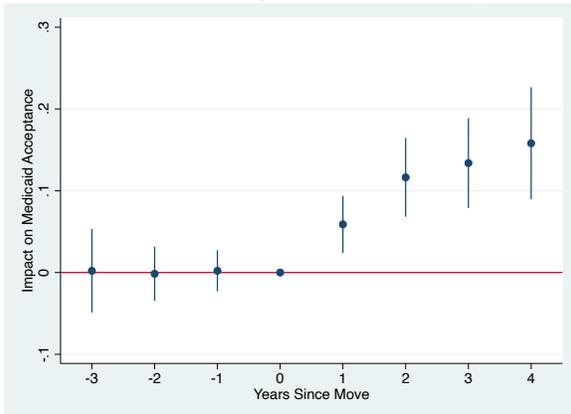
This figure shows the robustness of our movers-based supply estimates to alternate Medicaid fee (Panel A) and recovery rate (Panel B) indices constructed using versions of regressions (1) and (2) that additionally control for the zip code's median income. The remaining panels show robustness to versions of the indices calculated on samples excluding (Panels C and D) or limited to (Panels E and F) the first visit from each patient to each physician.

Figure 7: Supply Curve Estimates from Movers: Alternate Recovery Rates

Left column:
Impact of Medicaid fee

Right column:
Impact of Medicaid recovery rate

Panels A/B: Recovery rate defined relative to Medicare rates
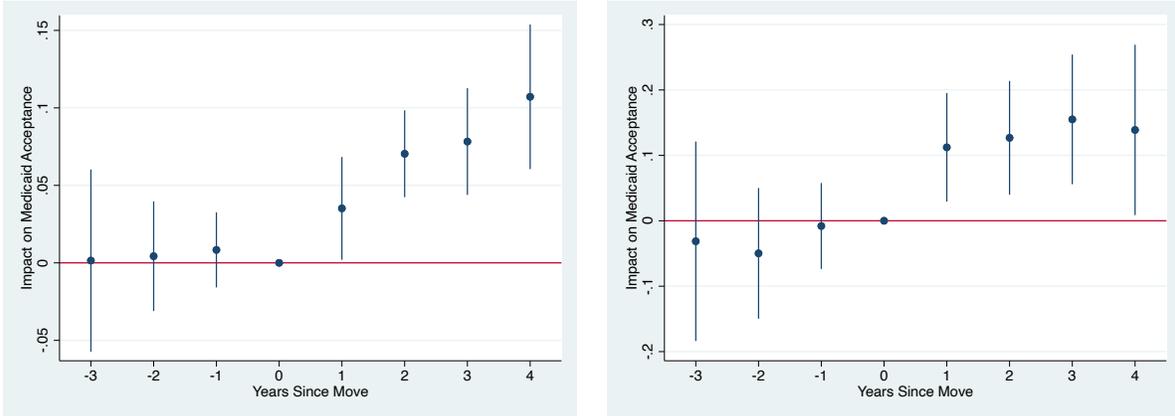
Panels C/D: Recovery rate defined relative to richest imputation available
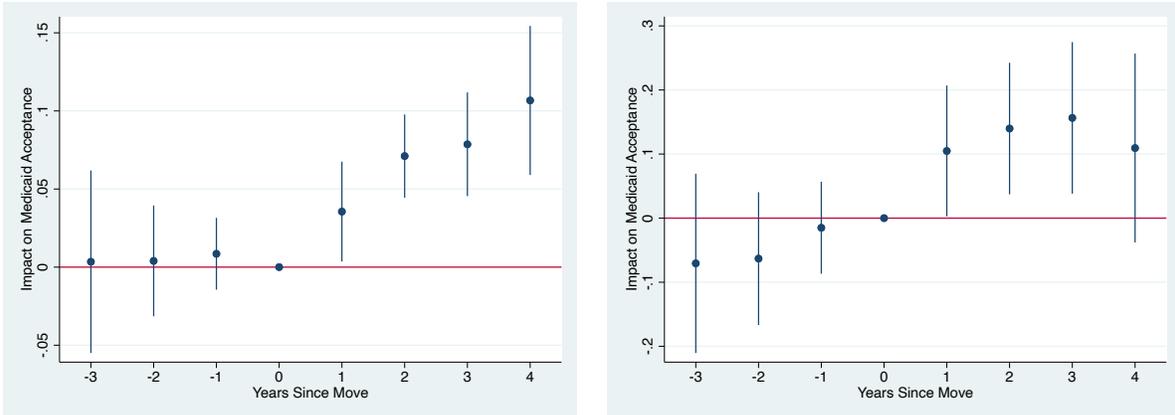
Panels E/F: Recovery rate aggregated across all claims for the visit
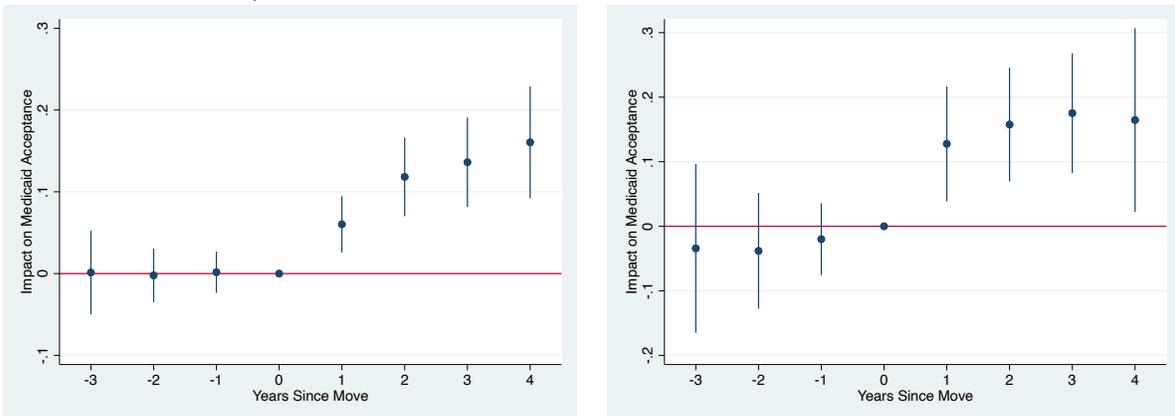
This figure shows the robustness of our movers-based supply estimates to different methods for constructing the recovery rate index, described in Appendix A.1.

