# Do Health Insurers Innovate?
# Evidence from the Anatomy of Physician Payments

Jeffrey Clemens
UCSD and NBER

Joshua D. Gottlieb
UBC and NBER

Tímea Laura Molnár*
UBC

July 2, 2017

**Abstract**

One of private health insurers' main roles in the United States is to negotiate physician payment rates on their beneficiaries' behalf. We show that these rates are often set in reference to a government benchmark, and ask how often private insurers customize their fee schedules away from this default. We exploit changes in Medicare's payments and dramatic bunching in markups over Medicare's rates to address this question. Although Medicare's rates are influential, 25 percent of physician services in our data, representing 45 percent of covered spending, deviate from the benchmark. Heterogeneity in the pervasiveness and direction of deviations suggests that the private market coordinates around Medicare's pricing for simplicity but abandons it when sufficient value is at stake.

*Clemens: jeffclemens@ucsd.edu, Gottlieb: Joshua.Gottlieb@ubc.ca, Molnár: molnar.timea.laura@alumni.ubc.ca. We are extremely grateful to Luisa Franzini, Cecilia Ganduglia-Cazaban, Osama Mikhail, and the UTSPH/BCBSTX Payment Systems and Policies Research Program at the University of Texas School of Public Health for data access, and for their extensive assistance in navigating the BCBSTX claims data. Clemens and Gottlieb thank the Stanford Institute for Economic Policy Research and the Federal Reserve Bank of San Francisco for their hospitality. We thank SSHRC, Jon Skinner and the Dartmouth Institute for support, and Victor Saldarriaga and Andrew Vogt for excellent research assistance. Finally, we thank Leila Agha, Liran Einav, Jeff Emerson, Amy Finkelstein, John Friedman, Matt Gentzkow, R.B. Harris, David Laibson, Neale Mahoney, Matt Panhans, Maria Polyakova, Jon Skinner, Ashley Swanson, numerous seminar participants, three referees and the editor Luigi Siciliani for helpful comments.

Health insurers have a powerful ability to shape the efficiency of health care delivery. Insurers straddle the relationship between patients and medical providers, and enter into contracts with both sides of the market. Consumers or employers purchase insurance plans whose copayments and deductibles influence subsequent demand for care. At the same time, contracts with physicians and hospitals govern how these providers will be compensated for treating insured patients, and hence the caregivers' financial incentives.

The literature on optimal consumer cost-sharing is long and well-developed (Feldstein, 1973; Besley, 1988). Only recently, however, has an empirical literature begun to explore how private insurers set copayments in practice. Einav, Finkelstein and Polyakova (2016) show that private insurers provide more risk-protection for drugs subject to less moral hazard. They contrast this with public insurance plans, which offer relatively uniform coverage with regards to cost-sharing. Starc and Town (2015) show that insurers responsible for patients' non-pharmaceutical spending provide more generous coverage for drugs that can keep people out of the hospital.

We investigate whether insurers apply a similar logic to the payment schedules they negotiate with providers. To what extent do private insurance carriers adopt Medicare's cost-based approach to physician payment? Looking at the flip side of this question, we analyze the extent to which private insurers customize their physician reimbursements relative to Medicare's industry standard.

Despite recent high-level changes in the U.S. health insurance market, the incentive structure through which physicians are paid remains predominantly "fee for service."[1] A physician's income depends on the quantity and intensity of the treatment she provides—even when part of a larger managed care plan or "accountable care organization" (Zuvekas and Cohen, 2016).[2] A growing body of evidence finds that these high-powered incentives

[1]Data from the 2004-2005 Community Tracking Study (CSHSC 2006, 46) show that 52 percent of physicians earn zero revenue from capitated contracts, and 79 percent earn less than a quarter of their revenue from such contracts.

[2]Among those same physicians referenced in footnote 1, who earn little from capitated contracts, 65

for incremental care provision help drive the level and composition of medical spending (McClellan, 2011).

The structure of physician payments is a potentially powerful tool for insurers to encourage more efficient care. Relatively little is known, however, about the extent to which private insurers customize their fee-for-service payments for this purpose. Clemens and Gottlieb (2017) find that private payments rise and decline quite strongly with Medicare's payments, which raises the question of whether private insurers' payments are meaningfully independent from Medicare's rate schedule. In this paper, we shed light on this question using detailed physician payment data from a large insurer.

Medicare compensates physicians and outpatient providers through a detailed fee-for-service pricing system. Physicians submit bills for each instance in which they provide any of 13,000 recognized services. The system assigns each service a certain number of Relative Value Units, which determine the payment. These relative values aim to measure average cost but not medical value. This procurement model thus has little capacity to steer treatment towards effective—let alone cost-effective—care. It has particular difficulty managing the use of capital-intensive diagnostic imaging services, for which average-cost payments significantly exceed providers' marginal costs—as they must in order to facilitate entry.

Private reimbursement arrangements are less transparent than Medicare's. To peer into the black box of these business-to-business contracts, we begin by developing a cross-sectional method for systematically assessing whether payments are benchmarked to Medicare's rate structure. Our first approach involves a classification algorithm motivated by the bunching literature.[3] Using the outpatient claims data of Blue Cross Blue Shield of Texas (BCBS-TX), we begin by computing the ratio of each private payment to the Medicare payment

percent earn more than one quarter of their revenue from managed care (CSHSC 2006, 47). CSHSC (1999) reports similar estimates from 1996-1997.

[3]Our setting differs from standard bunching applications in that the bunching we observe is not driven by kinks or notches in budget sets (Kleven, 2016). Instead, it results from clustering around reference points.

for that service. Within the payments to individual physician groups, the distributions of these ratios reveal spikes that indicate exceptionally common markups. We use these spikes to identify which payments are likely benchmarked to Medicare's relative rate structure.

We complement this cross-sectional method with an analysis of updates to Medicare's structure of relative payments. If the Medicare links we identify are accurate, then payments for Medicare-benchmarked services should update when Medicare's schedule of Relative Value Units is revised. We are able to assess this pass-through at a high frequency by applying institutional knowledge of the exact dates on which BCBS-TX implements Medicare's updates to the relative value scale. The relatively high frequency at which we can conduct our analysis allows us to limit, if not eliminate, concerns about potential confounders including active contract renegotiations and payment changes connected to substantive technological advances. We find that the payments associated with 55 percent of in-network, outpatient spending (and around three quarters of services) are linked to Medicare. These estimates are quite similar to those we obtain using our cross-sectional bunching approach.

We continue our analysis with an effort to understand the circumstances under which payments are more likely, or less likely, to be benchmarked to Medicare's relative rate structure. Deviations from benchmarking exhibit several distinctive patterns. Looking across physician groups, payments to relatively large groups are less tightly benchmarked to Medicare than payments to small groups. Payments for only ten percent of services provided by the smallest firms, representing 20 percent of their spending, deviate from Medicare's relative values. The same is true of 40 percent of services—and two-thirds of spending—from firms with total billing exceeding $1 million per year.

Looking across service categories, payments are more likely to deviate from Medicare's relative values for capital-intensive services, like diagnostic imaging, than for labor-intensive services like standard office visits. Payments for roughly 45 percent of imaging services, but only 15 percent of evaluation and management services, deviate from Medicare's menu.

Within imaging, Medicare distinguishes between two types of services: a capital-intensive component for taking the image and a labor-intensive component for interpreting the image. Medicare explicitly amortizes the fixed cost of the imaging equipment into the former. We find that private insurers' payments for interpretation are far less likely to deviate from Medicare rates than payments for taking the image itself. The directions of these deviations reveal that the adjustments narrow likely gaps between marginal costs and Medicare's average-cost payments. We find that payments for labor-intensive services tend to be adjusted up while payments for capital-intensive services are adjusted down.

One plausible interpretation of these findings emphasizes the complexity of the insurer-physician contracting environment. Specifically, there is a tension between gains from fine-tuning payments and costs from making contracts complex. To manage this tension, insurers may draw on Medicare's relative value scale for the purpose of contract simplification, while strategically adapting their contracts where the value is highest. This view is consistent with the heterogeneity we observe: the benefits of fine-tuning payments will tend to be largest within contracts with large physician groups and for the capital-intensive services for which Medicare's average-cost payments deviate most from marginal cost. A complementary explanation is that large firms may use their bargaining power to obtain high service-specific markups rather than, or in addition to, high overall payments per unit of care. The information content of the relative value scale on which Medicare's payments rely can also be interpreted as a knowledge standard and, more generally, as a public good.

Our results are relevant in two broader contexts. Learning how prices are set in health care—a sector comprising 18 percent of the economy—is essential for understanding macroeconomic price-setting dynamics.[4] The service sector in general (Nakamura and Steinsson, 2008), and medical care in particular (Bils and Klenow, 2004), have especially sticky prices.

[4]Clemens, Gottlieb and Shapiro (2014, 2016) show how much Medicare price regulation can impact overall inflation.

We provide evidence on how this stickiness arises.[5] Consistent with Anderson, Jaimovich and Simester's (2015) evidence from retail, the complexity of physician contracting may explain both the long duration of these prices and the public-private linkages we identify.

Public policies' residual influence on private firms is relevant in a wide range of contexts. Outside of the health care context, labor contracts sometimes benchmark wage rates to the statutory minimum.[6] Within the health sector, Medicare has been found to shape aspects of private players' behavior in the pharmaceutical, hospital, and physician marketplaces (Duggan and Scott Morton, 2006; Alpert, Duggan and Hellerstein, 2013; White, 2013; Clemens and Gottlieb, 2017). The forces we investigate here differ conceptually from those analyzed in prior work, including Clemens and Gottlieb's (2017) analysis of physician payments. Clemens and Gottlieb (2017) analyze Medicare's influence on physicians' bargaining positions through its effects on their incentives and outside options. The current paper emphasizes the Medicare payment model's role as an industry standard. More specifically, it emphasizes the Medicare payment model's role as a benchmark or default around which contract negotiations coordinate. Our analysis provides insights into the overall pervasiveness of benchmarking against Medicare's relative cost schedule and into the types of contracts in which customization is most prevalent.

We continue in section 1 by presenting institutional background on price setting in U.S. physician markets. Section 1 concludes with a discussion of several potential explanations for the benchmarking phenomenon we examine. Section 2 introduces the claims data we analyze. Section 3 presents the empirical strategies we implement, while sections 4 and 5 present our results. Section 6 briefly concludes.

[5] In particular, our empirical evidence supports price-setting mechanisms with the flavour of Christiano, Eichenbaum and Evans (2005) or Smets and Wouters (2003, 2007).

[6] A publicly posted contract template of the United Food and Commercial Workers Union (2002), for example, includes the requirement that "At no time during the life of this Agreement will any of the bagger/carry-out rates be less than twenty-five ($0.25) cents an hour above the Federal minimum wage."

# 1 Medical Pricing Institutions

Public and private payments for health care services are set through very different mechanisms. Medicare reimbursements are set based on administrative estimates of the resource costs of providing care, which we describe in section 1.1. For patients with private health insurance, providers' reimbursements are determined through negotiations between the insurers and providers, which we describe in section 1.2. Section 1.3 discusses several potentially complementary economic rationales for a link between reimbursement rates across these two segments of the market.

## 1.1 Medicare Price Determination[7]

In 1992, Congress established a system of centrally administered prices to reimburse physicians and other outpatient providers. This Resource-Based Relative Value Scale (RBRVS) is a national fee schedule that assigns a fixed number of Relative Value Units (RVUs) to each of 13,000 distinct health care services. Legislation specifies that the RVUs for service $j$ are supposed to measure the resources required to provide that service. Since the costs of intermediate inputs differ across the country, RBRVS incorporates local price adjustments, called the Geographic Adjustment Factor (GAF), to compensate providers for these differences. The payment for service $j$ to a provider in geographic region $i$ is approximately:

$$\begin{aligned}\text{Reimbursement rate}_{i,j,t} = \text{Conversion Factor}_t &\times \text{Geographic Adjustment Factor}_{i,t} \\ &\times \text{Relative Value Units}_{j,t}. \qquad (1)\end{aligned}$$

The "reimbursement rate," a term we use interchangeably with "price," is the amount Medicare pays for this service. The Conversion Factor (CF) is a national scaling factor, usually

[7] This section draws from Clemens and Gottlieb (2014).

updated annually.

Variation in Medicare's payment rates is driven primarily by the number of RVUs assigned to a service. This assignment is constant across areas while varying across services. Medicare regularly updates the RVUs assigned to each service, primarily based on input from the American Medical Association, using the formal federal rule-making process. These updates are intended to account for technological and regulatory changes that alter a service's resource intensity. We exploit these changes in the empirical strategy described in section 3.

## 1.2 Private Sector Price Setting

U.S. private sector health care prices are set through negotiations between providers and private insurers.[8] The details of these negotiations are not transparent, and our limited knowledge about private sector prices comes from claims data that reveal the reimbursements paid once care is provided.[9] A common feature of physician contracts, central to both our theoretical and empirical analyses, is a form of benchmarking to Medicare.

Practitioners emphasize that Medicare's administrative pricing menu features prominently in private insurers' contracts. Newsletters that insurers distribute to participating providers frequently draw explicit links between Medicare's fee schedule and the insurer's maximum allowable charges. For example, reimbursement rates might be linked to Medicare by default unless the contract specifies otherwise. But the relative value scale does not determine an absolute price level. As in Medicare, computing private reimbursements requires multiplying RVUs by a dollar scaling factor. Practitioners describe physician contracts as involving negotiations over markups relative to Medicare, combined with payments for particular services or service bundles (Nandedkar, 2011; Gesme and Wiseman, 2010; Mertz,

[8]In rare exceptions, such as in Maryland, the state government determines all hospital payment rates.

[9]A growing literature finds that physician concentration significantly affects this bargaining process. Payments are higher in markets where physicians are more concentrated (Dunn and Shapiro, 2014; Baker, Bundorf, Royalty and Levin, 2014; Kleiner, White and Lyons, 2015; Clemens and Gottlieb, 2017).

2004). Our empirical work examines the pervasiveness of this benchmarking phenomenon and the circumstances under which customization occurs.

## 1.3 Potential Rationales

Why might contracts between physicians and private insurers use Medicare's relative rate structure as a benchmark? We consider several explanations, which are broadly complementary. A first explanation is that benchmarking to Medicare's relative rate structure enables insurers and physicians to greatly simplify their contracts. A fully benchmarked contract requires negotiating over a single parameter—the markup or scaling factor. Alternative contract structures could require negotiating payments for hundreds or thousands of distinct billing codes. Medicare's payment model may serve as an industry-standard benchmark with which all parties are familiar. In Appendix A, we present a formal model of this idea, which generates predictions that our empirical analysis supports.

A second, strongly complementary explanation involves the Medicare schedule's informational content. By design, Medicare's payment model contains substantial information about the relative costs of providing physicians' services. If average-cost reimbursement is more or less what insurers desire to implement, Medicare's payment model provides useful information for private insurers to adopt. Put differently, Medicare's relative cost estimates can be interpreted as a public good. Although they may fail to reflect variations in local cost structures, the expense to insurers of independently calculating these costs may be high.

A third possibility is that providing care for Medicare beneficiaries represents physicians' primary outside option when they negotiate with private insurers (Clemens and Gottlieb, 2017). Because Medicare accounts for a large share of the market, its payments may loom large in insurer-physician negotiations. Benchmarking private payments to Medicare's payments may be a straightforward way for contracts to acknowledge and readily adjust to changes in the value of that alternative.

A fourth possibility emphasizes insurance regulations. Regulations require insurers to ensure access to "medically necessary" services. Benchmarking payments to Medicare's rate structure may be the easiest approach to satisfying this requirement.[10] Private payments are almost universally marked up rather than marked down relative to Medicare's rates. Such a payment structure ensures that private insurers are paying sufficiently high rates to generate at least as much care access as Medicare beneficiaries enjoy.

# 2 Medical Pricing Data

Our main analysis considers firm-to-firm pricing in the context of medical claims processed by one large insurer, Blue Cross Blue Shield of Texas (BCBS). BCBS is by far the largest carrier in Texas, commanding over 40 percent of each insurance market segment (Kaiser Family Foundation, 2015). Our main database covers the universe of BCBS's payments for outpatient care in 2010; we expand our sample to cover 2008–2011 for one analysis.[11] For each claim, the database details the treatment provided, location, physician, physician group, and BCBS's payment to that group. We restrict this universe along several dimensions. The full 2010 dataset contains 57,613,494 claim lines and $4.29 billion in spending, which we clean as described in Appendix B.1. This initial cleaning, which eliminates payments made to out-of-network physicians (who have not reached a negotiated agreement with BCBS on reimbursement rates) leaves us with 44,055,829 service lines and $2.63 billion of spending. We will subsequently examine the other segment of the data separately.

We next merge the remaining claims with Medicare billing codes. In order for private insurers to benchmark prices to Medicare, those services would need to be billed using

[10]Beneficiaries may lack access to care if payment rates are too low to induce physicians to treat them. The Medicaid program's low payment rates, for example, are often linked to the possibility that beneficiaries will lack access to essential services.

[11]Our empirical results for other years are very similar to those for 2010. We focus on this one year for brevity and show other years' results in the appendix.

Medicare's billing codes. The services we cannot merge are thus clearly not benchmarked to Medicare's relative value scale. The merge retains over 97 percent of claims for evaluation and management, diagnostic imaging, and surgical services. We lose notable portions of one broad spending category, namely laboratory tests, for which both Medicare and BCBS frequently base payments on alternative codes. The remaining analysis sample includes 3,681 unique codes from the Healthcare Common Procedure Coding System (HCPCS), which comprise 23,933,577 service lines and $2.05 billion of spending.[12]

The claims data also allow us to describe the provider groups serving BCBS beneficiaries, at least in terms of the care they provide to that sample. To enable our subsequent investigation of heterogeneity in Medicare benchmarking, we measure the total value of the care each group provides to BCBS patients in a given year. Our final dataset, which is summarized in Table 1, includes care provided by over 80,000 physician groups as identified by their billing identification number.[13] 15,000 of these groups bill more than $10,000 annually and account for 97 percent of BCBS spending. These 15,000 groups filed an average of just under 1,000 claims, had an average of 4 physicians, and saw an average of just over 400 patients.

# 3 Empirical Approach

Our primary empirical goal is to estimate the frequency with which private reimbursement rates are benchmarked directly to Medicare's relative rate structure. We begin by presenting visually striking evidence of bunching in the ratios of physician groups' payments relative to Medicare's payments. We then develop an approach to formalizing this visual evidence in section 3.1. Next, we present an empirical approach for exploiting policy-driven changes

[12] The HCPCS coding system is used by Medicare and many private insurers. The set of codes includes those developed for the Current Procedural Terminology (CPT) system by the American Medical Association.
Appendix Table B.1 shows the exact data loss resulting from each step of cleaning. The key conclusion from this table is that, once we restrict ourselves to the relevant universe of data, additional losses from merging in Medicare codes and eliminating infrequent codes are not substantial.

[13] This is analogous to the commonly used tax ID number in Medicare claims data, but our version is anonymized.

to Medicare's Relative Value Units (RVUs) in section 3.2. Finally, section 3.3 relates these approaches.

## 3.1 Measuring Medicare Benchmarking with Bunching

We begin our empirical analysis by examining the relationship between private and Medicare pricing in the cross-section. To do so, we first divide BCBS's payment to group $g$ for service $j$ at time $t$ ($P_{g,j,t}$) by Medicare's allocation of RVUs to that service at that time. This defines an "Implied Conversion Factor" (ICF) as:

$$ICF_{g,j,t} = \frac{P_{g,j,t}}{RVU_{j,t}}. \tag{2}$$

While an ICF is defined for every claim, simply computing the ICF does not tell us whether the claim was, in fact, contractually benchmarked as a markup relative to Medicare's RVUs. To gauge the prevalence of contractually specified benchmarking, we analyze the regularity with which a particular group's payments reflect *the same* ICF. Specifically, we investigate the prevalence of sharp bunching in the ratio of a group's payments relative to Medicare.

Graphical illustrations, as presented in Figure 1, can help to build intuition regarding our bunching approach's strengths and weaknesses. Panel A shows payment rates for the services provided regularly by a single physician group in the 2010 BCBS claims data.[14] Each circle on the graph is a unique payment amount for a unique service code. That is, if the group sometimes received \$45 for a standard office visit (HCPCS code 99213), and other times received \$51, those two amounts would show up as separate circles. The Blue Cross payment amount is on the $y$-axis and the Medicare payment for the service is on the $x$-axis, both shown on log scales. Taking logs of equation (2) reveals that Medicare-linked pricing

[14] The figures exclude any code-by-payment combination that appears ten times or fewer in the data for the relevant physician group. The more systematic analysis presented below has no such exclusion. Throughout this analysis, we restrict to data from the period before BCBS implemented the RVU updates (January 1–June 30, 2010). This way our calculations are not confounded by RVU changes.

implies a one-for-one relationship between the log Medicare payment and the log private reimbursement.[15] Panel A shows the data from a mid-sized group (billing BCBS between \$200,000 and \$1 million in 2010) for which a single ICF dominates the payment picture. The most natural interpretation of this graph is that those services on the solid line are priced according to Medicare's fee schedule with a common ICF, while the remaining services are priced separately. Several of the circles below the solid line plausibly involve instances of a less common, but still contractually specified, ICF for this group. A conservative estimate would view these and other circles off the solid line as deviations from Medicare-linked pricing.

