Volume Responses to Medicare Reimbursement Changes*

Aileen M. Devlin

November 29, 2021

[Latest version available here]

Abstract

Medicare sets reimbursement rates defining provider payments and patient cost-sharing for the healthcare of the United States elderly and disabled, but the impact of these rates on healthcare volume, and hence spending and access is not well understood. Theoretically the impact of reimbursement on volume could be either negative or positive depending on whether provider substitution effects, provider income effects, or patient demand effects dominate. And the empirical literature has found mixed results in a variety of settings. I study this using relatively modern and exogenous variation from an Affordable Care Act provision that increased reimbursement for professional care in four states in 2011. I estimate two difference-in-difference analyses focusing on office-based care, which received the largest price increases. First, I compare affected states and a matched set of control counties. And second, within affected states, I compare more and less affected providers. These two analyses with very different identifying assumptions both yield a positive estimate for the effect of reimbursement on volume. The positive response is driven by increases in the number of providers and services rather than increases in service intensity. The positive relationship suggests that Medicare should be concerned about access when it decreases reimbursement rates.

*I am grateful to Amy Finkelstein, Jonathan Gruber, and James Poterba for their advice and support. I thank Angie Acquatella, Alden Cheng, Jeremy Majerovitz, Grace McCormack, Anna Russo, Parinitha Sastry, Martina Uccioli, Sean Wang, and especially Hector Blanco for helpful comments and conversations. This work was supported by the National Institute on Aging under Grant Number T32-AG000186 and the National Science Foundation Graduate Fellowship Program.
1 Introduction

Medicare’s reimbursement rates, a key policy lever affecting both spending and access to care, have a theoretically ambiguous impact on healthcare volume, and the empirical literature estimating the impact finds mixed results. Medicare historically assumed cost savings from reimbursement decreases would be offset by providers increasing volume, but more recently that assumption has been called into question with concerns that reimbursement reductions could limit access to care (Meara et al., 2017).

Theoretically, reimbursement would positively affect volume if provider substitution effects dominate. While the volume-reimbursement relationship could be negative if provider income effects caused the supply curve to bend backwards or if demand responses caused by coinsurance linking patient cost-sharing to reimbursement rates dominated. Hence, identifying the direction of the relationship is an empirical question. But the literature empirically investigating the relationship is mixed and the most clearly exogenous variation is from the late 1990s and earlier.

This paper estimates the impact of Medicare reimbursements on healthcare volume in a relatively modern setting. An Affordable Care Act (ACA) policy increased reimbursement rates in four states in 2011. I use these reimbursement changes, which were plausibly exogenous, to estimate two difference-in-difference (DID) analyses relying on very different identifying assumptions. First, I compare the affected states to matched control counties. And second, I compare providers within affected states who were more and less exposed to the reimbursement increase to each other. Both methods show that the reimbursement increase positively affected volume and suggest this was driven by extensive margin changes in the number of providers and services, not by service intensity.

I focus my analysis on office-based care paid through Medicare’s Physician Fee Schedule (PFS). The PFS is large, both as a share of Medicare and in absolute terms. Approximately 38 million of the United States elderly and disabled are covered by traditional fee-for-service Medicare and the PFS sets reimbursements for their office-based and other professional care (Freed et al., 2021). The PFS covers about 10% of total Medicare spending, and in 2018 had a total volume of $70.5 billion (MedPAC, 2020; Cubanski et al., 2019). The PFS updates reimbursement rates annually, changing rates for some procedures or geographic areas, which makes the impact of the changes continually of policy interest. I focus on office-based PFS care in particular because, office-based care is likely more discretionary than professional care provided in outpatient or inpatient hospitals, and the providers are more likely to have their incomes closely tied to the reimbursement for the