Figure 8: Supply Curve Estimates from Cross-State Groups

Left column:
Impact of Medicaid fee

Right column:
Impact of Medicaid recovery rate

Panels A/B: Groups defined based on SK&A system and group variables, plus Tax ID
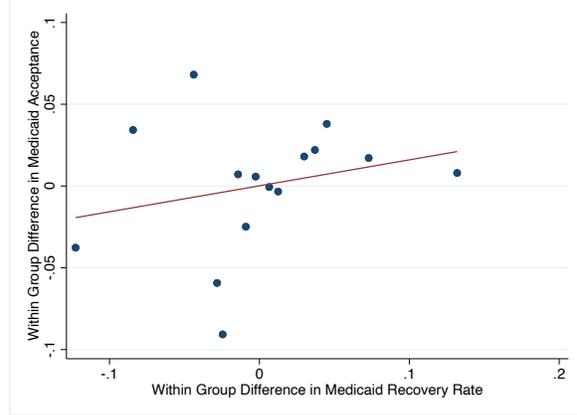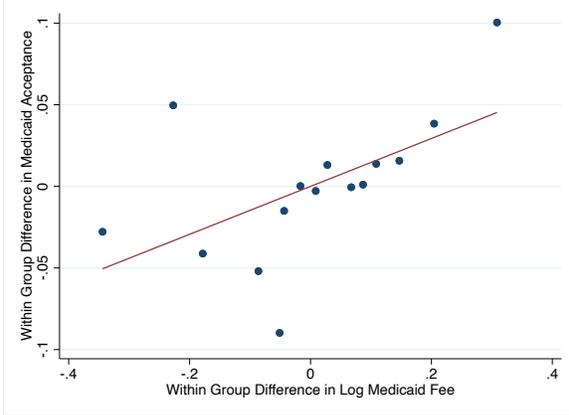
Panels C/D: Groups defined only based on SK&A system and group variables
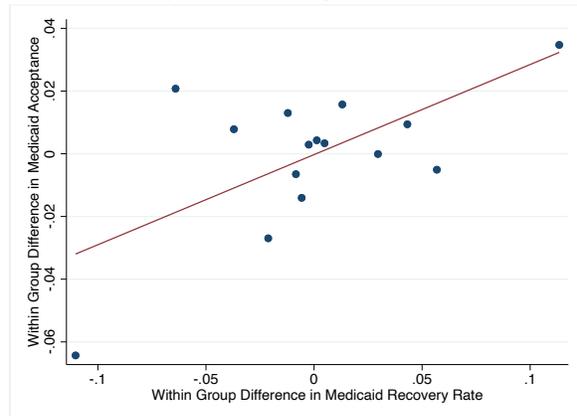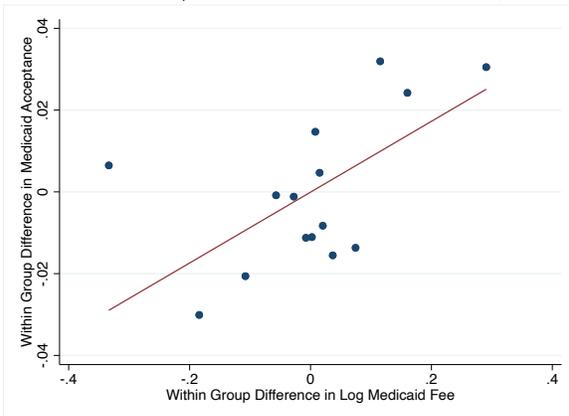
Panels E/F: Groups defined only based on Medicare Tax ID
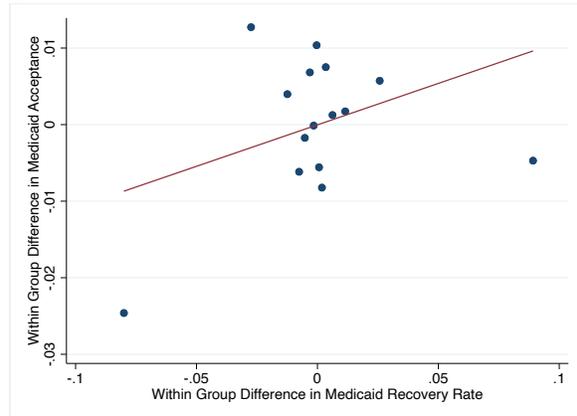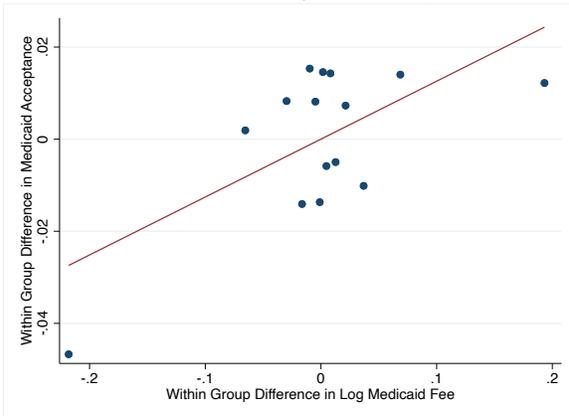
This figure shows how physicians' willingness to treat Medicaid patients varies among physicians within the same group or system, when that group or system includes physicians in multiple states. In each panel, the y axis value represents the residual from regressing each physician's Medicaid acceptance decision on a group-level fixed effects. The x axis value shows the residual from regressing the relevant state index (fee or recovery rate) on the same fixed effects. In other words, both the x and y axes show the physicians' within-group difference, relative to a group-level mean, in the variables listed. We then make a binned scatterplot by grouping these residuals into 15 bins based on their x axis values, and ploting the x and y means within each bin.