Panel B shows the full distribution of this group's markups relative to Medicare's rates. To calculate these markups, we simply divide the $y$-value of each dot in Panel A (the private payment) by its $x$-value (the Medicare rate).[16] Panel B shows a clear spike in the distribution of these ratios at around 1.4, indicating that most claims were paid based on a 40 percent markup over Medicare. This spike includes all of the services along the red line in Panel A. Other scattered values in the histogram reflect the deviations away from that line.

Panels C through F show graphs constructed analogously, but for two larger groups that provide more unique services at more distinct prices. The group shown in Panels C and D exhibits two clear spikes in the ICF frequency distribution, with a smattering of other values. The group shown in Panels E and F has a range of ICFs, none of which visually dominates the payment picture. These plots indicate a remarkably complicated contract with BCBS.

Estimating the pervasiveness of "common" ICFs requires a definition of "common."

[15]Rearranging (2) and then taking logs yields:

$$\ln(P_{g,j,t}) = \ln(ICF_{g,j,t}) + \ln(RVU_{j,t}), \tag{3}$$

which has an implied coefficient of 1 on $\ln(RVU_{j,t})$. The $y$-intercept (in logs) is simply the log of the ICF.

[16]This distribution has the same sample restrictions as in Panel A; see footnote 14 for details. Note that each observation from Panel A has equal weight in the distribution in Panel B, so the distributions in Panels B, D, and F are not weighted to reflect the frequency with which we observe each markup. A weighted version would increase the relative heights of the highest bars, since the common ICFs are, by definition, more common than other markups.

When presenting our results, we will explore sensitivity to the threshold we impose for the frequency with which an ICF must appear in a group's payments. This also requires an assumption on our rounding of the ratio of private to public payments. We explore sensitivity to the choice of rounding as well.

In addition to estimating the share of ICFs that are benchmarked, we run descriptive regressions to investigate correlates of average mark-ups. That is, we estimate

$$\ln(ICF_{g,j}) = \boldsymbol{X_{g,j}}\gamma + e_{g,j} \tag{4}$$

where $\boldsymbol{X_{g,j}}$ contains characteristics of the physician group or local market, such as firm size or concentration. We measure firm size as log total billings to the insurer. We compute firm market share within a local health care market (hospital service area) and specific service, and we measure the degree of concentration across all physician practices within that market (using the HHI at the service-by-area level). We estimate this equation at the claim level, and compute clustered standard errors that allow $e_{g,j}$ to have arbitrary correlation within clusters at the physician group level, since that is the level at which the ICFs are negotiated.

## 3.2 Framework for Analyzing Benchmarking Using RVU Updates

We next develop an estimation framework based on changes in Medicare's relative value scale. A committee of the American Medical Association, composed of representatives of various physician specialties, recommends RVU updates to Medicare (Government Accountability Office, 2015). These updates come in two main forms: reassessments of the resources required to provide a single service, and revisions to part of the underlying methodology. For example, a revision to the method for computing physician effort can change the weights assigned to many service codes. At least one broad update of this sort appears to occur annually over the period we study, as do hundreds of larger service-specific reassessments.

The vast majority of updates to Medicare payments go into effect on January 1 each year. But even when relying on these rates, private insurers have a choice about whether and when to shift from one year's relative value scale to the next year's (Borges, 2003). BCBS informs its providers of the date on which such updates go into effect through its provider newsletter, the *Blue Review*. During the period of our primary analysis sample, the newsletter announced updates taking place on July 1, 2010 (BCBS 2010).[17]

Panel A of Figure 2 presents an example of how these changes impact physician payments in our BCBS data. This graph shows average log payments by day for the most commonly billed service code, a standard office visit with an established patient (code 99213). The average log payment jumps distinctively on July 1, 2010, the day on which BCBS implemented the 2010 relative values. Medicare's log RVUs for this service rose by 0.068 between the 2009 and 2010 fee schedules. BCBS's average log payment rose by just under 0.05. To study BCBS's payment updates systematically, we next develop a method for using high frequency payment changes to infer the share of private reimbursements linked to Medicare.

When a payment $P_{g,j,t}$ is linked to Medicare's relative values, we can write

$$P_{g,j,t} = \varphi_{g,t} \cdot RVU_{j,t} \tag{5}$$

or, taking logs,

$$\ln(P_{g,j,t}) = \ln(\varphi_{g,t}) + 1 \cdot \ln(RVU_{j,t}), \tag{6}$$

where $\varphi_{g,t}$ is the Implied Conversion Factor (ICF) from section 3.1. Equation (6) describes a linear relationship between log private insurance payments and log RVUs for a service. It describes the one-for-one relationship between log RVUs and log private payments that obtains when contracts are specified in this manner. If the markup is a constant, it will be reflected in the constant term of a regression version of (6). If the mark-up varies across

[17]For one empirical extension, we incorporate additional payment changes implemented on July 1, 2008, on August 15, 2009, and on September 1, 2011 (BCBS 2008; 2009; 2011). In all four years, the standard deviations of changes in log RVUs are around 0.07, or approximately 7 percent. This means that the Medicare changes contain sufficient variation for us to exploit and generate reasonably precise estimates.

physician groups, then it will be captured by group fixed effects. If it varies both across groups and across time, then it will be captured by group-by-time fixed effects.

Payments may alternatively be negotiated without reference to RVUs. In this case, we have

$$P_{g,j,t} = \rho_{g,j,t} \qquad \text{or} \qquad \ln(P_{g,j,t}) = \ln(\rho_{g,j,t}), \tag{7}$$

with no role for $\varphi_{g,t}$ or $RVU_{j,t}$.

When RVU allocations change, equations (6) and (7) contain predictions for how private reimbursements will adjust. Consider two time periods, across which Medicare shifts payments by $\Delta \ln(RVU_{j,t})$. Let $\varepsilon_{g,j,t} = \Delta \ln(\rho_{g,j,t})$ be any change in the alternative non-benchmarked payment (as in equation 7). We can now write both types of prices in terms of service fixed effects and changes as follows. For Medicare-linked services, we have:

$$\ln(P_{g,j,t}) = \phi_j \mathbb{1}_j + \phi_g \mathbb{1}_g + \phi_{g,j} \mathbb{1}_g \cdot \mathbb{1}_j + \Delta \ln(RVU_{j,t}) \cdot \mathbb{1}_{\{t=\text{post}\}}. \tag{8}$$

For services not linked to Medicare, we have:

$$\ln(P_{g,j,t}) = \phi_j \mathbb{1}_j + \phi_g \mathbb{1}_g + \phi_{g,j} \mathbb{1}_g \cdot \mathbb{1}_j + \varepsilon_{g,j,t} \cdot \mathbb{1}_{\{t=\text{post}\}}. \tag{9}$$

In these equations, $\mathbb{1}_{\{t=\text{post}\}}$ is an indicator for the second time period. Under both pricing schemes, the fixed effects capture baseline payments to group $g$ for service $j$ in the first period, while the interaction with $\mathbb{1}_{\{t=\text{post}\}}$ captures the change between the two periods.

The linearity of equations (8) and (9) implies a straightforward way to estimate the fraction of services with payments benchmared to Medicare's relative values. Equation (8) says that a linear regression of log private payments on changes in log Medicare RVUs, for services with prices linked to Medicare, should yield a coefficient of 1 after controlling for the relevant sets of fixed effects. Equation (9) shows that the same regression should yield a coefficient of 0 for services not priced based on Medicare, as long as the non-Medicare payment changes ($\varepsilon_{g,j,t}$) are uncorrelated with RVU updates.

More generally, suppose that both types of payments exist, and specifically that a constant share $\sigma$ of payments are benchmarked to Medicare prices, while $1 - \sigma$ are set independently. (We will subsequently allow for heterogeneity in $\sigma$.) The average of log reimbursements is then given by a weighted average of equations (8) and (9), and the coefficient on log RVU updates can reveal the linked share $\sigma$:

$$\ln(P_{g,j,t}) = \phi_j \mathbb{1}_j + \phi_g \mathbb{1}_g + \phi_{g,j} \mathbb{1}_g \cdot \mathbb{1}_j + \sigma \cdot \Delta \ln(RVU_{j,t}) \cdot \mathbb{1}_{\{t=\text{post}\}} + \eta_{g,j,t}, \qquad (10)$$

where we define $\eta_{g,j,t} = (1 - \sigma) \cdot \varepsilon_{g,j,t} \cdot \mathbb{1}_{\{t=\text{post}\}}$. Equation (10) suggests that, in a linear regression with appropriate fixed effects, we can infer the Medicare-linked share from the coefficient on log RVU changes. This motivates our baseline specification for estimating $\sigma$. We use data at the level of individual claims, indexed by $c$, to estimate:

$$\ln(P_{c,g,j,t}) = \beta \Delta \ln(RVU_j) \cdot \mathbb{1}_{\{t=\text{post}\}} + \phi_t \mathbb{1}_{\{t=\text{post}\}} + \phi_j \mathbb{1}_j + \phi_g \mathbb{1}_g + \phi_{g,j} \mathbb{1}_g \cdot \mathbb{1}_j + \eta_{c,g,j,t}. \quad (11)$$

Equation (11) is a claims-level version of equation (10) where $\hat{\beta}$ estimates the share of payments based on Medicare rates. It adds a time period fixed effect $\mathbb{1}_{\{t=\text{post}\}}$ in case private payments shift broadly across the two time periods. This parametric difference-in-differences specification also incorporates full sets of group ($\mathbb{1}_g$), service ($\mathbb{1}_j$), and group-by-service ($\mathbb{1}_g \cdot \mathbb{1}_j$) effects to account for all time-invariant group- and service-specific terms. Thus our estimate of $\hat{\beta}$ is identified only using changes in RVUs across the two time periods. The time effect further limits the identifying variation exclusively to relative changes in RVUs across services. To obtain the share of spending linked to Medicare, we can estimate equation (11) weighted by the average pre-update price of each service.[18] We compute standard errors allowing for clustering in the errors $\eta_{c,g,j,t}$ at the service-code level, which is the level at which variation in RVU updates occurs.

To describe the timing with which BCBS incorporates Medicare updates into its reimbursements, we also present dynamic estimates from the following parametric event study:

$$\ln(P_{c,g,j,t}) = \sum_{t \neq 0} \beta_t \Delta \ln(RVU_j) \cdot \mathbb{1}_t + \phi_t \mathbb{1}_t + \phi_j \mathbb{1}_j + \phi_g \mathbb{1}_g + \phi_{g,j} \mathbb{1}_g \cdot \mathbb{1}_j + \eta_{c,g,j,t}. \tag{12}$$

When estimating equation (12), we again cluster standard errors at the service code level. We normalize $t$ such that $t = 1$ is the month in which BCBS has announced that it will implement Medicare's fee updates. We thus expect to see $\hat{\beta}_t = 0$ for periods preceding the updates' incorporation, $t < 0$, while the $\hat{\beta}_t$ for $t > 0$ are our estimates of how often Medicare updates are incorporated into private payments. A flat profile of the post-update $\hat{\beta}_t$ estimates would suggest that all price changes correlated with Medicare changes are implemented instantaneously. An upward trend in these coefficients might suggest that our baseline estimates are affected by ongoing renegotiations between BCBS and firms whose bargaining positions are affected by Medicare updates. We discuss this concern in detail in Appendix C.

### 3.3 Relating Our Approaches

The analyses we implement have complementary strengths and weaknesses. A shortcoming of the cross-sectional analysis of bunching is that it requires us to observe a constant markup across many services. Thus it may fail to detect genuine Medicare linkages involving markups that are common across relatively small numbers of services. These linkages would

[18]Since the unweighted regression treats each claim equally, it effectively weights service codes by the frequency with which they are used.

be detected, however, by our analysis of Medicare payment updates. The latter approach does not depend on the commonality of the markup. If the service is benchmarked, the payments will change when the underlying relative values change.

A shortcoming of the payment update approach, on the other hand, is that it could be biased if Medicare updates occur contemporaneously with changes driven by new contract negotiations. Our robustness analysis and our investigation of the precise timing of Medicare-linked changes provide evidence that new contract negotiations are unlikely to underlie our results. Nonetheless, these analyses cannot rule out active contract renegotiations altogether.

The bunching and changes approaches are thus complementary in that they have distinctive strengths and weaknesses. We relate these two analyses, and demonstrate consistency between them, by showing that the service-firm pairs we identify as benchmarked are strongly correlated across our approaches. We do this by dividing the data into subsamples according to the benchmarking results we obtain in our bunching analysis. We then estimate equation (11) separately on these subsamples.

# 4 Baseline Benchmarking Results

## 4.1 Bunching Estimates

Table 2 presents estimates of the share of services linked to Medicare according to the bunching method from section 3.1. The results show that at least two-thirds of spending, and three-quarters of services, have prices linked to Medicare. The full table explores sensitivity to two key assumptions. First, we round the value of each $ICF_{c,g,j,t}$ to the nearest 20 cents, 10 cents, or 2 cents to explore sensitivity to rounding error. Second, we define "common ICFs" as those that rationalize a sufficiently large share of the insurer's payments to a single physician group. In Figure 1, for example, the red line in Panel A should undoubtedly qualify as common. Other values may also qualify depending on the strictness of the threshold we apply. We consider thresholds ranging from 5 to 20 percent of a group's claims, then calculate the share of the insurer's payments associated with *any* of a group's common ICFs.

The Medicare-benchmarked shares range from 65 to 90 percent depending on the rounding and frequency thresholds; they decrease substantially with the stringency of the definition for a common ICF, but are not sensitive to the choice of rounding threshold. Appendix Table B.2 shows that alternative measures generate qualitatively similar results.[19]

Going forward, we require as our baseline that common ICFs account for 10 percent of a group's claims, when rounded to the nearest $0.02. The motivation for adopting a stringent rounding threshold is to be conservative in the extent to which our method detects false positives. At the same time, the 10 percent threshold ensures that multiple ICFs can be readily detected. Using this definition, over half of firms have just one common ICF. Fewer than 5 percent have more than 2 common ICFs.

Having identified these ICFs, we use them to describe how the generosity of BCBS reimbursements relates to firm and market characteristics. Table 3 presents estimates of equation (4), which regresses the ICF values themselves (in logs) against physician group and market characteristics.[20] Columns 1 through 3 reveal that each of firm size, market share, and market concentration is, by itself, positively correlated with the generosity of the firms' payments. Consistent with other work on health care pricing (Dunn and Shapiro, 2014; Baker et al., 2014; Kleiner et al., 2015; Cooper, Craig, Gaynor and Van Reenen, 2015; Clemens and Gottlieb, 2017), payments to large firms in markets with high levels of concentration are more generous than payments to small firms in markets with low levels of concentration.

[19] If we only count the single most common ICF for each group, the estimates are very similar to those reported in Table 2 when imposing a 20 percent threshold. Unfortunately, theory does not provide guidance as to which threshold is most appropriate, and the choice of threshold substantially affects our estimate of the linked share. Our changes-based estimation strategy is not sensitive to choices of this sort.

[20] Appendix Table B.3 also shows how these same characteristics relate to the frequency of deviations from Medicare benchmarking, and the value of the deviations when they occur.

Columns 4 and 5 include all three characteristics together, with column 5 also adding fixed effects for service codes and geographic areas. Firm size remains a strong predictor of the average generosity of a firm's payments, as does overall market concentration. Market share switches signs, likely because of collinearity with log firm size.

## 4.2 Results from Medicare Fee Change Analysis

We next move on from estimating ICFs to exploiting Medicare's RVU changes. Using the method from section 3.2, Panel B of Figure 2 presents event study estimates of the link between Medicare's relative value scale and BCBS reimbursements. It shows estimates of equation (12) for the Medicare payment changes implemented on July 1, 2010. The regression underlying the figure weights each claim by the dollar value of the service.[21]

The estimates reveal substantial—but not universal—links between Medicare updates and the payments providers receive from BCBS. The coefficients imply that $\hat{\sigma} = 75$ percent of services have payments linked to Medicare's relative values. The dramatic dynamics in the figure suggest that this reflects a contractual link between Medicare's relative values and BCBS payments. As in the raw data for standard office visits presented in Panel A, we see that payment changes occur when we expect. Importantly, the estimates of $\sigma$ are both economically and statistically larger than 0 and smaller than 1, implying that payments for a substantial share of services deviate from strict benchmarking to Medicare's relative values. Sections 5.1 and 5.2 will investigate these deviations in detail.

Column 1 of Table 4 presents our baseline estimates of equation (11), which summarizes this result in a single coefficient. The estimate in column 1 of Panel A shows that roughly 55 percent of BCBS's spending is linked to Medicare's relative values. This estimate corresponds with the analysis reported if panel B of figure 2. In Panel B, we weight service codes equally rather than according to baseline payments. The unweighted estimate implies that roughly

[21] Appendix Figure C.1 shows an unweighted version of this graph, for each year.

three quarters of BCBS's physician claims are paid based on Medicare's relative value scale. The difference in coefficients between Panels A and B implies that payments for relatively expensive services are less likely to be benchmarked to Medicare than are payments for low-cost services.[22]

## 4.3 Robustness and Cross-Validation of the Two Approaches

Table 4 probes the robustness of our changes-based estimates to a variety of specification checks. Column 1 of each panel reports our baseline specification, which includes a full set of group-by-HCPCS code fixed effects. In this baseline, we control for time effects with a simple post-update indicator. Column 2 drops the group-by-HCPCS code fixed effects in favor of a more parsimonious set of HCPCS code fixed effects. Column 3 augments the baseline specification by controlling for a cubic trend in the day of the year, which we interact with the size of each service's Medicare fee change. Column 4 allows the cubic trend in day to differ between the periods preceding and following the fee schedule update, as in a standard regression discontinuity design. The table shows that these specification changes have essentially no effect on the estimated coefficient $\hat{\beta}$. This reinforces the interpretation that, among services billed using standard HCPCS codes, roughly 55 percent of BCBS's spending is linked to Medicare's relative value scale.

Figure 3 shows dynamic estimates that pool together data from 2008–11, and use RVU changes from 2008–10 simultaneously. This estimation allows us to check whether any given change is offset by other changes in subsequent years. The figure shows no evidence of such an offset. The short-run responses to RVU changes persist for many quarters thereafter. The graph also shows flat pre-trends over long time periods, such as 6 quarters before the 2009 RVU changes and 10 quarters before the 2010 changes.

The estimates presented thus far may differ from the true Medicare benchmarking param-

[22]Appendix Tables C.1 and C.2 replicate Panels A and B, respectively, in other years' data.

eter $\sigma$ if changes in other terms of providers' contracts covary with the Medicare changes. Indeed, payment changes that significantly alter physician groups' average Medicare payment can move private payments in subsequent years, due in part to the resulting changes to their bargaining positions (Clemens and Gottlieb, 2017). In Appendix C.4, we thus draw on institutional detail and theoretically motivated specification checks to explore how much our estimates might deviate from the true share of payments benchmarked to Medicare's relative values. We find no evidence that renegotiations confound the relationship between BCBS's and Medicare's payments over the time horizons we analyze. Appendix C.4 thus bolsters the case for interpreting our estimates of $\hat{\beta}$ as measuring the fraction of services tied directly to Medicare.

To validate that our classification algorithm correctly captures services whose prices are actually linked to Medicare rates, we estimate our baseline changes-based regression separately for services identified as being more and less likely to be benchmarked according our bunching methodology. We classify each group-service pair $(g, j)$ as Medicare-linked if all of group $g$'s claims for service $j$ in the pre-update period appear to be linked to Medicare rates, and as non-linked otherwise. We estimate equation (11) separately for these two samples.

Panel D of Figure 2 shows two binned scatterplots analogous to Panel C, relating log BCBS price changes to log Medicare fee changes separately for the two samples. The linked sample is shown with red triangles and has a slope of 0.9, indicating that BCBS prices for 90 percent of linked services update in response to Medicare changes. The non-linked sample is shown with blue circles, and has a much smaller slope of 0.3.

In Appendix D.1 we consider the extenal validity of our baseline results using data from Colorado. Using data from one insurer in the Colorado All-Payer Claims Database, we obtain results generally in line with those from BCBS. This highlights that the phenomenon we investigate is not unique to our setting. Deviations from Medicare's payment structure are somewhat more common in the Colorado insurer, but the basic fact—substantial benchmarking, but far from universal—appears broadly relevant.

# 5 How Do Private Payments Deviate from Medicare?

In order to illuminate the economic determinants of benchmarking, we next consider variation in the strength of the link between private payments and Medicare's relative values. We consider the two primary dimensions along which payments vary: differences across physician groups and categories of services.

## 5.1 Deviations from Benchmarking across Physician Groups

One key difference across groups is the scale of their business with BCBS. Size, which is likely related to market power, could influence physician-insurer negotiations in multiple ways. Table 3 found that larger firms obtain higher ICFs. Larger group size could also lead to more deviations away from this benchmarking, in particular positive deviations (higher reimbursements).