---

1The 38 million people covered by fee-for-service Medicare are the elderly and disabled who choose to enroll in the publicly run Parts A and B—they are 61% of total Medicare beneficiaries and exclude those who choose Part C or Medicare Advantage which is publically funded but privately provided insurance. Medicare Advantage does not administratively set rates, so these people are excluded from this paper (Freed et al., 2021).
Most of Medicare’s frequent reimbursement changes are endogenous responses to changing costs, but I leverage a plausibly exogenous ACA policy that raised reimbursements in four states. This policy introduced a binding floor for one of the geographic adjustments that raised rates in Montana, North and South Dakota, and Wyoming in 2011 [Federal Register, 2010]. The policy raised rates unequally across procedures generating additional variation within affected states. Individual providers specializing in the procedures with larger increases were more exposed to the policy change, and likewise areas with different distributions of procedures experienced difference average rate changes. The formula Medicare uses to set PFS reimbursements and the ACA policy change interacted such that within states affected by the legislated floor, counties average reimbursement increased from 5% to 12% while average reimbursement in counties in neighboring states increased only from 0.5% to 4.5%. And likewise providers in affected states faced average increases in reimbursement rates ranging from 3% to 15%.

I use this variation to run a DID analysis comparing affected counties to matched control counties. The DID depends on the identifying assumption of counterfactually parallel trends between treated and control groups, and the treated states are relative outliers compared to other states in the United States[^1]. I thus use a covariate matching algorithm to select counties from unaffected states that are most likely to satisfy the parallel trends assumption. This analysis finds that utilization in the affected states increased with reimbursement compared to the control states. The point-estimate for the elasticity is 2.6 for the number of services provided, 2.4 for total Relative Value Units (a measure of utilization incorporating both service intensity and count), and 1.1 for the number of physicians serving Medicare patients.

Since the area level DID depends crucially on the parallel trends assumption, I also run a provider level DID with a different parallel trends assumption. I use variation within treated states comparing providers who are more and less affected. The policy change increased rates more for procedures like x-rays and electrocardiograms where a greater share of the input costs come from practice expenses (materials, equipment, and nursing staff) as opposed to physician labor, while services like office visits and psychotherapy that are primarily derived from physician labor received smaller increases. Thus, variation in the types of services across providers generates variation

---

[^1]: Physicians can be employed by hospitals, group practices, or other associations and make a purely wage income, they can be sole proprietors whose income changes one-for-one with Medicare reimbursement for their service, or they can be basically anything in between in group practices with varied profit-sharing incentives [Gottlieb et al., 2020]. Since physicians integrated with hospitals change their coding and referral patterns to benefit the hospital, they clearly optimize jointly and hence may respond less strongly to incentives from professional payments or may not be the residual claimant on as large a share of the procedure level reimbursement [Baker et al., 2016; Sacarny, 2018].

[^2]: The policy applied to “Frontier States” which are defined as states where at least half of the counties have a population per square mile less than six. The affected states were thus more rural and sparsely populated than potential control states, but potential control states have many counties that look similar to the Frontier States.
in their average reimbursement change. Radiologists who generally receive large reimbursement increases are unlikely to be a good control group for psychiatrists facing generally lower reimbursement increases. But within a specialty, it seems plausible that family medicine doctors with different practice styles or patient cases mixes and hence different average reimbursement rate changes would have had parallel utilization trends in the absence of the ACA policy. I thus run a DID comparing the more and less affected providers within specialty. A particular strength of this analysis is that I can validate the parallel trends assumption by running an analogous DID in unaffected states.

The provider level DID corroborates the area level DID suggesting that providers respond positively to the increased reimbursement. At the provider level, the reimbursement increase causes a predominantly extensive margin response; for example, the approximately 3 percentage point spread in reimbursement increases between the 10th and 90th percentile family medicine physicians causes a 7 to 24 percentage point increase in the probability of continuing to provide office-based care depending on the time frame. Though this is to some extent offset by increase in facility care so there is little impact on the provision of any care at all. The estimates for intensive margin changes in service counts or Relative Value Units conditional on continuing to provide care are quite noisy, but they are suggestive of a positive response.

Considering the area and provider level results together, the effect of the ACA reimbursement increase in the affected states seems to have been to decrease provider exit and to increase the total number of services provided to Medicare patients. There is no evidence that the reimbursement increase changed service intensity; i.e., that individual procedures became more resource intensive. This contrasts with Clemens and Gottlieb (2014) who found that service intensity drove almost the entirety of the positive response they found.