52

Figure 9: Empirical Value Functions after Denials

Empirical CDF of continuation value for different actions available to provider after denial
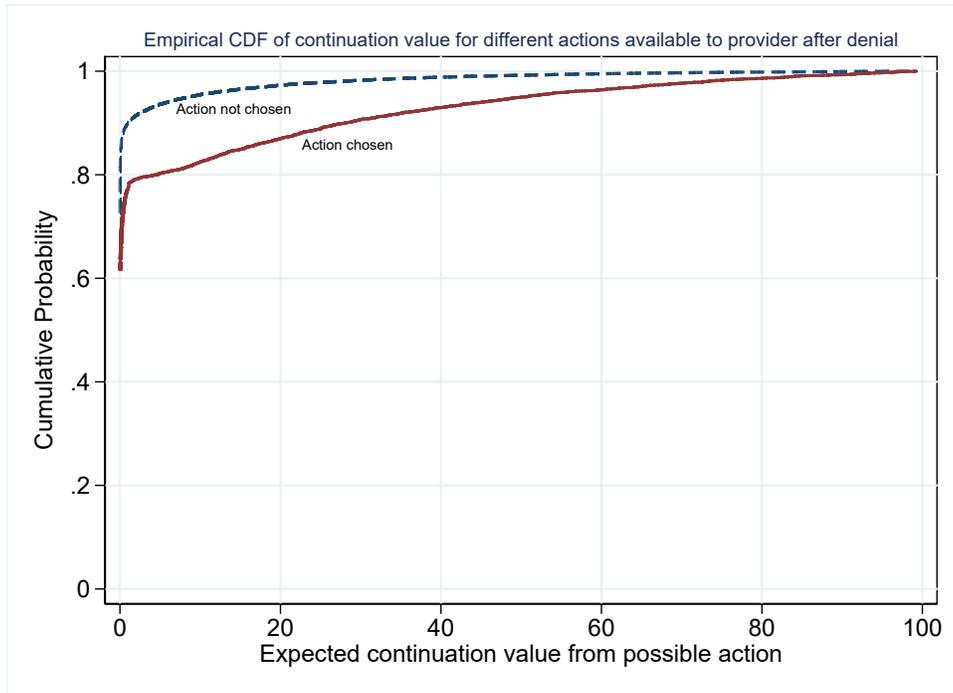
This figure shows the empirical CDF of the value function of different feasible actions for the provider after a denial. We apply the CCP method developed in Hotz and Miller (1993) to the sequence of resubmission and payment decisions computing empirical value functions conditioning on diagnosis, procedure, and state. The procedure is described in Section 6. The difference between the value of observed actions and the value of counterfactual actions identifies the parameters governing resubmission costs.

Figure 10: Model Estimation Results

Panel A: Cost of resubmissions if more than one claim for the same visit
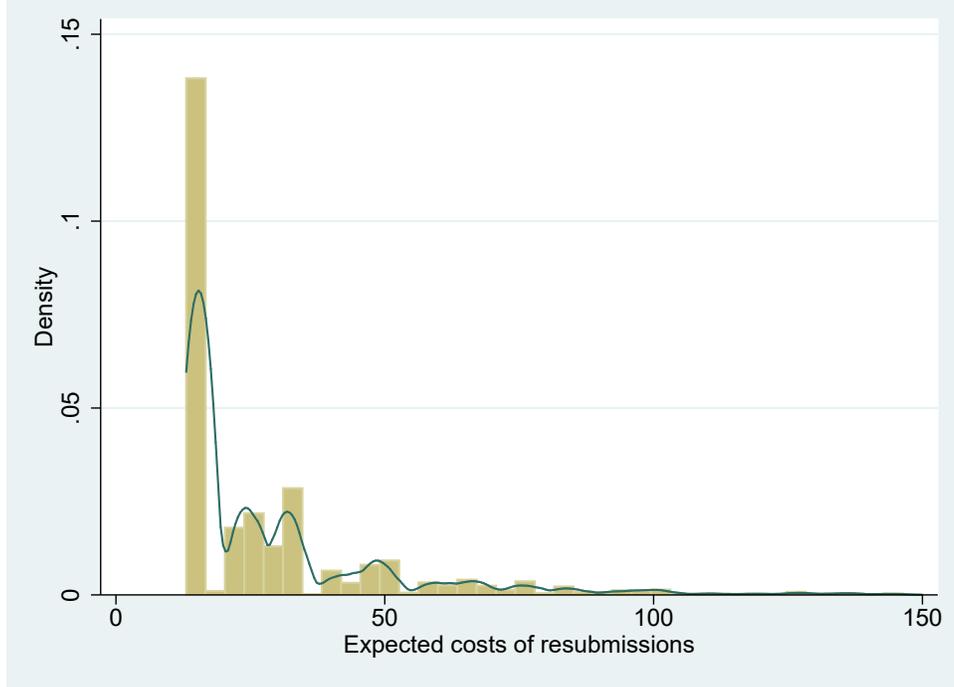
Panel B: Cost of increasing the probability of full payment by 0.1
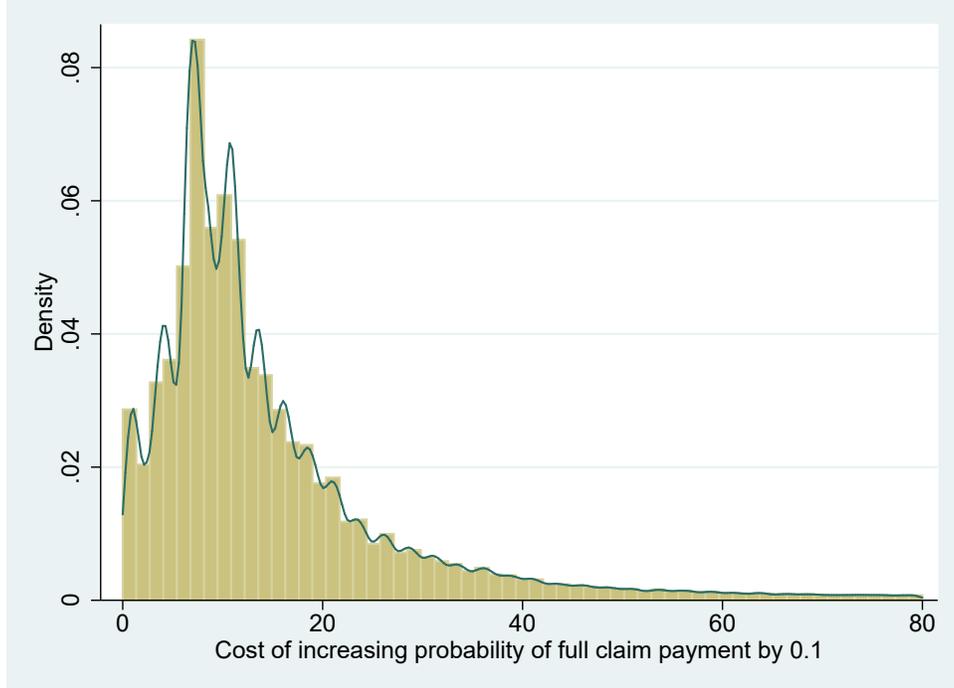
This figure shows the results of our model estimates. The first panel shows the distribution of expected resubmission costs across visits, conditional on the visit entering the denial-and-resubmission game. The second panel shows the distribution of the estimated cost of the *ex ante* (i.e. before the first claim submission) effort that the physician would have to spend to increase the probability of successful payment of the first claim by 0.1.