To determine how size relates to benchmarking, we measure the quantity of care each group provides in our data. We then add interactions with practice size to our baseline changes regression, equation (11). Table 5 shows the results. The first column reports the baseline, equally weighted regression from Table 4. The second column introduces interactions between the Medicare updates and indicators for the size of the physician group providing the care. We define mid-sized firms as those with \$200,000 to \$1,000,000 in annual billing with BCBS, and large firms as those with more than \$1,000,000 in annual billing. Each of these categories comprises one-quarter of the sample, with the remaining half of claims coming from smaller firms. The estimates imply that nearly 90 percent of services provided by firms billing less than \$200,000 are benchmarked to Medicare, while roughly 60 percent of services provided by firms billing more than \$1,000,000 are benchmarked. Columns 3 and

4 present similar, but dollar-weighted, estimates. The results in column 4 suggest that 77 percent of payments to firms billing less than $200,000 are benchmarked to Medicare, while one-third of payments to firms billing more than $1,000,000 are benchmarked.[23]

Figure 4 shows that we find a similar relationship between the share linked to Medicare and physician group size using our cross-sectional bunching approach. The series in the figure reveal that this is true in both the equally-weighted and payment-weighted series. It is also true whether or not we adjust for the underlying composition of each group's services, to which we now turn.[24]

## 5.2 Which Services Deviate from the Medicare Benchmark?

The value of improving on Medicare's menu depends on the severity of that menu's inefficiencies. Because it is difficult to systematically quantify Medicare's inefficiencies across a large range of individual services, we focus on one of the Medicare fee schedule's most salient problems. Medicare rates are computed based on average-cost reimbursement, so its reimbursements will hew closer to marginal costs for labor-intensive services than for capital-intensive services. Standard optimal payment models suggest that the latter would

[23]Appendix Table C.5 shows similar results in data from other years.

[24]To check whether the relationship between benchmarking and group size is affected by the composition of large and small groups' services, we run a regression that allows group size and service composition to enter simultaneously. We define fixed effects $\mathbb{1}_{b(j)}$ using the "Betos" classification defined by Berenson and Holahan (1990). This hierarchical classification system goes from the broad categories we use here (such as Evaluation & Management and Imaging) to 2-digit (e.g. Advanced Imaging [MRIs and CAT scans]) and 3-digit classifications (e.g. CAT Scan: Head). We categorize all of the medical services in our data at the level of the 1-digit Betos categories.

To measure the relationship between group size and the Medicare-linked share, we categorize physician groups $g$ according to vigintiles of their aggregate private billing in a year, using $\mathbb{1}_{s(g)}$ to denote vigintile fixed effects. We then estimate the following regression at the group-code level:

$$\text{Medicare-Linked Share}_{j,g} = \nu_b \mathbb{1}_{b(j)} + \zeta_s \mathbb{1}_{s(g)} + \upsilon_{j,g}. \tag{13}$$

The orange diamonds in Figure 4 show the estimates of $\hat{\zeta}_s$. These illustrate the relationship between Medicare links and group size, adjusted for service composition. The composition-adjusted relationship remains strongly negative. The remaining measures in Figure 4 show similar results in terms of dollars spent, rather than number of services.

be more appropriately reimbursed through combinations of up-front financing of fixed costs and incremental reimbursements closer to marginal cost (Ellis and McGuire, 1986). We can proxy for services' capital and labor intensity by comparing the frequency of benchmarking across categories of care produced with different inputs, such as labor-intensive Evaluation & Management services versus capital-intensive Imaging.[25]

Table 6 estimates equation (11)—the relationship between private prices and changes in Medicare's relative values—separately across broad categories of services. The estimates imply that nearly 30 percent more of the payments for Evaluation & Management services are linked directly to Medicare's relative values than for Imaging services.[26]

Second, we divide Imaging codes into subcomponents with high capital and high labor content. Providers often bill separately for taking an image (the capital-intensive part, since it requires an imaging machine) and interpreting it (the labor-intensive part). When the same group supplies both components, it submits the bill as a "Global" service. The results in columns 5 through 7 show that payments for the labor-intensive Professional Component are more tightly linked to Medicare's relative values than are the payments for the capital-intensive Technical Component. These patterns support the hypothesis that physicians and insurers are more likely to contract away from Medicare's menu for capital-intensive services than for labor-intensive ones.

Table 7 shows that we find a similar relationship between the share linked to Medicare and service categories using our cross-sectional approach. Benchmarking is 30–50 percent less frequent for Imaging, Procedures, and Tests than for Evaluation & Management services. The results across columns reveal that we find similarly substantial differentials whether or not we control for firm size and whether services are weighted according to the spending they represent.

[25] These categories are defined using the Betos categories described in footnote 24.
[26] Appendix Table C.4 replicates this analysis in other years' data.

These results suggest that private contracts deviate when Medicare's rates are most problematic from an efficiency perspective. One way to interpret this is in light of negotiation and adjustment costs. Private bargaining can overcome these frictions more easily when Medicare's rates are farther from the efficient or equilibrium level that would obtain under unconstrained negotiations.

## 5.3 How Do Deviations Change Incentives Relative to Medicare?

What are physicians and insurers aiming to achieve when they negotiate reimbursements that deviate from Medicare's relative prices? In this section, we present evidence on the direction of deviations from strictly Medicare-benchmarked rates to investigate what services BCBS rewards through upward adjustments and discourages through downward adjustments. We do so by describing residuals from the following regression:

$$\ln(P_{g,j}) = \psi \ln(RVU_j) + \mu_g + e_{g,j}. \tag{14}$$

In a world of perfect benchmarking, we would find $\hat{\psi} = 1$ and $e_{g,j}$ uniformly equal to 0. So the empirical prediction errors $\hat{e}_{g,j}$ contain information about the direction of deviations from strict Medicare benchmarking. We examine heterogeneity in this prediction error across categories of services.

A subtle but important point is that this approach captures deviations from Medicare's relative prices that come through the introduction of multiple Medicare-benchmarked conversion factors. If an insurer thinks the Medicare menu's primary inefficiency is that it uniformly overpays for diagnostic imaging services relative to other services, for example, its preferred contract may simply set a low conversion factor for imaging services and a high conversion factor for other services. Our previous analyses would describe such a contract as being fully linked to Medicare. The analysis in the current section will capture the fact that this is structured to discourage the use of imaging services relative to other services.

Table 8 presents means of the residuals $\hat{e}_{g,j}$ from equation (14) across Betos categories.[27] The table shows that payments for Evaluation & Management and Testing services generally have positive residuals while payments for Imaging services have negative residuals. Figure 5 Panel A plots the distributions of these residuals by service category. The distribution for Imaging shows far more density of negative residuals than those for other services. Testing has more positive residuals, although that is largely driven by one outlier code.[28] Compared to the relative payments implied by Medicare's relative values, BCBS systematically adjusts its contracts to discourage imaging services. This coincides with the conventional wisdom that Medicare's relative values underpay for labor-intensive services relative to other services, and suggests that BCBS aims to partly rectify that mispricing.

Differences in BCBS's adjustments for labor- and capital-intensive services are particularly sharp across the subcategories of diagnostic imaging. Payment adjustments for the labor-intensive Professional Component of these services are positive, at around 0.015 in logs (approximately 1.5 percent). Payment adjustments for the capital-intensive Technical Component of these services are substantially negative, averaging $-0.12$ in logs. Figure 5 Panel B shows that this pattern holds throughout the distribution. While it is clear that BCBS reimbursements lean heavily on Medicare's relative values for their basic payment structure, these results suggest that BCBS adjusts its contracts to increase the generosity of payments for labor-intensive services and decrease its payments for capital-intensive services. This is consistent with deviating from Medicare with an eye towards more closely targeting either marginal costs or medical value.[29]

[27]Betos categories are aggregates of related services, defined in footnote 24. To be precise, these means are $\overline{\hat{e}_{g,j}} = \frac{1}{N_b}\sum_{j\in b}\hat{e}_{g,j}$, where each Betos group $b$ comprises $N_b$ claims for all services $j \in b$ in that group.

[28]In the Testing category the vast majority of residuals are negative, with the exception of one of the more common tests, which has a large and positive average residual. Recall from section 2, however, that Testing is the one category with significant missing data problems.

[29]Appendix D.2 shows that the changes in BCBS prices due to Medicare benchmarking matter in practice for the care that physicians supply to BCBS patients.

# 6 Conclusion

This paper uses physician payments from a large private insurer as a window into the structure of private insurers' contracts with physicians. Using two empirical strategies, we show that payments exhibit pervasive linkages to the relative payment model of the federal Medicare program. We find that three quarters of the services and 55 percent of the spending we analyze are benchmarked to Medicare. The analysis thus reveals that benchmarking to Medicare's relative rate structure is complemented by substantial customization.

In the contracts we analyze, customization is most prevalent when its value appears likely to be highest. Deviations from benchmarking occur disproportionately in contracts with large physician groups, where significant value may be at stake. Deviations also involve reductions in payments for diagnostic imaging services, a category of care for which many academics and policy makers believe marginal benefits are low relative to costs (Winter and Ray, 2008; MedPAC, 2011). The benchmarking phenomenon is strongest in payments for services where average-cost reimbursements will be most aligned with marginal costs, such as labor-intensive primary care services. When contracts deviate from Medicare, the direction of payment adjustments would tend to encourage the provision of primary care and discourage care for which over-utilization is a more widespread concern. The results are thus suggestive of effort to improve the payment structure through customization.

A number of factors, including contracting frictions, market power, the information content of Medicare's fee schedule, and the burden of regulatory compliance, may contribute to the contracting patterns we observe. Disentangling these forces is an attractive goal for future research. Further insights may be gained by analyzing the relationship between hospital contracting arrangements and Medicare's inpatient payment model. Variations in the degree of contract complexity, regulatory burdens, and both the size and market power of the parties involved may generate economically interesting variations in the connection between Medicare's inpatient payment model and private insurers' hospital payments. Early work along these lines includes an analysis by Baker, Bundorf, Devlin and Kessler (2016), who study the degree to which hospitals are paid retrospectively (rather than prospectively) by private insurers. Their results echo some of our findings here, in that Medicare-style prospective payments are more pervasive in similar economic conditions to those where we find more benchmarking.[30]

Regardless of the explanation, the Medicare-benchmarking phenomenon implies that many inefficiencies in Medicare's reimbursements spill over into private fee schedules. By extension, the value of improvements to public payment systems may ripple through private contracts in addition to improving the performance of Medicare itself. At the same time, the customization we observe would tend to curb what policy analysts regard as Medicare's greatest inefficiencies. Both public and private players thus appear to have important roles in the process of fee schedule improvement and payment system reform.

[30]Baker et al. (2016) analyze the extent to which private hospital payments reflect the Medicare program's primarily "prospective" approach. They find that private payments are more likely to be prospective when there is significant competition across the hospitals within a region, when a hospital's patients are covered primarily through managed-care arrangements, and when a large fraction of the hospital's patients are covered by Medicare.

# References

**Alpert, Abby, Mark Duggan, and Judith K. Hellerstein**, "Perverse reverse price competition: Average wholesale prices and Medicaid pharmaceutical spending," *Journal of Public Economics*, 2013, *103*, 44–62.

**Anderson, Eric, Nir Jaimovich, and Duncan Simester**, "Price Stickiness: Empirical Evidence of the Menu Cost Channel," *Review of Economics and Statistics*, October 2015, *97* (4), 813–826.

**Baker, Laurence C., M. Kate Bundorf, Anne B. Royalty, and Zachary Levin**, "Physician Practice Competition and Prices Paid by Private Insurers for Office Visits," *Journal of the American Medical Association*, 2014, *312* (16), 1653–1662.

**Baker, Laurence, M Kate Bundorf, Aileen Devlin, and Daniel P Kessler**, "Why Dont Commercial Health Plans Use Prospective Payment?," Working Paper No. 22709, National Bureau of Economic Research 2016.

**Berenson, Robert and John Holahan**, "Using A New Type-of-Service Classification System to Examine the Growth in Medicare Physician Expenditures, 1985–1988," NTIS Publication PB91-188599, Urban Institute Health Policy Center December 1990.

**Besley, Timothy J.**, "Optimal reimbursement health insurance and the theory of Ramsey taxation," *Journal of Health Economics*, 1988, *7* (4), 371–336.

**Bils, Mark and Peter J. Klenow**, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, October 2004, *112* (5), 947–985.

**Blue Cross Blue Shield of Texas**, "Fee schedule update," *Blue Review*, 2008, (1), 8.

_ , "Fee schedule updates effective in August," *Blue Review*, 2009, (2), 5.

_ , "Fee schedule updates," *Blue Review*, 2010, (3), 15.

_ , "Fee schedule update," *Blue Review*, 2011, (6), 10.

**Borges, Walt**, "Mo' Better Blues," *Texas Medicine*, April 2003, *99* (4).

**Brekke, Kurt R., Tor Helge Holmås, Karin Monstad, and Odd Rune Straume**, "Do Treatment Decisions Depend on Physicians' Financial Incentives?," Working Ppaer No. 07/2015, NIPE July 2015. Available online at http://www.nipe.eeg.uminho.pt/Uploads/WP_2015/NIPE_WP_07_2015.pdf (accessed December 9, 2015).

**Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller**, "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 2011, *29* (2), 238–249.

**Carroll, John**, "How Doctors Are Paid Now, And Why It Has to Change," *Managed Care*, December 2007, (12).

**Center for Studying Health System Change**, "Physician Survey Summary File: User's Guide and Codebook," 1999. Community Tracking Study Physician Survey, 1996–1997 (ICPSR 2597). Available at http://doi.org/10.3886/ICPSR02597.v3.

_ , "2004–05 Physician Survey Summary File: User's Guide and Codebook," 2006. Community Tracking Study Physician Survey, 2004–2005 (ICPSR 4584). Available at http://doi.org/10.3886/ICPSR04584.v2.

**Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans**, "Nominal rigidities and the dynamic effects of a shock to monetary policy," *Journal of Political Economy*, 2005, *113* (1), 1–45.

**Clemens, Jeffrey and Joshua D. Gottlieb**, "In the Shadow of a Giant: Medicare's Influence on Private Payment Systems," Working Paper No. 19503, National Bureau of Economic Research October 2013.

_ **and** _ , "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?," *American Economic Review*, April 2014, *104* (4), 1320–1349.

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

_ **,** _ **, and Adam Hale Shapiro**, "How Much Do Medicare Payment Cuts Reduce Inflation?," *FRBSF Economic Letter*, September 22 2014, *2014-28.*

_ **,** _ **, and** _ , "Medicare Payment Cuts Continue to Restrain Inflation," *FRBSF Economic Letter*, May 9 2016, *2016-15.*

**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," Working Paper No. 21815, National Bureau of Economic Research December 2015.

**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.

**Duggan, Mark and Fiona Scott Morton**, "The Distortionary Effects of Government Procurement: Evidence for Medicaid Prescription Drug Purchasing," *Quarterly Journal of Economics*, February 2006, *121*, 1–30.

**Dunn, Abe and Adam Hale Shapiro**, "Do Physicians Possess Market Power?," *Journal of Law and Economics*, February 2014, *57* (1), 159–193.

_ **and** _ , "Physician Payments Under Health Care Reform," *Journal of Health Economics*, 2015, *39*, 89–105.

**Einav, Liran, Amy Finkelstein, and Maria Polyakova**, "Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D," Working Paper No. 22277, National Bureau of Economic Research May 2016.

**Ellis, Randall P. and Thomas G. McGuire**, "Provider Behavior Under Prospective Reimbursement: Cost Sharing and Supply," *Journal of Health Economics*, June 1986, *5* (2), 129–152.

**Feldstein, Martin S.**, "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, March–April 1973, *81* (2), 251–280.

**Geruso, Michael, Timothy Layton, and Jacob Wallace**, "Are All Managed Care Plans Created Equal? Evidence from Random Assignment to Medicaid HMOs," 2016. UT Austin, mimeo.

**Gesme, Dean H. and Marian Wiseman**, "How to Negotiate With Health Care Plans," *Journal of Oncology Practice*, July 2010, *6* (4), 220–222.

**Government Accountability Office**, "Medicare Physician Payment Rates: Better Data and Greater Transparency Could Improve Accuracy," Report GAO-15-434, GAO May 2015.

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

**Jacobson, Mireille, Craig C. Earle, Mary Price, and Joseph P. Newhouse**, "How Medicare's Payment Cuts For Cancer Chemotherapy Drugs Changed Patterns Of Treatment," *Health Affairs*, July 2010, *29*, 1391–1399.

**Kaiser Family Foundation**, "State Health Facts," 2015. Available online at http://kff.org/state-category/health-insurance-managed-care/ (accessed April 17, 2017).

**Kleiner, Samuel A., William D. White, and Sean Lyons**, "Market power and provider consolidation in physician markets," *International Journal of Health Economics and Management*, March 2015, *15* (1), 99–126.

**Kleven, Henrik J.**, "Bunching," *Annual Review of Economics*, forthcoming 2016, *8.*

**Mahoney, Neale**, "Bankruptcy as Implicit Health Insurance," *American Economic Review*, February 2015, *105* (2), 710–746.

**McClellan, Mark**, "Reforming Payments to Healthcare Providers: The Key to Slowing Healthcare Cost Growth While Improving Quality?," *Journal of Economic Perspectives*, Spring 2011, *25* (2), 69–92.

**MedPAC**, "Report to the Congress: Medicare and the Health Care Delivery System," Technical Report, Medicare Payment Advisory Commission June 2011. Available at http://www.medpac.gov/documents/reports/Jun11_EntireReport.pdf (accessed October 7, 2015).

**Mertz, Gregory J.**, "Can You Negotiate Better Reimbursement?," *Family Practice Management*, October 2004, *11* (9), 31–34.

**Nakamura, Emi and Jón Steinsson**, "Five Facts About Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, November 2008, *123* (4), 1415–1464.

**Nandedkar, Maithily**, "Key principles for negotiating with insurers, deciding whether to opt in or out of Medicare," September 2011. Available at http://www.aad.org/dw/monthly/2011/september/key-principles-for-negotiating-with-insurers-deciding-whether-to-opt-in-or-out-of-medicare (accessed August 21, 2013).

**Olea, José Luis Montiel and Carolin Pflueger**, "A Robust Test for Weak Instruments," *Journal of Business and Economic Statistics*, July 2013, *31* (3), 358–369.

**Smets, Frank and Rafael Wouters**, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, September 2003, *1* (5), 1123–1175.

_ **and** _ , "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, June 2007, *97* (3), 586–606.

**Starc, Amanda and Robert J. Town**, "Internalizing Behavioral Externalities: Benefit Integration in Health Insurance," Working Paper No. 21783, National Bureau of Economic Research December 2015.

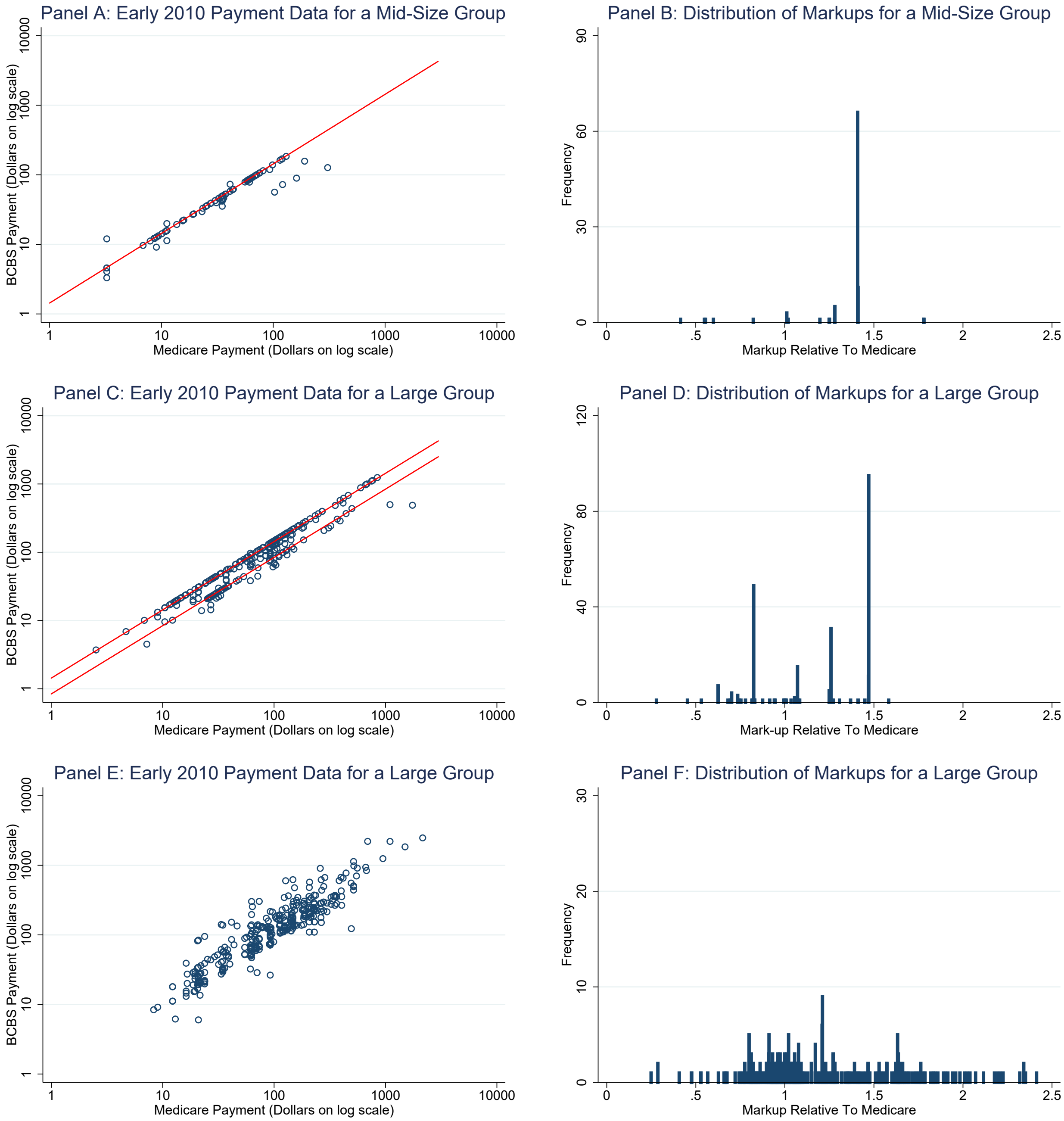
**United Food and Commercial Workers Union, The**, "UFCW 789 Grocery Contract," 2002. Available online at http://www.irle.berkeley.edu/library/pdf/0147.pdf.

**White, Chapin**, "Contrary To Cost-Shift Theory, Lower Medicare Hospital Payment Rates For Inpatient Care Lead to Lower Private Payment Rates," *Health Affairs*, May 2013, *32* (5), 935–943.

**Winter, Ariel and Nancy Ray**, "Paying Accurately For Imaging Services In Medicare," *Health Affairs*, 2008, *27* (6), 1479–1490.

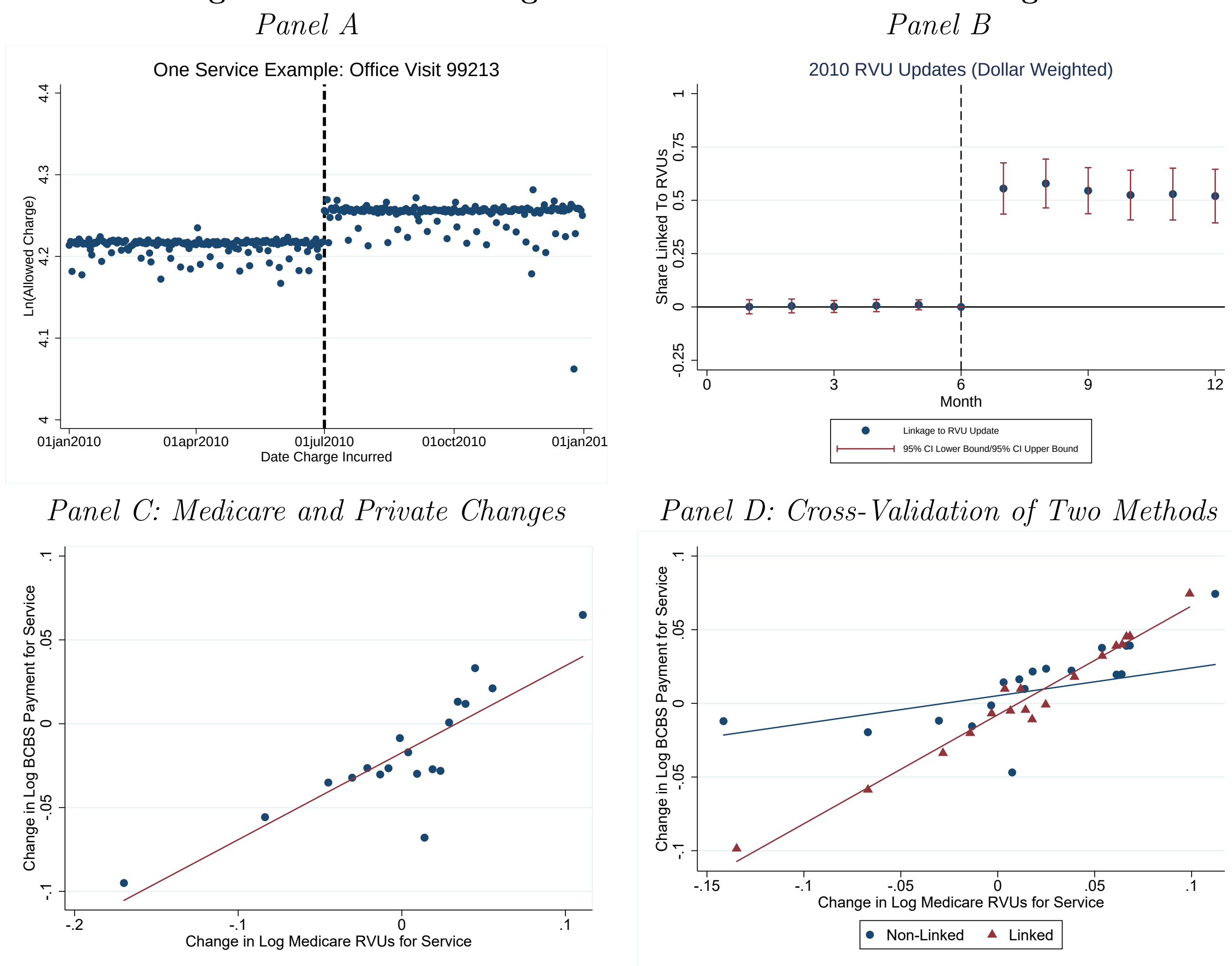
**Zuvekas, Samuel H. and Joel W. Cohen**, "Fee-For-Service, While Much Maligned, Remains The Dominant Payment Method For Physician Visits," *Health Affairs*, 2016, *35* (3), 411–414.

## Figure 1: Raw Payments For Illustrative Physician Groups



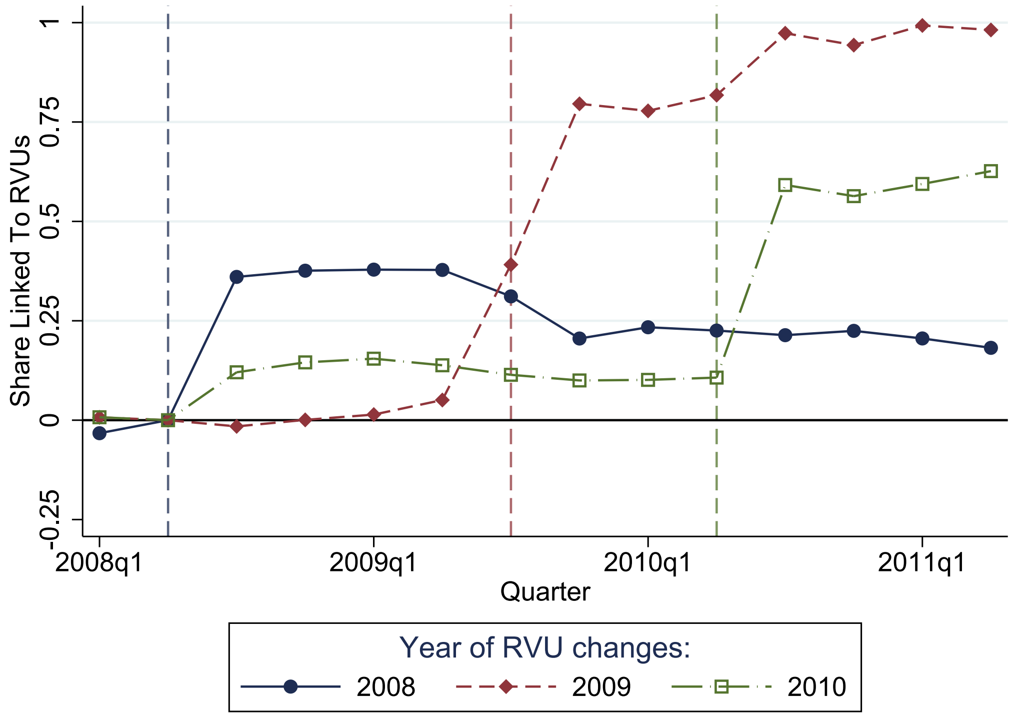

Note: Panels A through F present the raw data on BCBS reimbursement rates, and associated Medicare reimbursement, for 3 different physician groups in 2010. In Panels A, C, and E, each observation is a unique reimbursement paid for a particular service to the group. The lines have a slope of 1 (in logs) and represent the groups' most common Implied Conversion Factors. Panels B, D and F plot the distribution of markups relative to the Medicare rates for all payments each group received. They show clear spikes at the values that we identify as common Implied Conversion Factors in Panels A, C, and E. To comply with confidentiality rules, we omit from these graphs a small share of each group's claims. The share of claims whose observations are suppressed is 14.2% in Panels A and B, 1.94% in Panels C and D, and 2.95% in Panels E and F. Source: Authors' calculations using RVUs from the *Federal Register* and claims data from BCBS.

**Figure 2: Benchmarking Estimates Based on Price Changes**




Note: All panels use data from calendar year 2010. BCBS implemented its update from the 2009 to 2010 relative value scales on July 1, 2010, as indicated by the vertical dashed line in Panels A and B. Panel A presents daily averages of BCBS's log payment for a standard office visit. Panel B shows estimates of $\beta_t$ from equation (12), weighting observations equally. Standard errors for the estimates in Panel B are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Panel C presents a binned scatterplot of the relationship between Medicare payment updates (into 20 vigintiles) and changes in private payments. Private price changes are computed as the difference between service-level average payments after and before July 1, 2010. Panel D is similar, but with separate data and estimation for services that we identify as being linked to Medicare on the basis of their Implied Conversion Factors and those we identify as being non-linked. For presentation in the binned scatterplot, observations within each class of services (i.e., linked or non-linked) are grouped into twenty vigintiles on the basis of the log change in the service code's Medicare RVU allocation. The regression lines shown in Panels C and D are estimated at the underlying service-code level. Sources: Authors' calculations using RBRVS updates from the *Federal Register* and claims data from BCBS.

**Figure 3: Estimating Multiple Years' RVU Updates Simultaneously**

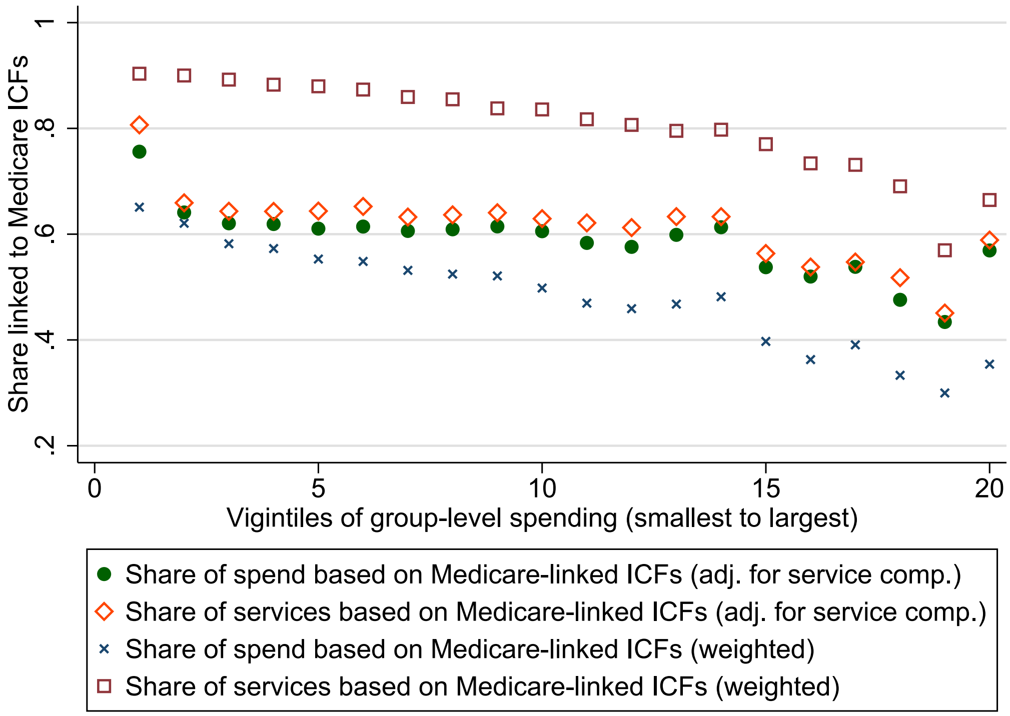



Note: The figure reports estimates of $\beta_t^{08}$, $\beta_t^{09}$ and $\beta_t^{10}$ from the following modification of equation (12):

$$\ln(P_{c,g,j,t}) = \sum_{t\neq 0} \beta_t^{08} \Delta \ln(RVU_{j,08}) \cdot \mathbb{1}_t + \sum_{t\neq 0} \beta_t^{09} \Delta \ln(RVU_{j,09}) \cdot \mathbb{1}_t^{10} + \sum_{t\neq 0} \beta_t \Delta \ln(RVU_{j,10}) \cdot \mathbb{1}_t + \phi_t \mathbb{1}_t + \phi_j \mathbb{1}_j + \phi_g \mathbb{1}_g + \phi_{g,j} \mathbb{1}_g \cdot \mathbb{1}_j + \eta_{c,g,j,t}. \tag{15}$$

In this specification, $\Delta \ln(RVU_{j,T})$ refers to the log of Medicare's RVU updates from calendar year $T-1$ to calendar year $T$. The corresponding coefficients $\beta_t^T$ indicate what share of the year-$T$ RVU updates were incorporated into BCBS payments during calendar quarter $t$. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). BCBS implemented its RVU updates on July 1, 2008, August 15, 2009, and July 1, 2010. The omitted interaction ($t = 0$) is 2008Q2 for all of the RVU update variables. The regression line is estimated at the underlying service-code level and is dollar-weighted. Sources: Authors' calculations using claims data from BCBS.

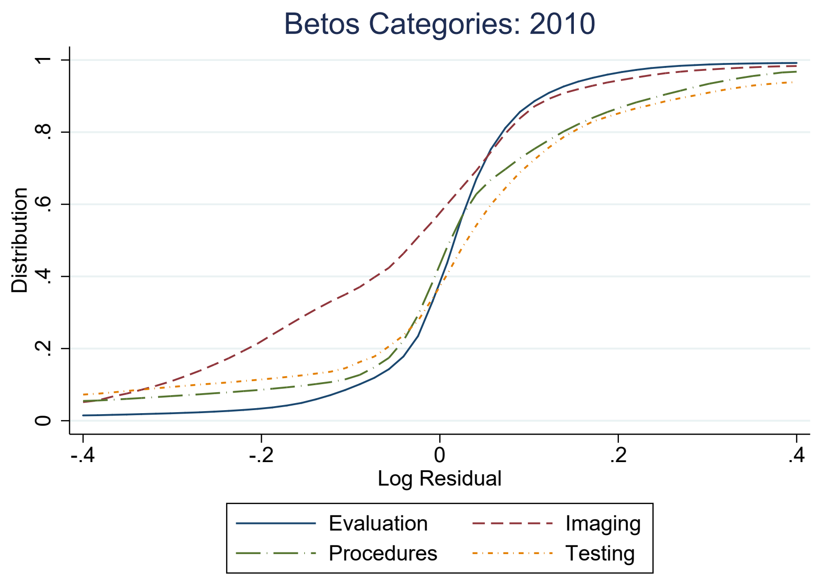
**Figure 4: Frequency of Benchmarking and Physician Group Size**



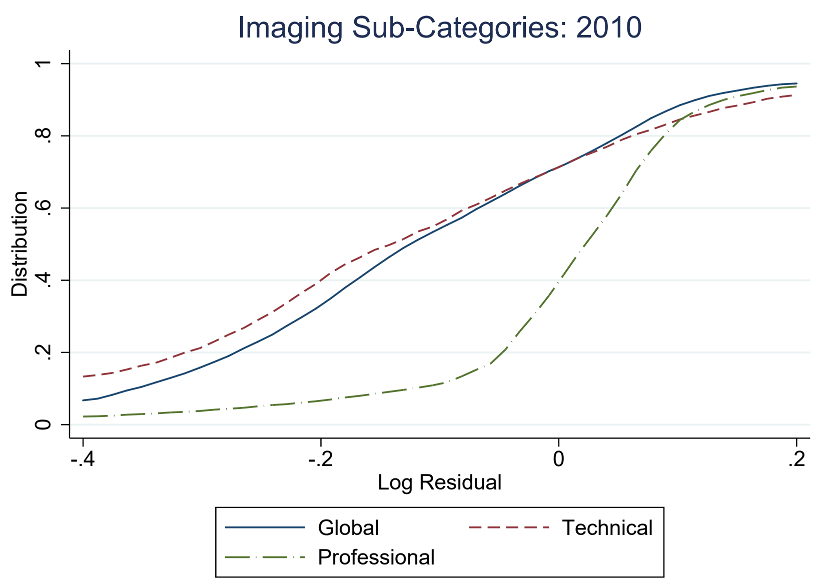

Note: This figure shows the relationship between a group's Medicare-linked service share and group size. Specifically, it plots variation in the share of services priced according to common Implied Conversion Factors (cICFs), as defined in section 3.1, according to physician group size. We measure group size by forming 20 vigintiles based on the group's BCBS billing. We require that cICFs account for 10 percent of a group's claims, when rounded to the nearest \$0.02. The green dots and orange diamonds show estimates of $\zeta_b$ from equation (13), which adjust for the composition of each group's services. The blue $\times$'s and red squares are unadjusted, but weighted to measure the Medicare-linked share of spending in dollar terms as opposed to the share of services. All data are from 2010. Sources: Authors' calculations using claims data from BCBS.

**Figure 5**: **Deviations from Medicare Benchmark by Service Category**

*Panel A*




*Panel B*




Note: The figure presents the distributions of empirical residuals $\hat{e}_{g,j}$ from estimates of equation (14). The distribution of residuals is shown within either broad Betos categories (Panel A), or within the subcategories of Imaging (Panel B). The distributions are smoothed using a local linear regression, with an Epanechnikov kernel and a bandwidth of 0.01. Source: Authors' calculations using claims data from BCBS.

**Table 1: Summary Statistics by Physician Group**

| | Mean | Median | Std. Dev. | Min. | Max. |
|---|---|---|---|---|---|
| *Panel A: All Groups* ($N$=80,675) | | | | | |
| | **Mean** | **Median** | **Std. Dev.** | **Min.** | **Max.** |
| Number of unique services | 9.70 | 3 | 27.23 | 1 | ~1,700 |
| Number of patients | 87.59 | 2 | 698.85 | 1 | ~61,930 |
| Number of doctors | 1.73 | 1 | 7.93 | 1 | ~1,100 |
| Number of claims | 201 | 3 | 1,763 | 1 | ~163,360 |
| Mean allowed amount | 108.91 | 84.43 | 125.16 | 0.64 | ~7,680 |
| Total BCBS revenue | 25,457 | 383 | 274,700.3 | 0.64 | ~43,000,000 |
| *Panel B: Groups with Billings* $>$ \$10,000 ($N$=15,235) | | | | | |
| | **Mean** | **Median** | **Std. Dev.** | **Min.** | **Max.** |
| Number of unique services | 35.99 | 24 | 53.12 | 1 | ~1,700 |
| Number of patients | 424.35 | 151 | 1,523 | 1 | ~61,930 |
| Number of doctors | 4.14 | 2 | 17.56 | 1 | ~1,100 |
| Number of claims | 981.13 | 386 | 3,860 | 1 | ~163,360 |
| Mean allowed amount | 105.52 | 84.65 | 136.3 | 10.75 | ~7,680 |
| Total BCBS revenue | 124,687 | 44,392 | 606,644 | 10,000 | ~43,000,000 |

Note: Table shows summary statistics for data by physician group. ~ indicates rounding. Source: Authors' calculations using claims data from BCBS.

**Table 2**: **Services Priced According to Common Implied Conversion Factors**

| | | | |
|---|---|---|---|
| *Panel A: Dollar-Weighted* | | | |
| | Frequency Threshold: | | |
| | *5%* | *10%* | *20%* |
| Rounding for ICFs: | | | |
| *\$0.02* | 83% | 76% | 66% |
| *\$0.10* | 86% | 80% | 71% |
| *\$0.20* | 87% | 80% | 71% |
| *Panel B: Service-Weighted* | | | |
| | Frequency Threshold: | | |
| | *5%* | *10%* | *20%* |
| Rounding for ICFs: | | | |
| *\$0.02* | 87% | 81% | 70% |
| *\$0.10* | 89% | 84% | 75% |
| *\$0.20* | 89% | 85% | 75% |

Note: Each cell shows the share of services for which payments are associated with a common Implied Conversion Factor (cICF), as defined in the main text. Data are from January 1—June 30, 2010, over which time BCBS used the 2009 version of Medicare's Resource Based Relative Value Scale. The cells within each panel show how the linked share varies as we apply different thresholds for the frequency required to quality as a cICF. The column labeled "Rounding" indicates the rounding applied to each estimated ICF. An ICF is defined as "common" for the payments to a physician group if it accounts for at least the fraction of services associated with the specified Frequency Threshold. Source: Authors' calculations using claims data from BCBS.

**Table 3**: **Firm Size and Implied Conversion Factors**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dependent variable: | | | Log implied conversion factor (ICF) | | |
| Firm Size (Log Spending) | 0.058** | | | 0.058** | 0.040** |
| | (0.004) | | | (0.005) | (0.006) |
| Firm Market Share | | 0.241** | | -0.158** | -0.092** |
| | | (0.015) | | (0.037) | (0.029) |
| Market Concentration | | | 0.238** | 0.318** | 0.159** |
| | | | (0.020) | (0.036) | (0.028) |
| $N$ | 20,736,449 | 20,736,449 | 20,736,449 | 20,736,449 | 20,736,449 |
| No. of Clusters | 23,098 | 23,098 | 23,098 | 23,098 | 23,098 |
| Code Effects | No | No | No | No | Yes |
| HSA Fixed Effects | No | No | No | No | Yes |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. Estimated at the claim level, this table shows the empirical relationship between the level of physicians' reimbursements, measured using Implied Conversion Factors (ICFs), and measures of firm size and/or market concentration. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each physician group. The construction of all variables is discussed in the main text. Source: Authors' calculations using claims data from BCBS.