Table 1: Descriptive Statistics

	Medicaid FFS	Medicaid MCO	Medicare FFS	Medicare Adv.	Private
Average Procedure Price	56.58	57.80	70.63	76.41	89.75
Average Visit Price	111.17	105.99	140.85	154.47	190.36
Average Probability of Procedure Denial	0.317	0.195	0.192	0.179	0.193
Average Probability of Full Claim Denial	0.187	0.087	0.088	0.094	0.148
Average Probability of Post-denial Resubmission	0.253	0.173	0.158	0.092	0.072

This table summarizes the data used for model estimation. The moments shown here, by insurer, are the key ones that enter into the model estimation and interpretation. Column 1 refers to traditional (state-run) Medicaid. Column 2 shows Medicaid managed care organizations (MCOs). Column 3 shows traditional Medicare (Medicare fee-for-service, or FFS). Column 4 shows privately run Medicare Advantage insurance, and the final column shows private commercial insurance.

Table 2: State-Level Descriptive Statistics

	Mean	SD	10th pctile	90th pctile
All Insurers				
Fee Index	3.62	0.27	3.27	3.87
Recovery Rate Index Def. 1	0.73	0.08	0.62	0.81
Recovery Rate Index Def. 2	0.75	0.08	0.64	0.83
Recovery Rate Index Def. 3	0.72	0.08	0.59	0.80
Medicaid				
Patient Acceptance Rate	0.80	0.10	0.69	0.93
Fee Index	3.32	0.20	3.02	3.58
Recovery Rate Index Def. 1	0.66	0.08	0.55	0.76
Recovery Rate Index Def. 2	0.68	0.09	0.58	0.79
Recovery Rate Index Def. 3	0.63	0.09	0.51	0.73
Medicare				
Patient Acceptance Rate	0.92	0.026	0.89	0.95
Fee Index	3.72	0.08	3.63	3.82
Recovery Rate Index Def. 1	0.75	0.06	0.67	0.81
Recovery Rate Index Def. 2	0.79	0.06	0.69	0.82
Recovery Rate Index Def. 3	0.75	0.06	0.67	0.81
Private Insurers				
Fee Index	3.97	0.21	3.77	4.19
Recovery Rate Index Def. 1	0.78	0.03	0.75	0.81
Recovery Rate Index Def. 2	0.81	0.02	0.79	0.85
Recovery Rate Index Def. 3	0.76	0.03	0.74	0.80

This table shows the distribution of key measures for our state level supply estimates. We see, both across all insurers and for each category of insurance, the fee indices, lost revenue index, and share of physicians accepting Medicaid and Medicare patients. The means, standard deviations, and percentiles reflect distributions across states.

Table 3: Mover Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Accept Medicaid Patients?				
Post-Move × Fee	0.1041***	0.0998***	0.1110***	0.0692***	0.0647***	0.0602***	0.0607***	0.0607***
	(0.0207)	(0.0190)	(0.0192)	(0.0142)	(0.0133)	(0.0134)	(0.0135)	(0.0135)
Post-Move × Recovery Rate	0.1794***	0.1807***	0.1723***	0.1122***	0.1316***	0.1399***	0.1411***	0.1411***
	(0.0347)	(0.0331)	(0.0342)	(0.0319)	(0.0355)	(0.0376)	(0.0381)	(0.0381)
N	56893	56893	56893	56893	56893	56893	56893	56893
R^2	0.600	0.602	0.603	0.604	0.605	0.605	0.605	0.605
$\hat{\gamma}/\hat{\beta}$	1.724	1.810	1.552	1.622	2.035	2.325	2.324	2.324
	(0.250)	(0.265)	(0.236)	(0.408)	(0.480)	(0.562)	(0.557)	(0.557)
δ s.t. Post-Move × Fee = 0	-14.001	-392.631	-24.005	2.524	2.248	1.901	1.917	1.917
δ s.t. Post-Move × Recovery Rate = 0	-15.040	-21.686	49.045	2.609	3.788	4.225	4.276	4.276
R^2 Within Physician	0.010	0.014	0.016	0.019	0.022	0.022	0.022	0.022
R_{max}	0.013	0.019	0.021	0.025	0.028	0.028	0.028	0.028
Clustering	State	State	State	State	State	State	State	State
Phys. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	No	No	No	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	No	No	No	No	Yes	Yes	Yes	Yes
Pop. Density Contr.	No	No	No	No	No	Yes	Yes	Yes
MD per capita Contr.	No	No	No	No	No	No	Yes	Yes

This table shows estimates of equation (8) with various sets of controls. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move and the difference between our fee and recovery rate indices estimated from the post-move and pre-move states. Columns 2 and following add additional controls for characteristics of the state or county in which the physician is located. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the physician fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within physician (shown as R_{max}) relative to the observed within-physician R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table 4: Mover Results Separated by Managed Care vs. Traditional

	Accept Medicaid Patients?							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-Move × Medicaid Fee	0.0971***	0.0926***	0.1025***	0.0660***	0.0605***	0.0549***	0.0557***	0.0557***
	(0.0216)	(0.0205)	(0.0198)	(0.0142)	(0.0132)	(0.0132)	(0.0133)	(0.0133)
Post-Move × FFS Recovery Rate	0.0966***	0.1002***	0.0963***	0.0733***	0.0776***	0.0845***	0.0842***	0.0842***
	(0.0282)	(0.0279)	(0.0299)	(0.0247)	(0.0233)	(0.0244)	(0.0246)	(0.0246)
Post-Move × MCO Recovery Rate	0.1533***	0.1225**	0.1262**	0.0742	0.0655	0.0612	0.0629	0.0629
	(0.0510)	(0.0492)	(0.0516)	(0.0443)	(0.0425)	(0.0424)	(0.0430)	(0.0430)
Post-Move × FFS Recovery Rate × FFS Share	0.0321	0.0421	0.0252	0.0085	0.0241	0.0304	0.0317	0.0317
	(0.0350)	(0.0361)	(0.0382)	(0.0358)	(0.0369)	(0.0374)	(0.0372)	(0.0372)
Post-Move × MCO Recovery Rate × MCO Share	0.1812**	0.1937**	0.1709**	0.1809**	0.1946**	0.1999**	0.2053***	0.2053***
	(0.0779)	(0.0776)	(0.0806)	(0.0737)	(0.0746)	(0.0751)	(0.0757)	(0.0757)
FFS Recovery Rate × FFS Share (Pre-move)	0.0098	0.0049	0.0137	0.0231	0.0130	0.0099	0.0093	0.0093
	(0.0240)	(0.0239)	(0.0227)	(0.0208)	(0.0209)	(0.0211)	(0.0212)	(0.0212)
MCO Recovery Rate × MCO Share (Pre-move)	-0.0436	-0.0518	-0.0390	-0.0323	-0.0378	-0.0403	-0.0430	-0.0430
	(0.0458)	(0.0445)	(0.0469)	(0.0454)	(0.0466)	(0.0467)	(0.0474)	(0.0474)
N	51277	51277	51277	51277	51277	51277	51277	51277
R^2	0.603	0.605	0.606	0.607	0.608	0.608	0.608	0.608
Clustering	State	State	State	State	State	State	State	State
Phys. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	No	No	No	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	No	No	No	No	Yes	Yes	Yes	Yes
Pop. Density Contr.	No	No	No	No	No	Yes	Yes	Yes
MD per capita Contr.	No	No	No	No	No	No	Yes	Yes

This table shows mover estimates using separate recovery rate indices for Medicaid Fee-for-Service (FFS) and Medicaid Managed Care Organizations (MCO). Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move, difference between our fee and recovery rate indices estimated from the post-move and pre-move states, and the recover rate differences interacted with states' shares of Medicaid patients in FFS vs. MCO. The share variables are standardized. Columns 2 and following add additional controls for characteristics of the state or county in which the physician is located.