**Table 4: Estimating Medicare Benchmarking Using RVU Changes**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | |
| | *Panel A: Weighted by Price* | | | |
| Log RVU Change × Post | 0.539** | 0.544** | 0.568** | 0.538** |
| | (0.061) | (0.061) | (0.060) | (0.061) |
| $N$ | 23,933,577 | 23,933,577 | 23,933,577 | 23,933,577 |
| No. of Clusters | 3,681 | 3,681 | 3,681 | 3,681 |
| | *Panel B: Unweighted* | | | |
| Log RVU Change × Post | 0.750** | 0.748** | 0.765** | 0.749** |
| | (0.038) | (0.038) | (0.043) | (0.038) |
| $N$ | 23,933,577 | 23,933,577 | 23,933,577 | 23,933,577 |
| No. of Clusters | 3,681 | 3,681 | 3,681 | 3,681 |
| Group-by-Code Effects | Yes | No | Yes | Yes |
| Code Effects | No | Yes | No | No |
| Cubic Time × RVU Change | No | No | Yes | No |
| Cubic Time × Post | No | No | No | Yes |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the form described in section 3.2. Each column in each panel reports an estimate of $\hat{\beta}$ from equation (11). Observations are at the claim level and are weighted according to each service's average payment during the baseline period (Panel A), or equally weighted (Panel B). Data are from 2010. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Table 5: Medicare Benchmarking by Firm Size**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | | *Log private reimbursement rate* | | |
| Log RVU Change | 0.750** | 0.882** | 0.539** | 0.775** |
| × Post-Update | (0.038) | (0.073) | (0.061) | (0.094) |
| Log RVU Change | | -0.074 | | -0.140* |
| × Post-Update × Midsize | | (0.098) | | (0.069) |
| Log RVU Change | | -0.293* | | -0.448** |
| × Post-Update × Large | | (0.117) | | (0.102) |
| *N* | 23,933,577 | 23,933,577 | 23,933,577 | 23,933,577 |
| Weighting: | Service | Service | Dollar | Dollar |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. Columns 1 and 3 report the baseline estimates from Table 4 Panels A and B respectively. In columns 2 and 4 we augment these specifications to include interactions between firm size indicator variables and both the "Post" indicator and the interaction between the "Log RVU Change" and "Post" indicator. The omitted category is small firms, defined as those with less than \$200,000 in billings. Mid-sized firms are those with billings between \$200,000 and \$1 million, and large firms are those with billings exceeding \$1 million. Data are from 2010 and at the claim level. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Table 6: Public-Private Payment Links Across Service Categories**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | | | | |
| | Evaluation | Imaging | Procedures | Tests | Imaging Sub-Categories: | | |
| | | | | | Global | Technical | Professional |
| Log RVU Change | 0.841** | 0.564** | 0.720** | 1.066** | 0.545** | 0.387* | 0.982** |
| × Post-Update | (0.036) | (0.084) | (0.081) | (0.066) | (0.109) | (0.152) | (0.066) |
| $N$ | 12,259,186 | 3,630,019 | 4,750,313 | 1,542,254 | 1,826,666 | 209,178 | 1,594,175 |
| No. of Clusters | 221 | 1,085 | 1,936 | 408 | 408 | 244 | 433 |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the form described in section 3.2. The cells in each panel report estimates of $\hat{\beta}$ from equation (11), with samples selected to contain the HCPCS codes falling into broad service categories. The name of the relevant service category accompanies each point estimate. Data are from 2010 and at the claim level. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Table 7: Medicare Benchmarking by Betos Category**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | Payments with Common Conversion Factors | | | |
| | Spending Share | | Service Share | |
| Imaging | -0.427** | -0.471** | -0.300** | -0.355** |
| | (0.053) | (0.047) | (0.030) | (0.024) |
| Procedures | -0.309** | -0.352** | -0.336** | -0.388** |
| | (0.030) | (0.028) | (0.054) | (0.052) |
| Tests | -0.383** | -0.415** | -0.258** | -0.297** |
| | (0.051) | (0.047) | (0.055) | (0.054) |
| Constant | 0.921** | 0.828** | 0.941** | 0.829** |
| | (0.015) | (0.015) | (0.020) | (0.017) |
| $N$ | 542,207 | 542,207 | 542,207 | 542,207 |
| Omitted Category | Evaluation & Management | | | |
| Additional Controls | Group Size | None | Group Size | None |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. This table shows estimates of the $\nu_b$ coefficients in equation (13), namely the relationship between Betos category and the Medicare-linked share of claim lines (columns 1 and 2) or spending (columns 3 and 4). Medicare links are measured using the common Implied Conversion Factors (cICFs) defined in section 3.1, using data from January 1 through June 30, 2010. We require that cICFs account for 10 percent of a group's claims, when rounded to the nearest $0.02. Columns 1 and 3 show estimates after controlling for vigintile of group size, as measured with BCBS spending, and columns 2 and 4 show estimates without group size controls. Standard errors are two-way clustered (Cameron, Gelbach and Miller, 2011) by Betos category and physician group. Sources: Authors' calculations using claims data from BCBS.

**Table 8: In What Direction Does BCBS Adjust Its Payments for the Various Service Categories?**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | *Distributions of Payment Residuals by Betos Categories* | | | | | | |
| | Evaluation & Management | Imaging | Procedures | Tests | Imaging Sub-Categories: | | |
| | | | | | Global | Technical | Professional |
| Residual Mean | 0.0112 | -0.0624 | 0.0107 | 0.0301 | -0.122 | -0.124 | 0.0150 |
| Residual SD | (0.169) | (0.246) | (0.279) | (0.319) | (0.272) | (0.281) | (0.177) |
| $N$ | 6,010,826 | 1,743,011 | 2,312,734 | 751,726 | 883,419 | 102,465 | 757,127 |

Note: The table presents means and standard deviations of residuals from estimates of equation (14) in data from 2010. That is, we regress the log of BCBS's payments on a set of physician-group fixed effects and the log of each HCPCS code's number of Relative Value Units. This table describes the residuals from that regression. We restrict the sample to the pre-update period (January 1 through June 30, 2010) so that the relative value units are constant for each service throughout the sample. Source: Authors' calculations using claims data from BCBS.

# Appendix For Online Publication Only

# A Conceptual Framework: Contracting Under Complexity

We present a conceptual framework to illuminate when payment rates will and won't be benchmarked perfectly to Medicare. Physicians and insurers can use Medicare's payments as a default relative price schedule, so that reimbursements are simply a markup over Medicare's rates. Adopting this default has costs if Medicare's relative payments are suboptimal, in a sense developed below. It may nonetheless be efficient to rely on this default due to negotiation and coordination costs.[31]

Consider an insurer that purchases $N$ types of medical services indexed by $j \in \{1, \ldots, N\}$, for treating its enrollees. These types could represent individual billing codes—very specific services such as a 20-minute office visit—or broader categories of care, such as all diagnostic imaging. We abstract from the physician-insurer bargaining process and assume that the insurer sets prices with full knowledge of the aggregate supply curve for each type of care. Let $r_j$ denote the reimbursement rate that the insurer pays to physicians for providing service $j$, and let $r_j^M$ be the corresponding Medicare rate. In aggregate, the physician market supplies care to the insurer's patients according to the supply functions $s_j(r_j)$. If the true price-setting process is not so simple—say, if physicians are not price-takers—the model's main ideas still hold. In that case, physicians and insurers strive to reach a pricing agreement that maximizes joint surplus, which they then split. We would simply view prices as jointly determined and adjustment costs defined below as those incurred by both parties.

We assume that the insurer aims to minimize its medical expenses while keeping patients, or their employers, satisfied with the insurance product. This constraint requires that the insurer purchase enough care to achieve the patients' reservation value $\overline{u}$.[32] We assume the patients have additively separable preferences over types of care, captured by $U(q_1, \ldots, q_N) = \sum_j u_j(q_j)$ where $q_j$ is the quantity of service $j$ supplied to a representative patient.

We will first consider the insurer's optimally chosen reimbursement rates and then consider deviations from that optimum. The insurer incurs costs of $C(r_1, \ldots, r_N) = \sum_j r_j q_j$ for treating patients. The physician supply curves determine quantities as $q_j = s_j(r_j)$. The insurer aims to minimize costs while keeping the patients satisfied, or:

$$\min_{r_1,\ldots,r_N} \sum_j s_j(r_j) r_j \text{ subject to } \sum_j u_j(s_j(r_j)) \geq \overline{u}. \tag{A.1}$$

[31] Providers themselves may find deviating from Medicare's menu costly due to increases in the non-trivial administrative expenses associated with billing (Cutler and Ly, 2011). A private alternative to Medicare's menu would not reduce these costs. It would also have trouble reflecting differences in market conditions which, as the model shows, influence optimal reimbursement rates.

[32] Geruso, Layton and Wallace (2016) take a similar conceptual approach to a health insurer's objective function.

Let $\theta_j = u_j'(q_j)$ denote the marginal value of care type $j$, and let $\epsilon_j = \dfrac{s_j'(r_j)r_j}{s_j(r_j)}$ denote the supply elasticity for that service. The first-order conditions for problem (A.1) then satisfy:

$$\frac{r_j^*}{r_k^*} = \frac{\theta_j / \left({\epsilon_j}^{-1} + 1\right)}{\theta_k / \left({\epsilon_k}^{-1} + 1\right)} \quad \text{for all } j, k. \tag{A.2}$$

Equation (A.2) shows that the insurer would optimally increase the relative payment for a service that is more highly valued (higher $\theta$) or whose supply is more elastic (higher $\epsilon$).

We next consider how much the insurer's costs increase if its reimbursement rates follow a different ratio. Suppose that the insurer is forced to adopt a scaled version of Medicare's pricing scheme, where reimbursements are set exogenously at $r_j^M$. The insurer scales all of the Medicare rates by $\varphi$ so that the patient satisfaction constraint $U(q_1, \ldots, q_N) = \overline{u}$ continues to bind with equality. Let $d_j = \varphi r_j^M - r_j^*$ represent the deviation from optimal reimbursements implied by this scheme, and $\delta_j = \dfrac{d_j}{r_j^*}$ the proportional deviation. Supposing that these deviations are reasonably small, we can approximate the insurer's costs to second order as

$$\begin{aligned} C(\varphi r_1^M, \ldots, \varphi r_N^M) &= \sum_j s_j(r_j^* + d_j)(r_j^* + d_j) \\ &\approx C(r_1^*, \ldots, r_N^*) + \sum_j \left[ s_j'(r_j^*) + \frac{1}{2} s_j''(r_j^*) \right] d_j^2 && \text{(A.3)} \\ &= C(r_1^*, \ldots, r_N^*) + \frac{1}{2} \sum_j s_j(r_j^*) \epsilon_j \left( \rho_j + \epsilon_j + 2 r_j^* - 1 \right) \delta_j^2 && \text{(A.4)} \end{aligned}$$

where $\rho_j = \frac{d\epsilon_j}{dr_j^*} \frac{r_j^*}{\epsilon_j}$ denotes the super-elasticity of supply. Note that, since $C(r_1^*, \ldots, r_N^*)$ defines the cost-minimizing reimbursements, the summation terms in equations (A.3) and (A.4) are positive. By the envelope theorem, small deviations from optimal pricing that continue to satisfy the patient satisfaction constraint don't generate any first-order cost increase. However larger deviations from the insurer's preferred reimbursements $r^*$ generate convex costs, which are increasing in the supply elasticity $\epsilon_j$ and super-elasticity $\rho_j$. This occurs because larger deviations from the insurer's unconditional optimum force the insurer to spend more to achieve the same patient satisfaction. This is especially so with more elastic supply, which causes any given deviation to shift physician behavior farther from the efficient service mix.

Suppose that Medicare's pricing is used as a benchmark and the insurer can adopt the scaled version ($r_j = \varphi r_j^M$) for free. Alternatively, the insurer can make costly adjustments to this default. Assume that it can pay a cost of $\alpha\theta$ to shrink the magnitude of the deviations relative to optimal prices from $\delta_j$ to $\dfrac{\delta_j}{\theta + 1}$. How much of a reduction will it pay for?[33]

[33] We can alternatively think of $\theta$ as determining a probability that the reimbursement for each service is

Including the adjustment costs, the insurer's total spending in excess of the first-best is now:

$$\Delta C(\theta) = \frac{1}{2} \sum_j s_j(r_j^*)\epsilon_j \left(\rho_j + \epsilon_j + 2r_j^* - 1\right) \left(\frac{\delta_j}{\theta + 1}\right)^2 + \alpha\theta. \tag{A.5}$$

The insurer will choose $\theta$ to minimize $\Delta C(\theta)$. This minimum is achieved at:

$$\theta^* = \max \left\{ 0, \left[ \frac{1}{\alpha} \sum_j s_j(r_j^*)\epsilon_j \left(\rho_j + \epsilon_j + 2r_j^* - 1\right) \delta_j^2 \right]^{1/3} - 1 \right\}. \tag{A.6}$$

Larger pricing errors $\delta_j$, higher quantities $s_j(r_j^*)$, more elastic supply, and lower adjustment costs $\alpha$ will all lead the insurer to spend more on mitigating the deviations.

switched from exactly the benchmark to exactly the insurer's preferred value. By paying for a higher $\theta$, the insurer can obtain the opportunity to adjust payment for additional randomly chosen services. Following the formulation in the text, a payment of $\alpha\theta$ allows the insurer to adjust a share $1 - 1/(\theta+1)^2$ of services, while the remaining share $1/(\theta+1)^2$ retain the Medicare default $\varphi r_j^M$.

# B Additional Detail on Implied Conversion Factors

## B.1 Data Cleaning

This section describes our process for cleaning and merging the BCBS claims data. Table B.1 shows the data lost as we progress from the raw claims data to the final analysis sample.

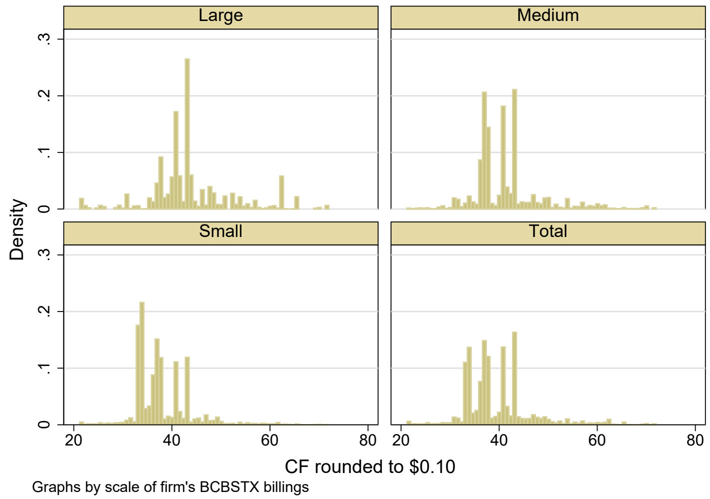
For concreteness, consider the 2009 claims data. The data for this year start with 54,724,994 claim lines and $4.01 billion in spending (row A). To reduce heterogeneity along several administrative margins, we analyze claim lines for which the payment is non-missing, the service quantity is 1, and the observation is an "original" claim line rather than an adjustment to a past payment.[34] This eliminates 5,090,024 claim lines and leaves us with $3.24 billion in spending (row B). Next, we want to ensure that our analysis focuses on reimbursements for services that are administratively equivalent from a payments perspective, and whose payments have been agreed upon through *ex ante* negotiations. We thus retain only observations that are explicitly coded as being "outpatient" and "in network." These criteria eliminate a total of 8,302,709 claim lines and leave us with $2.45 billion in spending (row C). Next we drop relatively rare service codes for which we have fewer than 10 observations prior to the RVU updates in a given year. In the 2009 data, this eliminates 149,269 claims and leaves us with $2.44 billion in spending (row D). The resulting sample of 41,182,992 service lines and $2.44 billion in spending constitutes the administratively comparable and sufficiently common billing codes we aim to understand.

In order for private insurers to benchmark prices to Medicare, at a minimum they would need to use Medicare's billing codes. On row (E), we thus merge the remaining claims with Medicare billing codes, which provides an upper bound on the potential benchmarking. The final analysis sample in 2009 includes 3,807 unique HCPCS codes, which comprise 21,941,227 service lines and $1.89 billion of spending. The key conclusion from row (E) is that, once we restrict ourselves to the relevant universe of data, additional losses from merging in Medicare codes and eliminating infrequent codes are not substantial. More specifically, this merge only loses notable portions of one broad spending category, namely laboratory tests, for which both Medicare and BCBS frequently base payments on non-standard codes. We retain over 97 percent of claims for evaluation and management, diagnostic imaging, and surgical services.

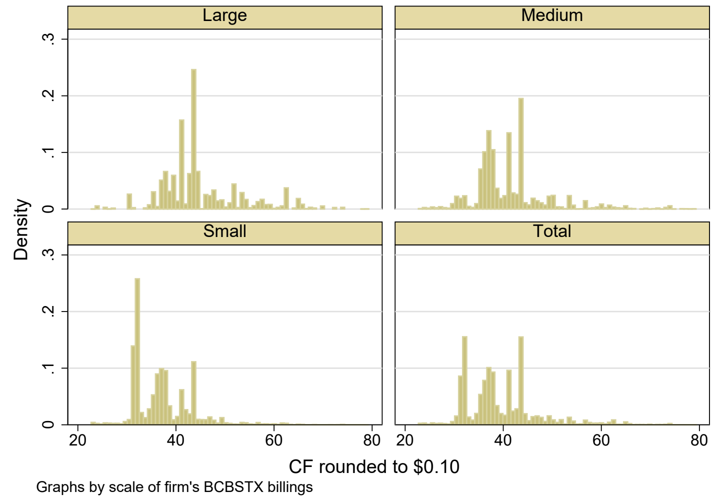
[34]Both Medicare and private sector payment policies generate nonlinear payments in certain circumstances when multiple instances of the same service are provided per claim.

## Appendix Figure B.1: Distribution of ICFs by Firm Size

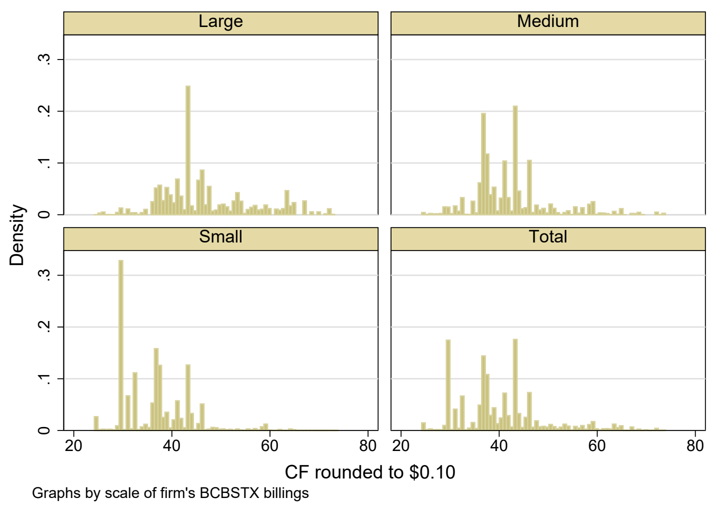
*Panel A: 2008*



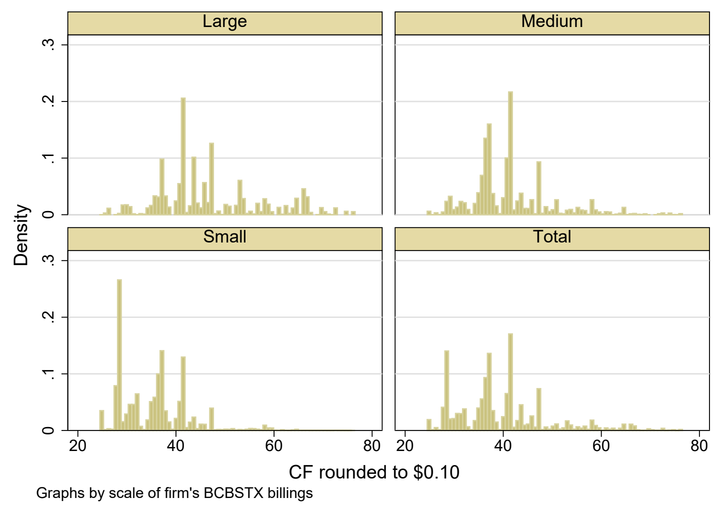

*Panel B: 2009*




*Panel C: 2010*




*Panel D: 2011*




Note: The figure reports the distributions of common Implied Conversion Factors that we compute in each year. We require that common ICFs account for 10 percent of a group's claims, when rounded to the nearest $0.02. Each year's distributions are split according to the sizes of the physician groups, measured as the dollar value of the group's BCBS billings.