Table 5: Cross-State Group Results

				Accept Medicaid Patients?			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Medicaid Fee	0.0900**	0.0921**	0.1313***	0.0853*	0.0743*	0.0681	0.0682
	(0.0403)	(0.0393)	(0.0400)	(0.0459)	(0.0436)	(0.0433)	(0.0434)
Medicaid Recovery Rate	0.2887***	0.2737**	0.2700**	0.2069**	0.2268**	0.2350**	0.2367**
	(0.1047)	(0.1078)	(0.1039)	(0.1018)	(0.0993)	(0.1028)	(0.1044)
N	12252	12252	12252	12252	12252	12252	12252
R^2	0.194	0.195	0.203	0.205	0.206	0.207	0.207
$\hat{\gamma}/\hat{\beta}$	3.206	2.972	2.056	2.427	3.054	3.453	3.469
	(1.064)	(0.995)	(0.597)	(1.017)	(1.355)	(1.617)	(1.618)
δ s.t. Medicaid Fee = 0	-19.172	-17.579	-6.189	13.426	6.662	5.057	5.090
δ s.t. Medicaid Recovery Rate = 0	-22.291	285.840	42.947	5.516	6.952	7.732	7.881
R^2 Within Group	0.006	0.007	0.017	0.020	0.021	0.021	0.021
R_{max}	0.007	0.009	0.022	0.026	0.027	0.028	0.028
Clustering	State	State	State	State	State	State	State
Group Definition				SK&A Group and System			
Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	No	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	No	No	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	No	No	No	Yes	Yes	Yes	Yes
Pop. Income Contr.	No	No	No	No	Yes	Yes	Yes
Pop. Density Contr.	No	No	No	No	No	Yes	Yes
MD per capita Contr.	No	No	No	No	No	No	Yes

This table shows estimates of equation (10) with various sets of controls. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who practice in groups that have locations in multiple states. The key independent variables are fee and recovery rate indices in the state where the physician is located. Columns 2 and following add additional controls for characteristics of the state or county in which the physician is located. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the practice fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within group (shown as R_{max}) relative to the observed within-group R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table 6: Empirical Value Functions and Resubmission Decisions

| | Conditioning Variables | | | |
	None (1)	Diagnosis (2)	Diagnosis State (3)	Diagnosis State Medical Group (4)
OLS coefficient (p value)	0.00328712 (0.00008)	0.00334352 (0.00007)	0.00352014 (0.00007)	0.00413057 (0.00006)
Observations	77846	76693	72897	66397

The table presents the OLS regression coefficient predicting the relationship between the value of a feasible action for the provider after a denial takes place and the probability that the action is observed in the data. Each column correspond to a different set of conditioning variables. The results highlight the monotone relationship between empirical continuation values and probability of action. This relationship identifies the parameters governing resubmission costs.

Table 7: Claim Amount and Probability of Payment

	Conditioning Variables				
	Procedure (1)	Diagnosis Procedure (2)	State Diagnosis Procedure (3)	Specialty State Diagnosis Procedure (4)	Medical Group Specialty State Diagnosis Procedure (5)
OLS coefficient (st error)	0.000408 (0.00002)	0.000521 (0.00002)	0.000489 (0.00004)	0.000425 (0.00004)	0.000578 (0.00004)
Observations	157306	157306	157306	157306	157306

The table presents the OLS regression coefficient predicting the relationship between the probability of full claim payment and the billed amount for the visit. Each column correspond to a different set of conditioning variables. The results highlight the monotone relationship between visit value and successfull billing, which we interpret through the model as higher billing effort from the provider. Along with the estimates from the continuation game, this relationship identifies the parameters of the initial effort cost.

Table 8: Summary of model estimates

	Mean	Median	St Dev
Expected costs of resubmissions (if any)	31.26	22.81	29.48
Expected revenues from resubmissions	41.11	76.81	139.39
Cost of 0.1 increase in probability of full 1st claim payment	10.51	12.58	10.45
Expected billing costs per visit	23.75	23.53	0.737

This table shows results from estimating the model in section 6. The first row shows the estimated costs to the physician practice of resubmitting claims, conditional on resubmitting the claim for a given visit at least once. The second row shows the expected revenues from these resubmissions; we see that the mean is about \$10 (30 percent) higher than the costs, but with a substantial variance. The third row shows the cost of effort the practice would have to expend to increase the probability of full payment of the first claim by 10 percentage points. This averages \$10.50, or around 10 percent of the average claim value. The final row puts these together and shows the overall expected billing costs for the visit as a whole. This row incorporates the initial effort cost to move from around the bottom to the top of the claim success distribution (from $p = 0.7$ to $p = 1$), plus the probability of engaging in resubmissions combined with the costs of those resubmissions.

A   Variable Construction

A.1   Recovery Rate

Recall that the recovery rate is defined on page 13 as $p_i = \tilde{c}_i/c_i$. We see the payment amounts $\tilde{c}_i$ in the data, but need to impute the claim value c_i that the physician could receive if there were no problems billing for the visit. In the definitions below, we refer to c_i as the imputed allowed amount. The difference $c_i - \tilde{c}_i$ is the "amount lost".

Definition 1   Amount lost is defined as the difference between c_i, the (imputed) allowed amount of the last claim we observe for a visit, and the amount paid on this last claim, $\tilde{c}_i$. The amount lost is constrained to be zero for cases for which the above definition would yield a negative amount lost. The observation is set to be missing if the observed payment in the last claim was negative. The amount lost measure is also capped at $10,000. We then compute the recovery rate as $p_i = \tilde{c}_i/c_i$. This definition is used in Panel B of Figure 7.

Definition 2   This is the same as Definition 1 but we use imputations based on Medicare rates. That is, the imputed claim value c_i is based on the observed markups for the physician-insurer pair relative to the rates implied by Medicare's fee schedule of RVUs. This definition is used in Panel D of Figure 7.

Definition 3   In this definition, the amount lost is defined as the difference between the (imputed) allowed amount of the last claim we observe for a visit and the total amount paid for the visit. That is, all payments observed for the visit, including reversals, are summed up. The amount lost is constrained to be zero for cases for which the above definition would yield a negative amount lost. The observation is set to be missing if the observed payment for the entire visit is negative. The amount lost measure is also capped at $10,000. This definition is used in Panel F of Figure 7.

A.2   Group and System Definitions

We use information on physician groups and their system from the SK&A data source and tax identification number (TIN) information from MD-PPAS to form four alternative definitions of group. The SK&A data contains two variables that are used to identify physicians working in different locations that belong to the same organization: a Medical Group Practice variable and Health System variable. The Medical Group Practice variable identifies different physician offices belonging to the same medical practice, often under a common specialty, while the system variable is a broader definition, in which multiple groups belong to the same health system. As it is not clear at what organizational level physician's choose to accept or reject Medicaid, we examine two alternative definitions of group:

1. Identify group using SK&A group, but use the SK&A system variable for groups belonging to a broader system.

2. Identify group using SK&A group, use the SK&A system variable for groups belonging to a broader system, and use TIN information from MD-PASS for groups not identified by SK&A.

In all cases, the group definition is defined on an annual basis, so identification is not based on changes in acceptance or rejection of insurance in an organization over time, but only different decisions by the same organization in a given time period. Definition 3 has the advantage over the alternatives as it is based on administrative data; however, it is possible for one large organization to have multiple TINs, so that this definition may be too narrow.

B   Additional Results

Figure B.1: Index Stability Across Years

Panel A: Fee Index   Panel B: Recovery Rate Index
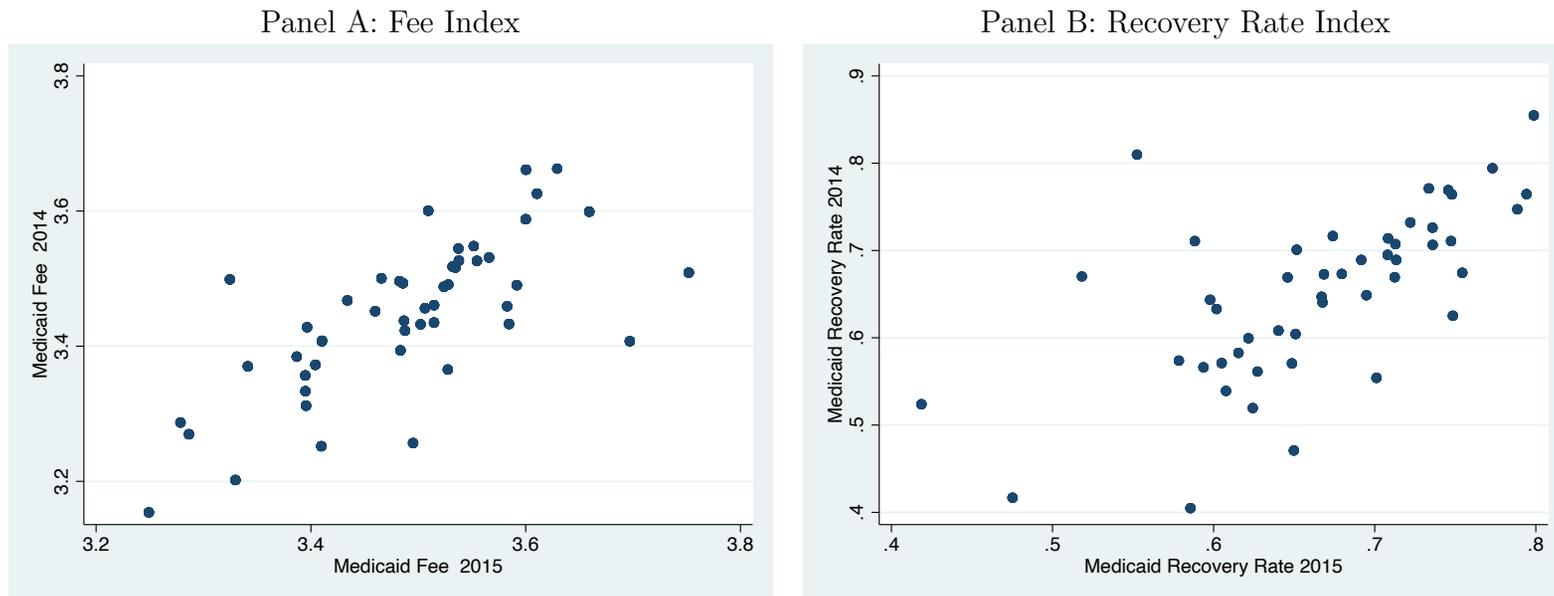

This figure shows the stability of our state Medicaid indices across years. We estimate our the fee and recovery rate indices from section 2.2 by state using separate annual regressions. We plot state-level scatterplots of the 2014 vs. 2015 index values.

Table B.1: Mover Results: Recovery Rate Definition 2

	Accept Medicaid Patients?							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-Move × Fee	0.1075***	0.1033***	0.1146***	0.0707***	0.0668***	0.0626***	0.0631***	0.0631***
	(0.0213)	(0.0196)	(0.0198)	(0.0141)	(0.0130)	(0.0130)	(0.0132)	(0.0132)
Post-Move × Recovery Rate	0.1675***	0.1702***	0.1597***	0.1069***	0.1294***	0.1364***	0.1370***	0.1370***
	(0.0371)	(0.0353)	(0.0357)	(0.0339)	(0.0394)	(0.0416)	(0.0421)	(0.0421)
N	56893	56893	56893	56893	56893	56893	56893	56893
R^2	0.600	0.602	0.603	0.604	0.605	0.605	0.605	0.605
$\hat{\gamma}/\hat{\beta}$	1.558	1.648	1.394	1.511	1.938	2.178	2.173	2.173
	(0.236)	(0.249)	(0.223)	(0.389)	(0.453)	(0.518)	(0.514)	(0.514)
δ s.t. Post-Move × Fee = 0	-8.202	-44.794	-16.835	2.600	2.354	2.015	2.025	2.025
δ s.t. Post-Move × Recovery Rate = 0	-7.313	-10.306	-31.005	4.147	7.412	8.732	8.854	8.854
R^2 Within Physician	0.010	0.014	0.016	0.019	0.022	0.022	0.022	0.022
R_{max}	0.012	0.018	0.021	0.025	0.028	0.028	0.028	0.028
Clustering	State	State	State	State	State	State	State	State
Phys. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	No	No	No	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	No	No	No	No	Yes	Yes	Yes	Yes
Pop. Density Contr.	No	No	No	No	No	Yes	Yes	Yes
MD per capita Contr.	No	No	No	No	No	No	Yes	Yes