**Appendix Table B.1: Data Cleaning**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Year: | 2008 | | 2009 | | 2010 | | 2011 | |
| Measure: | Claims | Spending | Claims | Spending | Claims | Spending | Claims | Spending |
| (A) Initial dataset | 45.5m | \$3.49b | 54.7m | \$4.09b | 57.6m | \$4.29b | 61.7m | \$4.64b |
| (B) Basic cleaning | 90.0% | 80.2% | 90.7% | 80.8% | 90.0% | 80.0% | 90.3% | 80.4% |
| (C) In-network outpatient | 74.0% | 59.6% | 75.5% | 61.1% | 76.5% | 61.5% | 77.3% | 62.3% |
| (D) Exclude rare codes | 73.9% | 59.3% | 75.3% | 60.8% | 76.5% | 61.3% | 77.3% | 62.1% |
| (E) Medicare code merge | 41.3% | 47.3% | 40.3% | 47.1% | 41.7% | 47.8% | 41.3% | 47.8% |

Note: This table quantifies the data lost at each step of our data cleaning and merge process. We show calculations for each of the four years of BCBS claims data. For each year, row (A) shows the raw number of claims (odd-numbered columns) and money spent (even-numbered columns) in that year's claims data. All subsequent rows show the share of claims on row (A) that remain after each set of cleaning steps. Row (B) shows the share of data remaining when we keep only claim lines for which the payment is non-missing, the service quantity is 1, and the observation is an "original" claim line rather than an adjustment to a past payment. These basic cleaning steps eliminate about ten percent of claims and twenty percent of spending. Row (C) further restricts our sample to the universe we consider, namely outpatient in-network claims. This eliminates approximately 15 percent more claims, and twenty percent more spending per year. Row (D) drops those relatively rare service codes for which we have fewer than 10 observations prior to the RVU updates in a given year; this has minimal effect on the sample sizes. Finally, row (E) drops claims that don't merge with Medicare's RBRVS codes. This loses 12–15 percent of observations per year. Source: Authors' calculations using claims data from BCBS.

**Appendix Table B.2**: **Alternative Measures of Pricing According to Common Implicit Conversion Factors**

| Benchmarking Measure: | Services | Dollars | Services Q1 | Dollars Q1 |
|---|---|---|---|---|
| | *Panel A: 2008* | | | |
| Benchmarking Measure: | Services | Dollars | Services Q1 | Dollars Q1 |
| Rounding for ICFs: | | | | |
| *\$0.02* | 67% | 60% | 68% | 62% |
| *\$0.10* | 73% | 66% | 74% | 67% |
| *\$0.20* | 77% | 71% | 78% | 72% |
| | *Panel B: 2009* | | | |
| Benchmarking Measure: | Services | Dollars | Services Q1 | Dollars Q1 |
| Rounding for ICFs: | | | | |
| *\$0.02* | 67% | 60% | 68% | 62% |
| *\$0.10* | 73% | 66% | 74% | 67% |
| *\$0.20* | 77% | 70% | 78% | 71% |
| | *Panel C: 2010* | | | |
| Benchmarking Measure: | Services | Dollars | Services Q1 | Dollars Q1 |
| Rounding for ICFs: | | | | |
| *\$0.02* | 87% | 83% | 88% | 84% |
| *\$0.10* | 89% | 86% | 89% | 86% |
| *\$0.20* | 89% | 87% | 90% | 87% |
| | *Panel D: 2011* | | | |
| Benchmarking Measure: | Services | Dollars | Services Q1 | Dollars Q1 |
| Rounding for ICFs: | | | | |
| *\$0.02* | 86% | 81% | 86% | 82% |
| *\$0.10* | 87% | 85% | 88% | 85% |
| *\$0.20* | 88% | 85% | 88% | 85% |

Note: Each cell shows the share of services for which payments are associated with a common Implied Conversion Factor (cICF), as defined in the main text. The different cells within a panel show this statistic according to slightly different measures and using different rounding thresholds to define cICFs. The column labeled "Rounding" indicates the rounding applied to each estimated ICF. We then declare an ICF to be "common" for the payments to a physician group if it accounts for at least 5 percent of the group's services in a given year. The first column shows the share of services priced using cICFs, just as in Table 2. The column labeled "Dollars" shows a dollar-weighted measure. The dollar-weighted estimates are lower than the service-weighted measure because lower-value services are more likely to be priced using common ICFs. The remaining columns report equivalent measures for which the claims data are restricted to the first quarter of a given year. Source: Authors' calculations using claims data from BCBS.

**Appendix Table B.3: Firm Size and Implied Conversion Factors**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | *Panel A: Heterogeneity in Medicare Benchmarking* | | | | |
| Dependent variable: | Share of claims linked to Medicare | | | | |
| Firm Size (Log Spending) | -0.042** | | | -0.043** | -0.017** |
| | (0.005) | | | (0.005) | (0.005) |
| Firm Market Share | | -0.149** | | 0.110** | 0.041 |
| | | (0.022) | | (0.031) | (0.032) |
| Market Concentration | | | -0.139** | -0.191** | -0.078* |
| | | | (0.024) | (0.035) | (0.039) |
| | *Panel B: Firm Size and Implied Conversion Factors* | | | | |
| Dependent variable: | Log implied conversion factor (ICF) | | | | |
| Firm Size (Log Spending) | 0.058** | | | 0.058** | 0.040** |
| | (0.004) | | | (0.005) | (0.006) |
| Firm Market Share | | 0.241** | | -0.158** | -0.092** |
| | | (0.015) | | (0.037) | (0.029) |
| Market Concentration | | | 0.238** | 0.318** | 0.159** |
| | | | (0.020) | (0.036) | (0.028) |
| | *Panel C: Firm Size and Deviations from ICFs* | | | | |
| Dependent variable: | Log deviation from ICF | | | | |
| Firm Size (Log Spending) | 0.047** | | | 0.045** | 0.030** |
| | (0.008) | | | (0.008) | (0.007) |
| Firm Market Share | | 0.183** | | 0.021 | $0.080^{+}$ |
| | | (0.038) | | (0.054) | (0.047) |
| Market Concentration | | | 0.151** | 0.068 | -0.048 |
| | | | (0.042) | (0.061) | (0.049) |
| *N* | 20,736,449 | 20,736,449 | 20,736,449 | 20,736,449 | 20,736,449 |
| No. of Clusters | 23,098 | 23,098 | 23,098 | 23,098 | 23,098 |
| Code Effects | No | No | No | No | Yes |
| HSA Fixed Effects | No | No | No | No | Yes |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. Estimated at the claim level, this table shows the empirical relationship between features of physicians' contracts and measures of firm size and/or market concentration. The construction of all variables is discussed in the main text. Source: Authors' calculations using claims data from BCBS.

**Appendix Table B.4: Medicare Benchmarking by Betos Category**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | Payments with Common Conversion Factors | | | |
| | Spending Share | | Service Share | |
| | *Panel A: 2008* ($N$=516,189) | | | |
| Imaging | -0.172** | -0.287** | -0.132* | -0.256** |
| | (0.046) | (0.048) | (0.053) | (0.057) |
| Procedures | -0.169** | -0.282** | -0.150* | -0.271** |
| | (0.039) | (0.045) | (0.059) | (0.069) |
| Tests | -0.188** | -0.275** | -0.111* | -0.206** |
| | (0.041) | (0.044) | (0.054) | (0.059) |
| Constant | 0.730** | 0.405** | 0.750** | 0.398** |
| | (0.028) | (0.042) | (0.037) | (0.055) |
| | *Panel B: 2009* ($N$=593,779) | | | |
| Imaging | -0.158** | -0.270** | -0.105* | -0.228** |
| | (0.040) | (0.042) | (0.047) | (0.050) |
| Procedures | -0.159** | -0.287** | -0.144** | -0.282** |
| | (0.033) | (0.040) | (0.050) | (0.059) |
| Tests | -0.174** | -0.261** | -0.098+ | -0.193** |
| | (0.036) | (0.042) | (0.050) | (0.060) |
| Constant | 0.712** | 0.397** | 0.736** | 0.393** |
| | (0.027) | (0.036) | (0.033) | (0.048) |
| | *Panel C: 2011* ($N$=651,901) | | | |
| Imaging | -0.391** | -0.441** | -0.273** | -0.336** |
| | (0.058) | (0.053) | (0.031) | (0.026) |
| Procedures | -0.318** | -0.371** | -0.351** | -0.417** |
| | (0.026) | (0.025) | (0.053) | (0.051) |
| Tests | -0.384** | -0.420** | -0.258** | -0.303** |
| | (0.040) | (0.034) | (0.048) | (0.045) |
| Constant | 0.895** | 0.808** | 0.921** | 0.814** |
| | (0.012) | (0.013) | (0.019) | (0.018) |
| Omitted Category | Evaluation & Management | | | |
| Additional Controls | Group Size | None | Group Size | None |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. This table shows estimates of the $\nu_b$ coefficients in equation (13), namely the relationship between Betos category and the Medicare-linked share of services (columns 1 and 2) or spending (columns 3 and 4) at the group-service code level. Medicare links are measured using the common Implied Conversion Factors (cICFs) defined in section 3.1. Columns 1 and 3 show estimates after controlling for vigintile of group size, as measured with BCBS spending, and columns 2 and 4 show estimates without group size controls. Standard errors are two-way clustered (Cameron, Gelbach and Miller, 2011) by Betos category and physician group. Sources: Authors' calculations using claims data from BCBS.

# C Estimation in Changes and Threats to Identification

This appendix justifies our measure of Medicare benchmarking based on estimation in simple differences, in section C.1. Appendix C.2 then discusses potential bias from active renegotiations of physician-insurer contracts contemporaneously with the implementation of Medicar RVU updates. Finally, Appendix C.3 computes the bias that would result in such a case.

## C.1 Estimation in Changes

We simplify our main estimating equations to two time periods in order to see the Medicare-private price relationships as transparently as possible. This approach will also clearly highlight the assumptions necessary for our estimate of $\hat{\beta}$ to equal the true Medicare-linked share $\sigma$. Averaging equation (11) within each time period, and then taking the difference across the two, yields:

$$\Delta\overline{\ln(P_{g,j})} = \alpha + \beta\Delta\ln(RVU_j) + (1-\sigma)\overline{\varepsilon_{g,j}}. \tag{C.1}$$

In the context of price changes for one service, this equation shows how we can directly interpret the evidence from Figure 2C. This graph showed BCBS average log payments for a standard office visit increasing by 70 percent of the Medicare log RVU change. Hence the implied estimate of $\sigma$, in the absence of contemporaneous active negotiations, is also 70 percent.

## C.2 Threats to Identification From Active Renegotiations

This interpretation is threatened by the possibility of actively negotiated changes in $\ln(\theta_g)$ and $\ln(\rho_{g,j,p})$, which would show up in the error term. If they also covary with the updates to Medicare's relative values, then our estimate of $\hat{\beta}$ would be biased relative to the true parameter $\sigma$. (We compute the bias in Appendix C.3 below.) This might arise endogenously because changes in Medicare's relative values could alter groups' bargaining positions, and perhaps do so differentially across services. We quantify the potential influence of these changes on our estimates of Medicare's benchmarking in two ways.

First, note that when we estimate $\beta$ on the full sample of physician groups, it could be biased away from $\sigma$ by active renegotiations of both $\ln(\rho_{g,j,t})$ and $\ln(\theta_{g,t})$. If we estimate $\beta$ on the data for a single firm, however, $\Delta\ln(\theta_g)$ is a constant. In the levels specification of equation (11), we can similarly account for changes in each group's average log payment by allowing for a full set of group-by-period effects. If estimates of $\beta$ change little as a result of adding firm-by-period effects to such a specification, we can rule out the possibility that changes in the overall level of each firm's payments are biasing our attempt to recover $\sigma$.

Second, the channel through which active renegotiations might bias our attempt to recover $\sigma$ involves changes in bargaining power *induced* by the RVU changes.[35] The threat to

[35] Actively negotiated payment changes that are driven by the RVU updates themselves may plausibly

our estimation takes the following form: BCBS may pursue renegotiations with firms whose average Medicare payment has fallen, with these negotiations resulting in declines in their payments. Similarly, physician groups whose average Medicare payment has increased may pursue renegotiations with BCBS, with these negotiations resulting in increases in their payments. This pattern would imply a positive bias to our estimates of $\sigma$. To investigate the potential relevance of this source of bias, we first construct the average change in the RVUs for the specific services provided by each firm. This allows us to gauge the extent to which each firm is affected. We then investigate whether we obtain larger estimates $\hat{\beta}$ on a sample of firms that were significantly affected compared with firms that experienced little change in their average RVUs.

## C.3 Deriving the Bias in our Medicare Link Estimate

The biased coefficient $\hat{\beta}$ we would estimate from equation (C.1) in the presence of simultaneous updates to non-benchmarked prices or group-specific markups is:

$$\begin{aligned}
\hat{\beta} &= \frac{\mathrm{Cov}[\Delta\overline{\ln(P_{g,j})}, \Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} \\
&= \frac{\mathrm{Cov}[\sigma\Delta\overline{\ln(\phi_g)} + \sigma\Delta\ln(RVU_j) + (1-\sigma)\Delta\overline{\ln(\rho_{g,j})} + \Delta\epsilon_{g,j,p}, \Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} \\
&= \sigma\frac{\mathrm{Cov}[\Delta\ln(RVU_j), \Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} + \sigma\frac{\mathrm{Cov}[\Delta\overline{\ln(\phi_g)}, \Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} \\
&+ (1-\sigma)\frac{\mathrm{Cov}[\Delta\overline{\ln(\rho_{g,j})}]}{\mathrm{Var}[\Delta\ln(RVU_j)]} + \frac{\mathrm{Cov}[\Delta\epsilon_{g,j,p}, \Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} \\
&= \sigma + \sigma\frac{\mathrm{Cov}[\Delta\overline{\ln(\phi_g)}, \Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} + (1-\sigma)\frac{\mathrm{Cov}[\Delta\overline{\ln(\rho_{g,j})}, \Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]},
\end{aligned} \tag{C.2}$$

where the third equality follows from the properties of covariances and the fourth from the fact that $\frac{\mathrm{Cov}[\Delta\ln(RVU_{j,t}),\Delta\ln(RVU_j)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} = 1$ and $\frac{\mathrm{Cov}[\Delta\epsilon_{g,j,t},\Delta\ln(RVU_j^M)]}{\mathrm{Var}[\Delta\ln(RVU_j)]} = 0$.

One separate source of bias in the estimate of $\hat{\beta}$ could arise if the linked share $\sigma$ varies across firms and services. This would imply additional terms in equation (C.2) describing our regression estimates, involving covariances between the RVU updates used for identification and the service-by-group linked shares $\sigma_{j,g}$. Recovering $\sigma$ also requires us to assume that these covariance terms are 0, which will be true if updates to Medicare's rates are uncorrelated with the $\sigma_{j,g}$. In section 5.2, we will allow for heterogeneity across various dimensions in the linked shares.

When thinking about our estimates and any potential bias, it is essential to remember that our estimates are based only on those services with RVU updates. The service code

covary with these changes. There is no *a priori* reason to suspect that changes renegotiated for other reasons would covary with the RVU updates and bias our estimates.

fixed effects ensure that codes without updates don't influence $\hat{\beta}$.

## C.4 Checks for the Relevance of Active Contract Renegotiation

The estimates presented in Figure 2 and Table 4 may differ from the true Medicare benchmarking parameter $\sigma$ if changes in other terms of providers' contracts covary with the changes in RVUs. Indeed, payment changes that significantly alter physician groups' average Medicare payment can move private payments in subsequent years, due in part to the resulting changes to their bargaining positions (Clemens and Gottlieb, 2017). We thus draw on institutional detail and theoretically motivated specification checks to explore how much our estimates might deviate from the true share of payments benchmarked to Medicare's relative values.

The most relevant institutional detail is the relatively short time horizon of our event studies. Dunn and Shapiro (2015) report that physician contracts tend to remain in force for around 3 years. Within each of our single-year event studies, we thus anticipate that roughly one-third of the groups in our sample engage in active contract re-negotiations, which could affect our estimates. Unlike the payment changes analyzed by Clemens and Gottlieb (2017), which significantly shifted certain specialties' average Medicare payments, those we consider here are relatively diffused across specialties, so unlikely to affect groups' overall outside options.

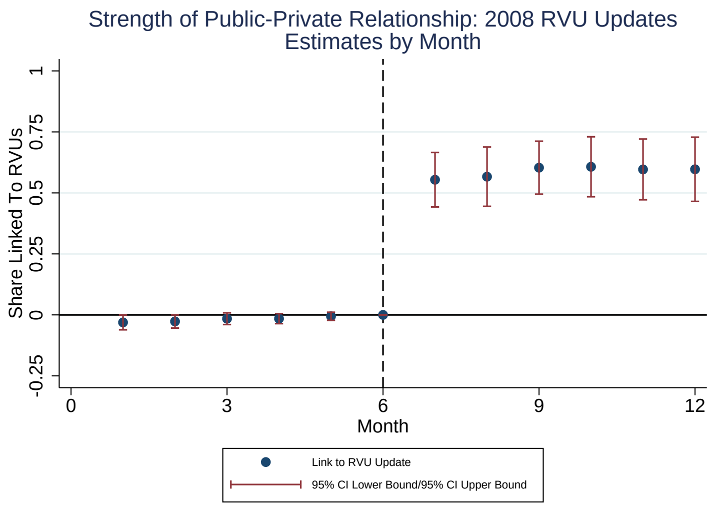
Nevertheless, we investigate the potential relevance of active contract renegotiation with two analyses. First, we consider the potential effect of scheduled RVU changes on a firm's bargaining position. We construct a variable that, for each firm, reports the average change in RVUs for the services it provides. Firms experiencing a negative average change have seen their bargaining positions deteriorate. Firms experiencing an average RVU increase have seen their bargaining positions improve. Using the average RVU change to which each firm was exposed, we construct an indicator for groups whose bargaining positions were significantly affected.

Second, we investigate the potential relevance of changes in groups' average log reimbursement by adding full sets of group-by-period fixed effects to our specification. For this regression, we restrict our sample to the 100 largest firms in each year, primarily for computational ease. Note, however, that large firms are precisely those for which we would expect active renegotiations to be most frequent.

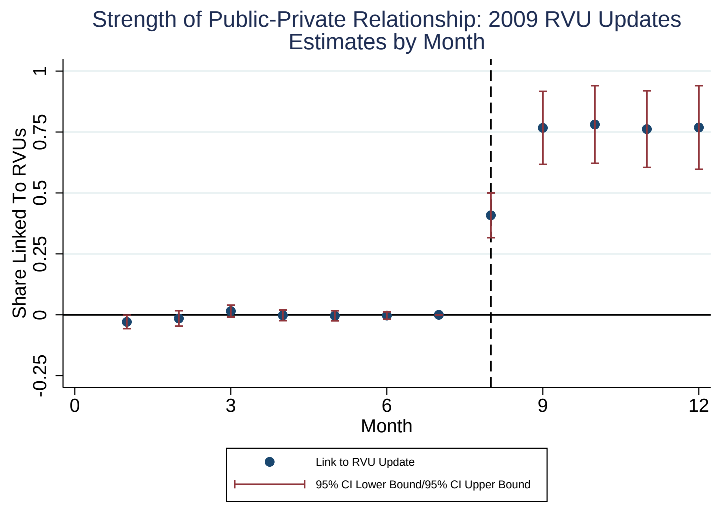
Table C.3 presents these results. Column 1 reports our baseline specification, unchanged from Table 4. Column 2 allows our coefficient of interest to vary with an indicator for whether a firm's average Medicare reimbursement rate was significantly affected by a year's RVU updates. The point estimate on this interaction varies across years, but is negative in each case. This is the opposite of what we would expect if significant RVU updates were driving active contract renegotiations. Column 3 limits the baseline specification to the services provided by the 100 largest physician groups. A comparison of column 3 with column 1 reveals that, on average across the years we analyze, the largest firms have contracts that are less linked to Medicare than are contracts in the full sample, a result that we explore further in section 5.1. Most relevant for our current purposes, however, column 4 reveals that adding group-by-period effects to the previous specification has essentially no impact on our coefficient of interest. These results provide evidence against the concern that that active contract renegotiations confound the relationship between BCBS's and Medicare's payments over the intervals we analyze. Thus they bolster the case for interpreting our estimates of $\hat{\beta}$ as unbiased estimates of the fraction of services tied directly to Medicare.