This table shows estimates of equation (8) with various sets of controls. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move and the difference between our fee and recovery rate indices estimated from the post-move and pre-move states. Columns 2 and following add additional controls for characteristics of the state or county in which the physician is located. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the physician fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within physician (shown as R_{max}) relative to the observed within-physician R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table B.2: Mover Results: Recovery Rate Definition 3

	Accept Medicaid Patients?							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-Move × Fee	0.1075***	0.1034***	0.1146***	0.0709***	0.0667***	0.0624***	0.0629***	0.0629***
	(0.0211)	(0.0194)	(0.0196)	(0.0140)	(0.0128)	(0.0129)	(0.0130)	(0.0130)
Post-Move × Recovery Rate	0.1683***	0.1722***	0.1612***	0.1104***	0.1309***	0.1387***	0.1395***	0.1395***
	(0.0346)	(0.0333)	(0.0339)	(0.0321)	(0.0371)	(0.0396)	(0.0401)	(0.0401)
N	56893	56893	56893	56893	56893	56893	56893	56893
R^2	0.600	0.602	0.603	0.604	0.605	0.605	0.605	0.605
$\hat{\gamma}/\hat{\beta}$	1.565	1.666	1.407	1.557	1.962	2.222	2.218	2.218
	(0.230)	(0.244)	(0.218)	(0.382)	(0.448)	(0.518)	(0.513)	(0.513)
δ s.t. Post-Move × Fee = 0	-8.429	-43.302	-16.835	2.622	2.359	2.010	2.022	2.022
δ s.t. Post-Move × Recovery Rate = 0	-8.090	-10.917	-46.131	4.040	6.568	7.629	7.738	7.738
R^2 Within Physician	0.010	0.014	0.016	0.019	0.022	0.022	0.022	0.022
R_{max}	0.013	0.019	0.021	0.025	0.028	0.028	0.028	0.028
Clustering	State	State	State	State	State	State	State	State
Phys. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	No	No	No	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	No	No	No	No	Yes	Yes	Yes	Yes
Pop. Density Contr.	No	No	No	No	No	Yes	Yes	Yes
MD per capita Contr.	No	No	No	No	No	No	Yes	Yes

This table shows estimates of equation (8) with various sets of controls. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move and the difference between our fee and recovery rate indices estimated from the post-move and pre-move states. Columns 2 and following add additional controls for characteristics of the state or county in which the physician is located. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the physician fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within physician (shown as R_{max}) relative to the observed within-physician R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table B.3: Mover Robustness to Different Indices: Recovery Rate Definition 1

	Accept Medicaid Patients?						
	Standard Index	Shrinkage Index	First Visit Excluded	First Visit Only	Local Income Control	Pregnant Patients Only	Index w/o Physician FE
Post-Move × Fee	0.0607***	0.0607***	0.0587***	0.0611***	0.0604***	0.0562***	0.0594***
	(0.0135)	(0.0135)	(0.0137)	(0.0134)	(0.0135)	(0.0130)	(0.0128)
Post-Move × Recovery Rate	0.1411***	0.1414***	0.1534***	0.1231***	0.1392***	0.0874**	0.1456***
	(0.0381)	(0.0381)	(0.0321)	(0.0419)	(0.0381)	(0.0420)	(0.0461)
N	56893	56893	56893	56893	56893	56893	56893
R^2	0.605	0.605	0.605	0.605	0.605	0.605	0.605
$\hat{\gamma}/\hat{\beta}$	2.324	2.328	2.614	2.013	2.305	1.555	2.451
	(0.557)	(0.558)	(0.606)	(0.519)	(0.558)	(0.513)	(0.631)
δ s.t. Post-Move × Fee = 0	1.917	1.917	1.823	1.926	1.900	1.669	1.847
δ s.t. Post-Move × Recovery Rate = 0	4.276	4.275	3.729	5.116	3.972	2.558	3.708
R^2 Within Physician	0.022	0.022	0.022	0.022	0.022	0.021	0.022
R_{max}	0.028	0.028	0.029	0.028	0.028	0.028	0.028
Clustering	State	State	State	State	State	State	State
Phys. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MD per capita Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table shows estimates of equation (8) with various of our indices. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move and the difference between our fee and recovery rate indices estimated from the post-move and pre-move states. Columns 2 and following construct the indices using different samples or adjustments indicated at the top of each column. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the physician fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within physician (shown as R_{max}) relative to the observed within-physician R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table B.4: Mover Robustness to Different Indices: Recovery Rate Definition 2

	Accept Medicaid Patients?						
	Standard Index	Shrinkage Index	First Visit Excluded	First Visit Only	Local Income Control	Pregnant Patients Only	Index w/o Physician FE
Post-Move × Fee	0.0631***	0.0631***	0.0627***	0.0622***	0.0628***	0.0561***	0.0605***
	(0.0132)	(0.0132)	(0.0135)	(0.0130)	(0.0132)	(0.0128)	(0.0127)
Post-Move × Recovery Rate	0.1370***	0.1374***	0.1459***	0.1194***	0.1345***	0.0798**	0.1208**
	(0.0421)	(0.0421)	(0.0361)	(0.0444)	(0.0419)	(0.0394)	(0.0485)
N	56893	56893	56893	56893	56893	56893	56893
R^2	0.605	0.605	0.605	0.605	0.605	0.605	0.605
$\hat{\gamma}/\hat{\beta}$	2.173	2.179	2.325	1.919	2.141	1.424	1.998
	(0.514)	(0.515)	(0.524)	(0.496)	(0.512)	(0.476)	(0.555)
δ s.t. Post-Move × Fee = 0	2.025	2.026	2.022	1.973	2.013	1.662	1.889
δ s.t. Post-Move × Recovery Rate = 0	8.854	8.845	6.043	16.345	7.768	4.277	6.078
R^2 Within Physician	0.022	0.022	0.022	0.022	0.022	0.021	0.022
R_{max}	0.028	0.028	0.028	0.028	0.028	0.028	0.028
Clustering	State	State	State	State	State	State	State
Phys. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MD per capita Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table shows estimates of equation (8) with various of our indices. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move and the difference between our fee and recovery rate indices estimated from the post-move and pre-move states. Columns 2 and following construct the indices using different samples or adjustments indicated at the top of each column. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the physician fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within physician (shown as R_{max}) relative to the observed within-physician R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table B.5: Mover Robustness to Different Indices: Recovery Rate Definition 3