## Appendix Figure C.1: Strength of Public Private Payment Relationships

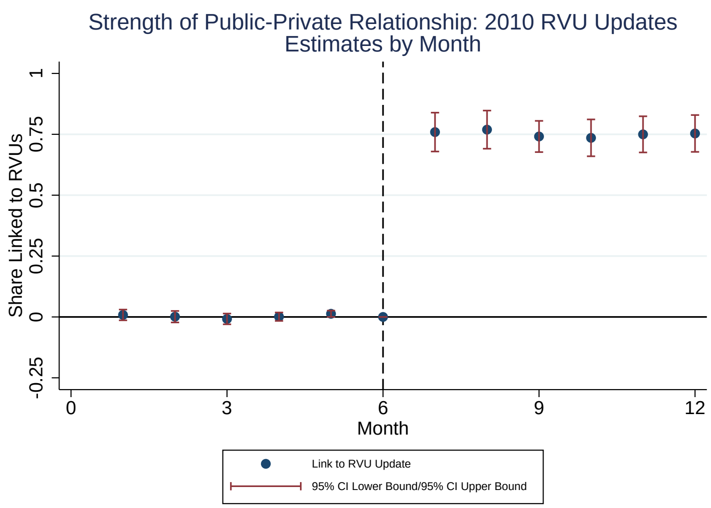
*Panel A*



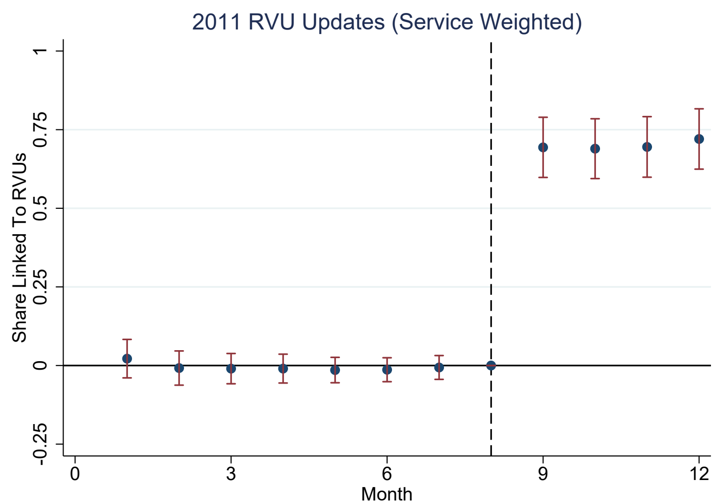

*Panel B*




*Panel C*

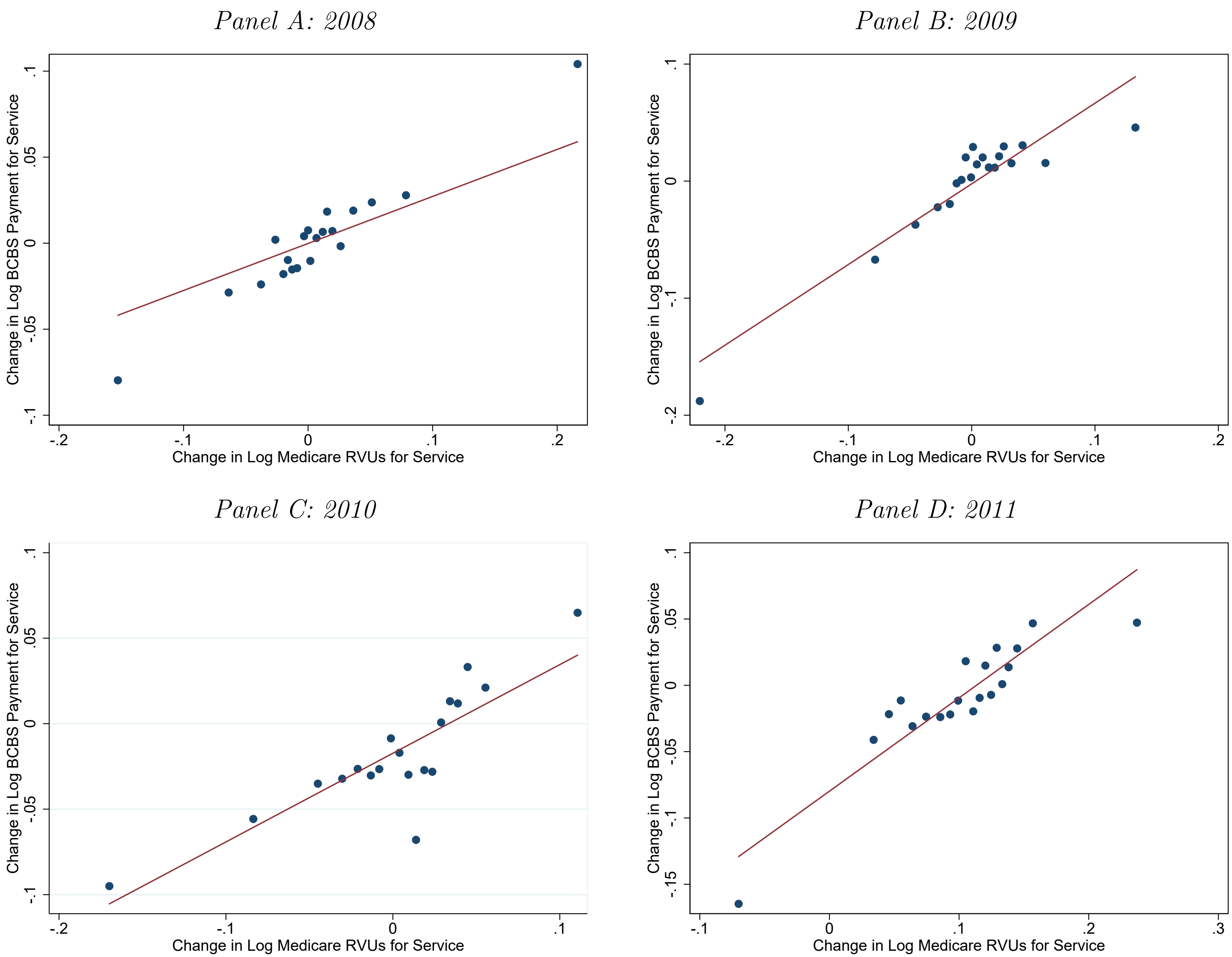



*Panel D*




Note: The figure reports estimates of the $\beta_p$ from estimates of equation (12). The vertical dashed line in each panel corresponds with the month during each year in which BCBS implemented its update from the prior year's relative value scale. These updates occurred on July 1, 2008, August 15, 2009, July 1, 2010, and September 1, 2011.

**Appendix Figure C.2: Benchmarking Estimates Based on Price Changes Across Services**




Note: The figure reports the relationships described by equation (C.1) for RVU updates in each year, and estimates of that equation. Each panel shows a separate year's estimates, measured as log differences between the period before BCBS implemented the Medicare RVU updates and the period after this update. The years are split at July 1, 2008, August 15, 2009; July 1, 2010; and September 1, 2011. The regressions are run at the underlying service level, but observations are grouped into twenty bins for each year, based on vigintiles of the Medicare log RVU change.

**Appendix Table C.1**: **Other Years' Estimates of Benchmarking Using RVU Changes**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | |
| | *Panel A: All Services: 2008 RVU Updates* | | | |
| Log RVU Change × Post | 0.602** | 0.597** | 0.539** | 0.602** |
| | (0.061) | (0.061) | (0.060) | (0.061) |
| $N$ | 19,552,096 | 19,552,096 | 19,552,096 | 19,552,096 |
| No. of Clusters | 3,505 | 3,505 | 3,505 | 3,505 |
| | *Panel B: All Services: 2009 RVU Updates* | | | |
| Log RVU Change × Post | 0.778** | 0.778** | 0.792** | 0.778** |
| | (0.081) | (0.078) | (0.070) | (0.081) |
| $N$ | 21,941,227 | 21,941,227 | 21,941,227 | 21,941,227 |
| No. of Clusters | 3,807 | 3,807 | 3,807 | 3,807 |
| | *Panel C: All Services: 2011 RVU Updates* | | | |
| Log RVU Change × Post | 0.704** | 0.689** | 0.679** | 0.704** |
| | (0.046) | (0.052) | (0.048) | (0.046) |
| $N$ | 25,404,007 | 25,404,007 | 25,404,007 | 25,404,007 |
| No. of Clusters | 4,091 | 4,091 | 4,091 | 4,091 |
| Group-by-Code Effects | Yes | No | Yes | Yes |
| Code Effects | No | Yes | No | No |
| Cubic Time x RVU Change | No | No | Yes | No |
| Cubic Time x Post | No | No | No | Yes |
| Weighting | Service | Service | Service | Service |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the forms described in section 3.2. Each column in each panel reports an estimate of $\hat{\beta}$ from equation (11). Panel A shows estimates using RBRVS updates and BCBS claims data for 2008, Panel B for 2009, and Panel C for 2011. Observations are at the claim-line level and are equally weighted. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Appendix Table C.2**: **Dollar-Weighted Estimates of Benchmarking Using RVU Changes**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | |
| | *Panel A: All Services: 2008 RVU Updates* | | | |
| Log RVU Change × Post | 0.421** | 0.413** | 0.359** | 0.420** |
| | (0.075) | (0.075) | (0.071) | (0.075) |
| $N$ | 19,552,096 | 19,552,096 | 19,552,096 | 19,552,096 |
| No. of Clusters | 3,505 | 3,505 | 3,505 | 3,505 |
| | *Panel B: All Services: 2009 RVU Updates* | | | |
| Log RVU Change × Post | 0.618** | 0.627** | 0.669** | 0.618** |
| | (0.046) | (0.045) | (0.052) | (0.046) |
| $N$ | 21,941,227 | 21,941,227 | 21,941,227 | 21,941,227 |
| No. of Clusters | 3,807 | 3,807 | 3,807 | 3,807 |
| | *Panel C: All Services: 2011 RVU Updates* | | | |
| Log RVU Change × Post | 0.749** | 0.739** | 0.738** | 0.749** |
| | (0.044) | (0.043) | (0.047) | (0.044) |
| $N$ | 25,404,007 | 25,404,007 | 25,404,007 | 25,404,007 |
| No. of Clusters | 4,091 | 4,091 | 4,091 | 4,091 |
| Group-by-Code Effects | Yes | No | Yes | Yes |
| Code Effects | No | Yes | No | No |
| Cubic Time × RVU Change | No | No | Yes | No |
| Cubic Time × Post | No | No | No | Yes |
| Weighting | Dollars | Dollars | Dollars | Dollars |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the forms described in section 3.2. Each column in each panel reports an estimate of $\hat{\beta}$ from equation (11). Panel A shows estimates using RBRVS updates and BCBS claims data for 2008, Panel B for 2009, and Panel C for 2011. Observations are at the claim-line level and are weighted according to each service's average payment during the baseline period. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Appendix Table C.3: Checks for the Relevance of Active Contract Negotiations**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | |
| | *Panel A: All Services: 2009 RVU Updates* | | | |
| Log RVU Change × Post | 0.778** | 0.847** | 0.696** | 0.666** |
| | (0.081) | (0.085) | (0.093) | (0.081) |
| Log RVU Change × Post × Update Impact | | -0.077 | | |
| | | (0.114) | | |
| $N$ | 21,941,227 | 21,941,227 | 4,097,283 | 4,097,283 |
| No. of Clusters | 3,807 | 3,807 | 3,496 | 3,496 |
| | *Panel B: All Services: 2010 RVU Updates* | | | |
| Log RVU Change × Post | 0.750** | 0.992** | 0.740** | 0.747** |
| | (0.038) | (0.076) | (0.048) | (0.052) |
| Log RVU Change × Post × Update Impact | | -0.393** | | |
| | | (0.099) | | |
| $N$ | 23,933,577 | 23,933,577 | 4,708,213 | 4,708,213 |
| No. of Clusters | 3,681 | 3,681 | 3,450 | 3,450 |
| | *Panel C: All Services: 2011 RVU Updates* | | | |
| Log RVU Change × Post | 0.704** | 0.804** | 0.544** | 0.523** |
| | (0.046) | (0.084) | (0.051) | (0.067) |
| Log RVU Change × Post × Update Impact | | -0.162 | | |
| | | (0.106) | | |
| $N$ | 25,404,007 | 25,404,007 | 5,069,260 | 5,069,260 |
| No. of Clusters | 4,091 | 4,091 | 3,825 | 3,825 |
| Group × Post-Update Effects | No | No | No | Yes |
| Sample | Full | Full | Largest Firms | Largest Firms |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the forms described in section 3.2. Column 1 replicates the baseline specification from column 1 of Table 4. Column 2 augments the baseline specification with interaction terms allowing the effect of RVU updates to vary with the extent of the average impact of each year's RVU updates on a physician group's average Medicare reimbursement rate. In columns 3 and 4 the sample is restricted to each year's 100 largest physician groups, as sorted by total bills submitted. The specification in column 3 is the baseline specification, while the specification in column 4 includes a full set of post-by-group interactions. Panel A shows estimates using RBRVS updates and BCBS claims data for 2009, Panel B for 2010, and Panel C for 2011. Observations are at the claim-line level and are equally weighted. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

## Appendix Table C.4: Public-Private Payment Links Across Service Categories

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | | | | |
| | *Panel A: 2008 RVU Updates by Betos Categories* | | | | | | |
| | Evaluation | Imaging | Procedures | Tests | Imaging Sub-Categories: | | |
| | | | | | Global | Technical | Professional |
| Log RVU Change | 0.541*** | 0.644*** | 0.495*** | 0.786*** | 0.665*** | 0.494*** | 0.945*** |
| × Post-Update | (0.115) | (0.092) | (0.116) | (0.055) | (0.103) | (0.112) | (0.228) |
| $N$ | 9,851,995 | 3,221,634 | 3,851,609 | 1,292,912 | 1,688,102 | 192,569 | 1,340,963 |
| No. of Clusters | 207 | 1,069 | 1,817 | 385 | 400 | 235 | 434 |
| | *Panel B: 2009 RVU Updates by Betos Categories* | | | | | | |
| | Evaluation | Imaging | Procedures | Tests | Imaging Sub-Categories: | | |
| | | | | | Global | Technical | Professional |
| Log RVU Change | 0.857** | 0.775** | 0.399** | 0.933** | 0.702** | 0.769** | 0.680** |
| × Post-Update | (0.209) | (0.066) | (0.064) | (0.052) | (0.072) | (0.068) | (0.184) |
| $N$ | 11,498,770 | 3,524,642 | 3,861,539 | 1,449,803 | 1,769,522 | 222,026 | 1,533,094 |
| No. of Clusters | 219 | 1,133 | 2,036 | 388 | 422 | 262 | 449 |
| | *Panel C: 2011 RVU Updates by Betos Categories* | | | | | | |
| | Evaluation | Imaging | Procedures | Tests | Imaging Sub-Categories: | | |
| | | | | | Global | Technical | Professional |
| Log RVU Change | 0.794** | 0.616** | 0.900** | 0.439* | 0.816** | 0.692** | 0.709** |
| × Post-Update | (0.065) | (0.100) | (0.075) | (0.221) | (0.048) | (0.067) | (0.058) |
| $N$ | 13,116,657 | 3,696,733 | 5,233,336 | 1,659,485 | 1,929,095 | 193,577 | 1,574,061 |
| No. of Clusters | 238 | 1,143 | 2,246 | 436 | 424 | 264 | 455 |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the forms described in section 3.2. The cells in each panel report estimates of $\hat{\beta}$ from equation (11), with samples selected to contain the HCPCS codes falling into individual broad service categories. The name of the relevant service category accompanies each point estimate. Panel A shows estimates using RBRVS updates and BCBS claims data for 2008, Panel B for 2009, and Panel C for 2011. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Appendix Table C.5**: **Medicare Benchmarking by Firm Size**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | |
| *Panel A: 2008 RVU Updates* (N = 19,552,096) | | | | |
| Log RVU Change | 0.602** | 0.560** | 0.421** | 0.418** |
| × Post-Update | (0.061) | (0.074) | (0.075) | (0.089) |
| Log RVU Change | | 0.130* | | -0.059 |
| × Post-Update × Midsize | | (0.065) | | (0.072) |
| Log RVU Change | | -0.000 | | 0.064 |
| × Post-Update × Large | | (0.101) | | (0.085) |
| *Panel B: 2009 RVU Updates* (N = 21,941,227) | | | | |
| Log RVU Change | 0.778** | 0.755** | 0.618** | 0.756** |
| × Post-Update | (0.081) | (0.090) | (0.046) | (0.070) |
| Log RVU Change | | 0.078 | | -0.110 |
| × Post-Update × Midsize | | (0.059) | | (0.071) |
| Log RVU Change | | -0.035 | | -0.271* |
| × Post-Update × Large | | (0.094) | | (0.109) |
| *Panel C: 2011 RVU Updates* (N = 25,404,007) | | | | |
| Log RVU Change | 0.704** | 0.812** | 0.749** | 0.774** |
| × Post-Update | (0.046) | (0.063) | (0.044) | (0.052) |
| Log RVU Change | | -0.140+ | | -0.036 |
| × Post-Update × Midsize | | (0.075) | | (0.100) |
| Log RVU Change | | -0.183* | | -0.023 |
| × Post-Update × Large | | (0.075) | | (0.116) |
| Firm Size × Post-Update Controls | No | Yes | No | Yes |
| Weighting | Services | Services | Dollars | Dollars |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. Columns 1 and 3 report the baseline estimates from Table 4 Panels A and B respectively. In columns 2 and 4 we augment these specifications to include interactions between firm size indicators variables and both the "Post" indicator and the interaction between the "Log RVU Change" and "Post" indicator. The omitted category is small firms, defined as those with less than $200,000 in billings. Mid-sized firms are those with billings between $200,000 and $1 million, and large firms are those with billings exceeding $1 million. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

# D Extensions

## D.1 External Validity: Data from Colorado

To gauge the external validity of our BCBS findings, we also use data from the state of Colorado's All-Payer Claims Database (APCD). The APCD is administered by the Center for Improving Value in Health Care (CIVHC), from whom we obtain the data. These data are structured similarly to the BCBS data, but include many insurers. Our analysis focuses on the payments of one Colorado insurer for which we were able to obtain the necessary information on fee schedule updates to implement our primary estimation framework.

In this exercise, we conduct a similar analysis to our baseline changes estimation using the claims data from this Colorado insurer. Just as in Table 4 in the main text, we use institutional detail on the timing with which the insurer adopted updates to Medicare's payments. This insurer adopts Medicare's payment updates less frequently than BCBS of Texas. Over the period for which we have the Colorado claims data, this insurer shifted from the 2008 version of Medicare's relative rate structure to the 2010 version on July 1, 2011.

The results are shown in Figure D.2 and Table D.1. The table reports our estimates of equation (12) using these changes. The estimates imply that roughly 40 percent of its payments are linked to Medicare's relative rate structure. This insurer thus appears to deviate from Medicare payments to a greater degree than does BCBS of Texas.

Beyond the insurers we analyze directly, anecdotal and documentary evidence suggest that our results apply more broadly. Provider newsletters and industry magazines describe price-setting that frequently uses Medicare as a benchmark, but sometimes deviates. Among many examples, *Managed Care* magazine writes that insurers' "talks with doctors on fee-for-service rates often begin with Medicare's rates" (Carroll, 2007). Clemens and Gottlieb (2017) provide other related examples.

## D.2 Supply Responses

To determine whether pricing at this granular level has impacts on real resource use, we estimate how relative price changes across services affect physicians' supply of care. We use the same Medicare price changes as in our main analysis, which means that our estimates have three economically salient features. First, they involve short-run responses within a calendar year. Second, they involve responses to changes in the profitability of some services relative to others rather than to across-the-board changes in reimbursement rates.[36] Finally, they involve private payment changes that result from contractual links to changes in Medicare's relative rates.[37]

[36]Our focus here on relative supply responses across services makes this analysis somewhat comparable to Gruber, Kim and Mayzlina (1999) or, more recently, Brekke, Holmås, Monstad and Straume (2015).

[37]This final characteristic makes it natural to think about a clear causal chain in our setting where prices influence the subsequent supply of care. When Medicare reimbursement changes lead to a renegotiation, as might happen over a longer time horizon, then we would have to consider the price renegotiations and supply decisions as jointly determined.

These features limit the external validity of these estimates. Changes in the overall structure of reimbursements, or in whether a given price is set through benchmarking, may engender different supply responses. These estimates are also subject to more significant endogeneity concerns than our price benchmarking estimates, since Medicare prices may respond to utilization, even if not to private insurance prices.

With these limitations in mind, we estimate an analogue of the changes regression shown in Panel C of Figure 2 in which the dependent variable is now the change in log quantity of care. We again split the year into two time periods: before and after BCBS implemented the year's Medicare updates. The change in the log number of instances that a given physician group provided a particular service across these two time periods is our dependent variable.

Figure D.3 shows the results of this estimation for each year, along with a binned scatterplot of the underlying data. Note that this is a reduced-form estimate; it relates the Medicare price change to the supply responses for privately insured patients.

In order to estimate the BCBS own-price supply elasticities, we next develop an IV framework. We use the same reimbursement changes that follow from Medicare's RVU updates in the following two-stage least squares setup:

$$\Delta\overline{\ln(P_{g,j})} = \alpha + \beta\Delta\ln(RVU_j) + \varepsilon_{g,j} \tag{D.1}$$

$$\Delta\ln(Q_{g,j}) = \gamma + \delta\Delta\widehat{\overline{\ln(P_{g,j})}} + \epsilon_{g,j}. \tag{D.2}$$

The first stage, equation (D.1), is taken from equation (C.1) in the text. This estimates the share of private prices that respond to the Medicare RVU updates. This generates a predicted price change, which we use in the second stage equation (D.2).

The coefficient $\delta$ that we estimate in equation (D.2) is close to providing an estimate of the physicians' supply elasticity for BCBS patients, in response to BCBS prices. It is somewhat confounded, however, by the fact that the BCBS prices are changing at the same time as the prices of physicians' outside option—treating Medicare patients.[38] This would tend to bias the estimates down relative to a pure own-price supply estimate.

Table D.2 shows the results. The IV estimates scale up the reduced form estimates substantially, and range from 0.15 to 0.66. The median estimate of 0.37 occurs in 2011. For comparison, the conceptually most similar estimates in the literature are those of Brekke et al. (2015). Brekke et al. (2015) estimate physicians' supply responses to a reimbursement change for one particular service, which is also the type of price change we consider here. These are different types of elasticities than those of Clemens and Gottlieb (2014), who consider market-wide changes, or the relative price changes of Gruber et al. (1999) and Jacobson, Earle, Price and Newhouse (2010).