	Accept Medicaid Patients?						
	Standard Index	Shrinkage Index	First Visit Excluded	First Visit Only	Local Income Control	Pregnant Patients Only	Index w/o Physician FE
Post-Move × Fee	0.0629***	0.0629***	0.0626***	0.0620***	0.0626***	0.0556***	0.0589***
	(0.0130)	(0.0130)	(0.0133)	(0.0128)	(0.0130)	(0.0128)	(0.0127)
Post-Move × Recovery Rate	0.1395***	0.1397***	0.1423***	0.1272***	0.1368***	0.0830**	0.1218**
	(0.0401)	(0.0401)	(0.0339)	(0.0433)	(0.0400)	(0.0402)	(0.0473)
N	56893	56893	56893	56893	56893	56893	56893
R^2	0.605	0.605	0.605	0.605	0.605	0.605	0.605
$\hat{\gamma}/\hat{\beta}$	2.218	2.222	2.273	2.052	2.185	1.492	2.067
	(0.513)	(0.514)	(0.509)	(0.510)	(0.511)	(0.491)	(0.576)
δ s.t. Post-Move × Fee = 0	2.022	2.023	2.017	1.970	2.008	1.639	1.816
δ s.t. Post-Move × Recovery Rate = 0	7.738	7.732	5.837	10.963	6.946	3.895	5.057
R^2 Within Physician	0.022	0.022	0.022	0.022	0.022	0.021	0.022
R_{max}	0.028	0.028	0.029	0.028	0.028	0.028	0.028
Clustering	State	State	State	State	State	State	State
Phys. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MD per capita Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table shows estimates of equation (8) with various of our indices. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move and the difference between our fee and recovery rate indices estimated from the post-move and pre-move states. Columns 2 and following construct the indices using different samples or adjustments indicated at the top of each column. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the physician fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within physician (shown as R_{max}) relative to the observed within-physician R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table B.6: Cross-State Group Results: Alternate Group Definition

	Accept Medicaid Patients?						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Medicaid Fee	0.0926**	0.0960**	0.1232***	0.0851*	0.0729*	0.0678	0.0727*
	(0.0418)	(0.0411)	(0.0390)	(0.0422)	(0.0407)	(0.0407)	(0.0392)
Medicaid Recovery Rate	0.2610***	0.2490***	0.2414***	0.1770*	0.1906**	0.1922**	0.2142**
	(0.0887)	(0.0907)	(0.0874)	(0.0882)	(0.0866)	(0.0889)	(0.0965)
N	14352	14352	14352	14352	14352	14352	14352
R^2	0.206	0.207	0.214	0.217	0.218	0.218	0.219
$\hat{\gamma}/\hat{\beta}$	2.819	2.593	1.958	2.080	2.614	2.836	2.946
	(0.887)	(0.822)	(0.581)	(0.881)	(1.159)	(1.312)	(1.250)
δ s.t. Medicaid Fee $= 0$	3814.199	-53.365	-10.025	7.357	4.170	3.420	3.869
δ s.t. Medicaid Recovery Rate $= 0$	96.562	27.236	16.546	3.798	4.417	4.467	5.401
R^2 Within Group	0.007	0.008	0.018	0.021	0.022	0.023	0.023
R_{max}	0.009	0.011	0.023	0.027	0.029	0.029	0.030
Clustering	State	State	State	State	State	State	State
Group Definition	SK&A Group and System augmented with TIN						
Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	No	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	No	No	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	No	No	No	Yes	Yes	Yes	Yes
Pop. Income Contr.	No	No	No	No	Yes	Yes	Yes
Pop. Density Contr.	No	No	No	No	No	Yes	Yes
MD per capita Contr.	No	No	No	No	No	No	Yes

This table shows estimates of equation (10) with various sets of controls. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who practice in groups that have locations in multiple states. The key independent variables are fee and recovery rate indices in the state where the physician is located. Columns 2 and following add additional controls for characteristics of the state or county in which the physician is located. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the practice fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within group (shown as R_{max}) relative to the observed within-group R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.

Table B.7: Cross-State Group Results: Robustness to Different Indices

	Standard Index	Shrinkage Index	First Visit Excluded	First Visit Only	Local Income Control	Pregnant Patients Only	Index w/o Physician FE
				Accept Medicaid Patients?			
Medicaid Fee	0.0727*	0.0725*	0.0690*	0.0749*	0.0723*	0.0654	0.0692*
	(0.0392)	(0.0392)	(0.0392)	(0.0394)	(0.0393)	(0.0393)	(0.0393)
Medicaid Recovery Rate	0.2142**	0.2159**	0.1634*	0.2439**	0.2079**	0.1979**	0.2911***
	(0.0965)	(0.0986)	(0.0871)	(0.1015)	(0.0962)	(0.0789)	(0.1049)
N	14352	14352	14352	14352	14352	14352	14399
R^2	0.219	0.219	0.218	0.219	0.219	0.219	0.219
$\hat{\gamma}/\hat{\beta}$	2.946	2.977	2.369	3.257	2.876	3.024	4.205
	(1.250)	(1.273)	(1.129)	(1.308)	(1.238)	(1.380)	(1.743)
δ s.t. Medicaid Fee $= 0$	3.869	3.848	3.499	4.094	3.825	3.170	4.276
δ s.t. Medicaid Recovery Rate $= 0$	5.401	5.308	3.727	7.943	4.648	20.341	14.540
R^2 Within Group	0.023	0.023	0.023	0.023	0.023	0.023	0.024
R_{max}	0.030	0.030	0.030	0.030	0.030	0.030	0.031
Clustering	State	State	State	State	State	State	State
Group Definition	SK&A Group and System augmented with TIN						
Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phys. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dem. Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Insurance Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Income Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Density Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MD per capita Contr.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table shows estimates of equation (10) with various of our indices. Each column represents a separate regression, and the dependent variable in all cases is the physician's reported willingness to treat Medicaid patients. The sample is all physicians who moved across states during the period in which we observe them, and the key independent variables are interactions between an indicator for years after the move and the difference between our fee and recovery rate indices estimated from the post-move and pre-move states. Columns 2 and following construct the indices using different samples or adjustments indicated at the top of each column. Under each regression, we report the sample size (physician-year observations), regression R^2 (including the physician fixed effects), ratio of the recovery rate coefficient (labeled $\hat{\gamma}$) to fee coefficient (labeled $\hat{\beta}$), and results of the Oster (2019) test for coefficient stability. Oster's δ shows the ratio of selection on unobservables that would be needed relative to the selection on observables that we estimate in order to drive the respective coefficient to zero. The test relies on an assumption about the maximum possible R^2 within physician (shown as R_{max}) relative to the observed within-physician R^2 (also shown); we use Oster's proposed value of $R_{max} = 1.3R^2$.