These positive supply elasticities imply that the pricing decisions we examine have meaningful implications for how physicians provide treatment. If Medicare sets prices inefficiently, then copying Medicare's relative prices leads to inefficient care. When insurers deviate from Medicare rates, these positive supply responses suggest that physicians respond to payment adjustments as the insurers presumably intend.

[38] Clemens and Gottlieb (2013, Appendix B) model these forces.

## D.3 Out-of-Network Payments

Our analysis thus far only includes in-network payments—those made to physician groups that have agreed with BCBS on mutually acceptable payment rates. We next show analogous results for out-of-network payments, which occur when providers have not reached any such agreement. When a BCBS-insured patient sees an out-of-network provider, the ultimate payment reflects a complex interaction of the provider's charge, after-the-fact negotiations (as in Mahoney, 2015), and the insurance plan's coverage. So out-of-network payments are less likely to depend on a convenient benchmark such as the Medicare fee schedule. This analysis allows us to determine whether the benchmarking that we document reflects active decisions as opposed to a purely mechanical force.

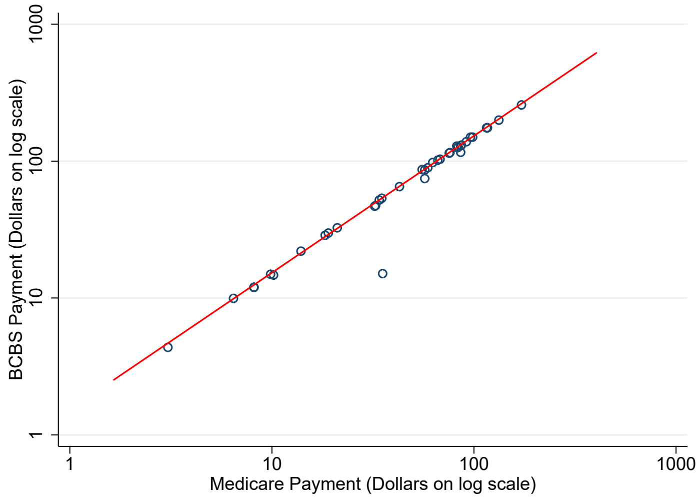
Table D.3 replicates Table 4 in the main text, but for out-of-network payments. Table D.4 is a dollar-weighted version of the same regressions. In both cases, we obtain small and precisely estimated coefficients. This means that out-of-network payments—which don't represent the outcome of the *ex ante* negotiations we described in section 1.2—are not priced in the same way.

Table D.5 complicates the analysis somewhat. It reveals that around half of out-of-network services appear to be priced according to cICFs. This share is much larger than the results from Tables D.3 and D.4 would suggest, though still far below the in-network results from Table 2 in the main text. The difference with the in-network results is especially pronounced in 2010 and 2011, and when using a more stringent cICF threshold (20 percent). In these cases, only 30 percent of out-of-network prices appear to be benchmarked to Medicare, compared with 70 percent of in-network payments. Nevertheless, the ambiguity over the correct definition again demonstrates the advantage of the update-based benchmarking measure in Tables D.3 and D.4.

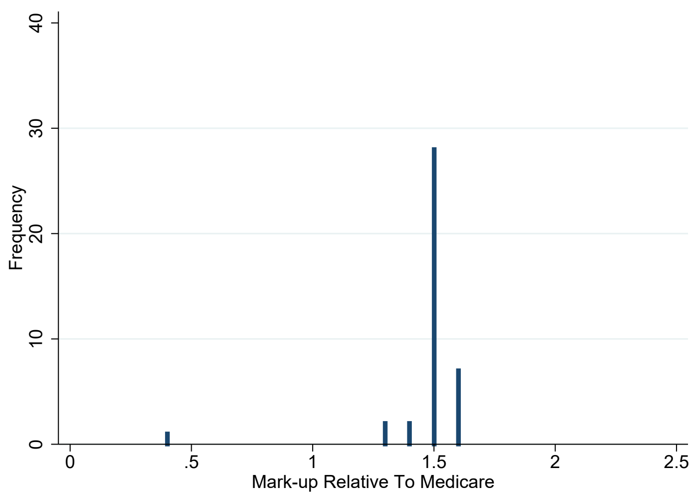
In short, we find much weaker—if any—Medicare benchmarking in out-of-network payments. The difference between these results and our in-network estimates suggests that the in-network prices reflect active efforts to negotiate around a simplified payment schedule.

**Appendix Figure D.1: Raw Payments For Illustrative Physician Group**

*Panel A: Early 2011 Payment Data for a Large Colorado Group*

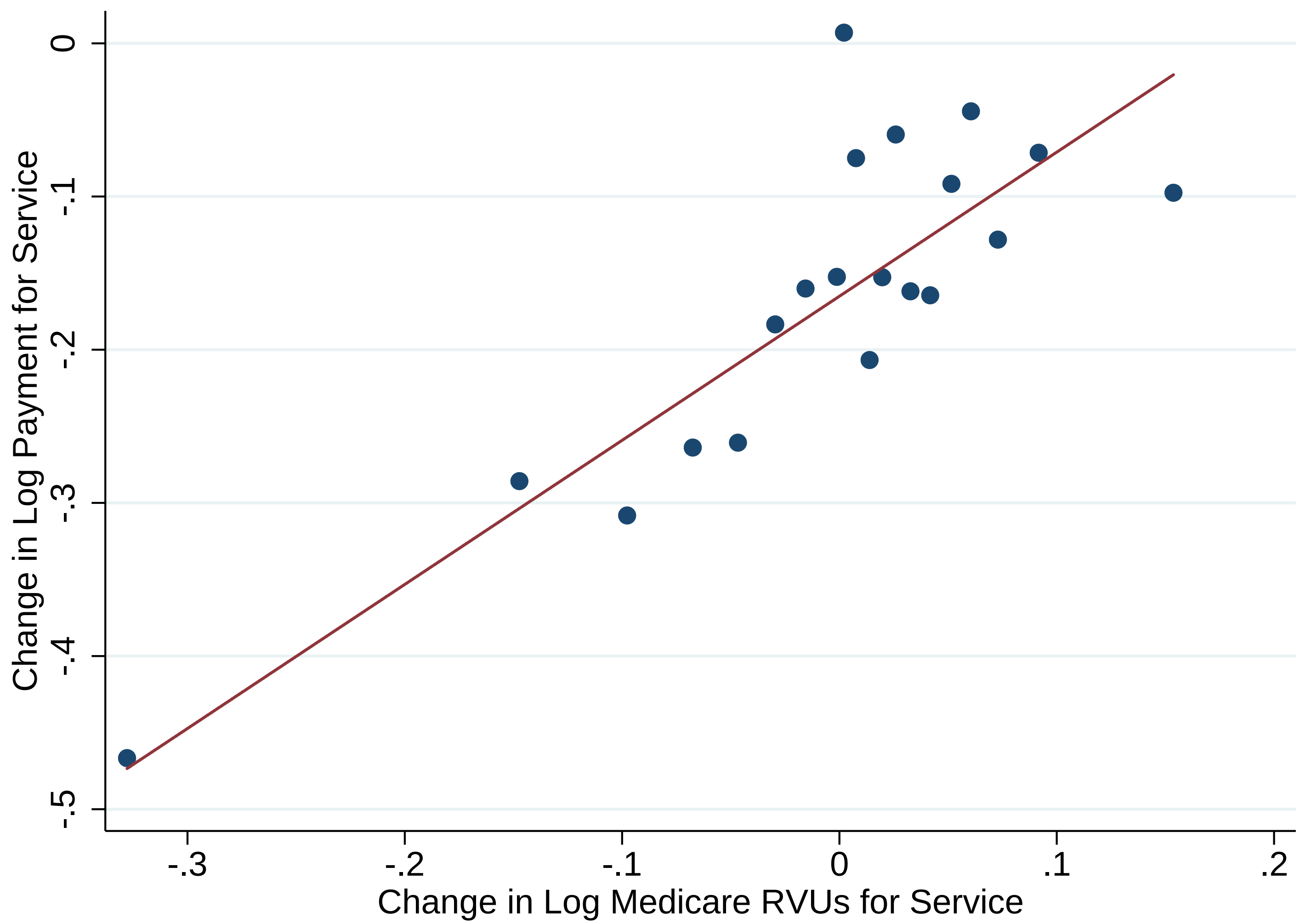



*Panel B: Distribution of Markups for a Large Colorado Group*



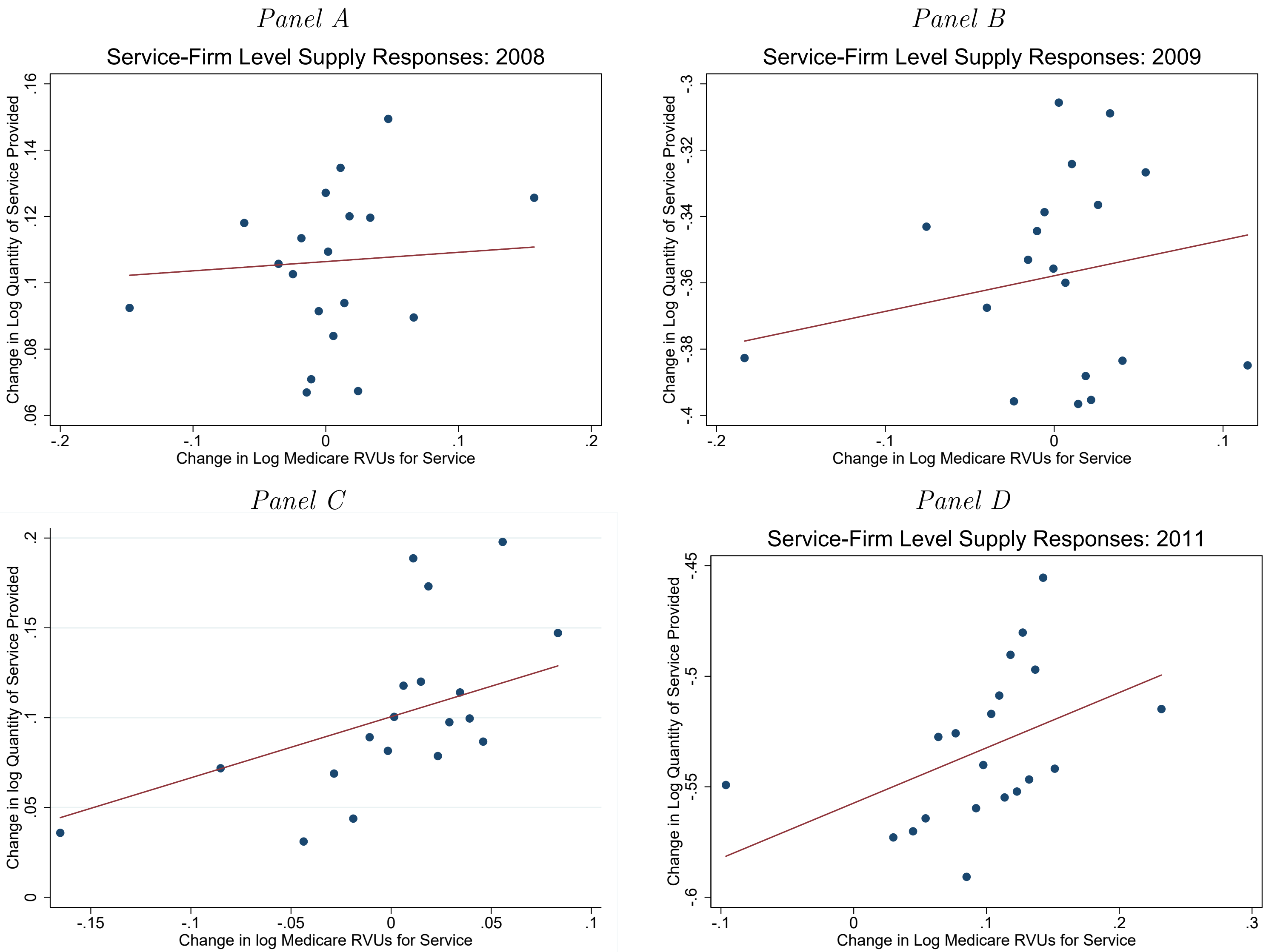

Note: This figure shows a payment scatterplot and distribution of markups analogous to those in Figure 1, but for a physician group in Colorado. In Panel A, each observation is a unique reimbursement paid for a particular service to the group. The line has a slope of 1 (in logs) and represent the group's most common Implied Conversion Factor. Panel B plots the distribution of markups relative to the Medicare rates for all payments the group received. It shows a clear spike at the value that we identify as a common Implied Conversion Factors in Panel A. To comply with confidentiality rules, we omit from these graphs a small share of each group's claims. The share of claims whose observations are suppressed is 17.17%. Sources: Authors' calculations using claims data from CO APCD.

**Appendix Figure D.2: Validating Bunching-Based Benchmarking Measure**




Note: This graph is analogous to Panel C of Figure 2, but using 2011 data from one insurer in the Colorado APCD, with the sample split at July 1, 2011. Private price changes are computed as the difference between service-level average payments after and before July 1, 2010. The regression is estimated at the underlying service-code level. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from Colorado APCD.

## Appendix Figure D.3: Short-Run Supply Responses to Medicare Price Changes




Note: The figure reports estimates of physicians' supply responses to Medicare price changes that BCBS implemented in a given year. Quantities, the dependent variable, are computed at the service-by-firm level. Each panel shows a separate year's estimates, measured as log differences between the period before BCBS implemented the Medicare RVU updates and the period after this update. The years are split at July 1, 2008, August 15, 2009; July 1, 2010; and September 1, 2011. The estimates have very different intercepts across the three panels because of the differences in the share of the year's data that are included in the periods before *vs.* after each year's update.

**Appendix Table D.1**: **Estimating Medicare Benchmarking Using RVU Changes: Colorado**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. variable: | Log private reimbursement rate | | | |
| Log RVU change | 0.368** | 0.354** | 0.455** | 0.408** |
| × Post | (0.145) | (0.146) | (0.098) | (0.102) |
| Provider-Code FE | | Yes | | Yes |
| Insurance Plan Controls | | | Yes | Yes |
| $N$ | 509,929 | 509,929 | 508,038 | 508,038 |
| No. Clusters | 1,471 | 1,471 | 1,471 | 1,471 |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the forms described in section 3.2. Each column reports an estimate of $\hat{\beta}$ from equation (11). Observations are at the claim-line level and are equally weighted. Data are from 2011. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from the Colorado APCD.

**Appendix Table D.2**: **Supply Elasticity Estimates**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Year: | 2008 | 2009 | 2010 | 2011 |
| Dependent variable: | Change in log service quantity | | | |
| | *Panel A: Reduced Form* | | | |
| Log RVU change for service | 0.027 | 0.095* | 0.339*** | 0.252*** |
| | (0.047) | (0.047) | (0.050) | (0.038) |
| | *Panel B: IV Estimates* | | | |
| Log BCBS payment change for service | 0.052 | 0.152* | 0.658*** | 0.365*** |
| | (0.090) | (0.076) | (0.102) | (0.055) |
| $N$ | 63,526 | 71,354 | 81,294 | 89,936 |
| First Stage $F$-Statistic | 358.9 | 776.2 | 483.9 | 1843.0 |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. Panel A estimates reduced form relationships analogous to the one shown in Panel D of Figure 2 in the main text. Panel B shows the second stage estimates from the IV framework in equation (D.2). The robust first-stage F-statistics all easily satisfy the weak instruments test of Olea and Pflueger (2013). Source: Authors' calculations using claims data from BCBS.

**Appendix Table D.3: Estimating Medicare Benchmarking for Out-of-Network Payments Using RVU Changes**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | |
| | *Panel A: All Services: 2009 RVU Updates* | | | |
| Log RVU Change × Post | 0.018 | 0.007 | 0.084* | 0.018 |
| | (0.043) | (0.047) | (0.037) | (0.043) |
| $N$ | 2,585,681 | 2,585,681 | 2,585,681 | 2,585,681 |
| No. of Clusters | 2,456 | 2,456 | 2,456 | 2,456 |
| | *Panel B: All Services: 2010 RVU Updates* | | | |
| Log RVU Change × Post | 0.302** | 0.351** | 0.170** | 0.302** |
| | (0.073) | (0.074) | (0.044) | (0.073) |
| $N$ | 2,386,575 | 2,386,575 | 2,386,575 | 2,386,575 |
| No. of Clusters | 2,051 | 2,051 | 2,051 | 2,051 |
| | *Panel C: All Services: 2011 RVU Updates* | | | |
| Log RVU Change × Post | 0.106* | 0.094+ | 0.047 | 0.105* |
| | (0.047) | (0.054) | (0.037) | (0.047) |
| $N$ | 2,626,264 | 2,626,264 | 2,626,264 | 2,626,264 |
| No. of Clusters | 2,473 | 2,473 | 2,473 | 2,473 |
| Group-by-Code Effects | Yes | No | Yes | Yes |
| Code Effects | No | Yes | No | No |
| Cubic Time × RVU Change | No | No | Yes | No |
| Cubic Time × Post | No | No | No | Yes |
| Weighting | Service | Service | Service | Service |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the forms described in section 3.2, except using data from out-of-network payments. Each column in each panel reports an estimate of $\hat{\beta}$ from equation (11). Panel A shows estimates using RBRVS updates and BCBS claims data for 2009, Panel B for 2010, and Panel C for 2011. Observations are at the claim-line level and are equally weighted. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Appendix Table D.4: Dollar-Weighted Estimates of Medicare Benchmarking for Out-of-Network Payments Using RVU Changes**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: | *Log private reimbursement rate* | | | |
| | *Panel A: All Services: 2009 RVU Updates* | | | |
| Log RVU Change × Post | 0.036 | 0.004 | -0.043 | 0.036 |
| | (0.048) | (0.055) | (0.079) | (0.048) |
| $N$ | 2,585,681 | 2,585,681 | 2,585,681 | 2,585,681 |
| No. of Clusters | 2,456 | 2,456 | 2,456 | 2,456 |
| | *Panel B: All Services: 2010 RVU Updates* | | | |
| Log RVU Change × Post | 0.244** | 0.315** | 0.203* | 0.242** |
| | (0.063) | (0.066) | (0.082) | (0.063) |
| $N$ | 2,386,575 | 2,386,575 | 2,386,575 | 2,386,575 |
| No. of Clusters | 2,051 | 2,051 | 2,051 | 2,051 |
| | *Panel C: All Services: 2011 RVU Updates* | | | |
| Log RVU Change × Post | -0.016 | -0.045 | 0.053 | -0.016 |
| | (0.068) | (0.075) | (0.066) | (0.067) |
| $N$ | 2,626,264 | 2,626,264 | 2,626,264 | 2,626,264 |
| No. of Clusters | 2,473 | 2,473 | 2,473 | 2,473 |
| Group-by-Code Effects | Yes | No | Yes | Yes |
| Code Effects | No | Yes | No | No |
| Cubic Time × RVU Change | No | No | Yes | No |
| Cubic Time × Post | No | No | No | Yes |
| Weighting | Dollars | Dollars | Dollars | Dollars |

Note: **, *, and + indicate statistical significance at the 0.01, 0.05, and 0.10 levels respectively. The table shows the results of OLS specifications of the forms described in section 3.2, except using data from out-of-network payments. Each column in each panel reports an estimate of $\hat{\beta}$ from equation (11). Panel A shows estimates using RBRVS updates and BCBS claims data for 2009, Panel B for 2010, and Panel C for 2011. Observations are at the claim-line level and are equally weighted. Standard errors are calculated allowing for arbitrary correlation among the errors associated with each HCPCS service code (including modifiers for the professional and technical components of diagnostic imaging services). Additional features of each specification are described within the table. The construction of all variables is further described in the main text. Sources: Authors' calculations using updates to Medicare's RBRVS as reported in the Federal Register and claims data from BCBS.

**Appendix Table D.5: Out-of-Network Services Priced According to Common Implied Conversion Factors**

| | Frequency Threshold: | | |
|---|---|---|---|
| | *5%* | *10%* | *20%* |
| *Panel A: 2009* | | | |
| Rounding for ICFs: | | | |
| *\$0.02* | 54% | 42% | 26% |
| *\$0.10* | 60% | 46% | 30% |
| *\$0.20* | 64% | 52% | 35% |
| *Panel B: 2010* | | | |
| Rounding for ICFs: | | | |
| *\$0.02* | 57% | 45% | 32% |
| *\$0.10* | 61% | 48% | 34% |
| *\$0.20* | 65% | 52% | 37% |
| *Panel C: 2011* | | | |
| Rounding for ICFs: | | | |
| *\$0.02* | 57% | 43% | 29% |
| *\$0.10* | 61% | 47% | 32% |
| *\$0.20* | 66% | 51% | 35% |

Note: Each cell shows the share of out-of-network services for which payments are associated with a common Implied Conversion Factor (cICF), as defined in the main text. The cells within each panel show how this share varies as we apply different thresholds for the frequency required to quality as a cICF. The column labeled "Rounding" indicates the rounding applied to each estimated ICF. An ICF is defined as "common" for the payments to a physician group if it accounts for at least the fraction of services associated with the specified Frequency Threshold. Source: Authors' calculations using claims data from BCBS.