The positive relationship suggests that the response is a provider driven movement along a supply curve (assuming healthcare is not a Giffen good). Medicare mechanically links patient out-of-pocket prices to provider reimbursement with a 20% coinsurance, so demand as well as supply is a function of Medicare reimbursement. But since Medicare administratively sets reimbursement there is no reason to assume the market is anywhere near equilibrium. If there were excess demand with reimbursement set below the market clearing rate, the positive response could cause a welcome decrease in rationing as the quantity supplied and demand come closer together. A positive supply response is not inconsistent with a reimbursement rate set above the market clearing rate and an excess supply disequilibrium because providers could take advantage of asymmetric information to induce demand above the quantity that a fully informed patient would desire. If the positive response were an induced demand response, that would have very different policy implications as Medicare would then being paying more inframarginally and simultaneously inducing inefficient

\footnote{As I discuss later, there are reasons to believe that patient demand may be high or inelastic since many patients have supplemental insurance to cover their cost-sharing.}
care. A third, more neutral explanation for the results is that reimbursement is not changing the volume of care, but simply affecting whether it is office or facility-based—this can be investigated by expanding my analysis to include facility-based care.

**Related Literature:** My paper contributes to the literature evaluating how health care utilization and provision responds to changes in reimbursement structure. This has been studied broadly—internationally, Norwegian general practitioners provide more visits when the payment increases (Brekke et al., 2017) and Japanese physicians distort their choice between branded and generic drugs to increase profits (Iizuka, 2012). In the United States private insurance incentives affect cesarean section rates (Johnson and Rehavi, 2016) and hospital choice (Ho and Pakes, 2014). Medicaid payments for primary care have been shown to increase appointment availability and utilization for Medicaid patients (Alexander and Schnell, 2019; Cabral et al., 2021; Polsky et al., 2015). The existence of Medicare itself increased the size of the healthcare sector (Finkelstein, 2007), and Medicare’s hospital payments affect how hospital care is provided (Dafny, 2005; Einav et al., 2018).

More specifically my paper adds to a large literature estimating the utilization response to Medicare reimbursement for outpatient care. The most closely related paper is Clemens and Gottlieb (2014), which studies how an exogenous Medicare policy that changed reimbursement at the geographic level in 1997 affected outpatient utilization. The papers estimating reimbursement elasticities within outpatient Medicare generally use either geographic or procedure level variation in reimbursement, and studies using both types of variation have estimates positive and negative elasticities.

Even the studies using the most clearly exogenous geographic level changes in reimbursement estimate both positive and negative elasticities. Clemens and Gottlieb (2014), Rice (1983), and Rice and McCall (1982) take advantage of geographic reimbursement rate changes that came from Medicare changing the size of the geographic areas at which it determines reimbursement. Clemens and Gottlieb (2014) estimates a positive elasticity, while Rice (1983) and Rice and McCall (1982) estimate negative responses. While the policies generating the variation they study were almost certainly exogenous, the reimbursement rate change was correlated with area characteristics—generally the more rural areas faced reimbursement increases while the urban areas faced decreases. Clemens and Gottlieb (2014) control flexibly for differential trends by urbanicity, but Rice (1983) and Rice and McCall (1982) may be biased by differential trends.

Other papers using more nuanced strategies to identify the reimbursement-utilization relationship without exogenous policy changes have estimated both positive and negative volume elastic-

---

5Slightly more tangentially related, providers who financially benefit from recommending more intensive care by being able to provide it themselves do so to a greater extent for MRIs and cardiac problems (Afendulis and Kessler, 2007; Baker, 2010).
ities. Hadley et al. (2009) and Brunt (2015) create a “generosity” measure by subtracting their best estimate of cost from Medicare’s payment rate and assume that time and area fixed effect will account for any unobserved variables that might cause bias, but Hadley et al. (2009) estimates a positive elasticity while Brunt (2015) estimates a negative one. In a more recent, and quite resourceful paper using only public data, Brunt and Hendrickson (2021b) match physician practices with similar characteristics across borders of Medicare payment regions that face different reimbursement rates and estimate a negative elasticity—though this depends on the assumption that no unobserved variables impacting utilization vary discreetly at locality borders, which are often states lines.\(^6\)

Papers estimating the volume response to reimbursement using procedure level variation tend to find positive estimates, with some notable exceptions. Usually procedure-level time series variation is endogenous, but it can be reasonably assumed that the exact timing of large, abrupt changes are exogenous so that short event studies around the changes identify the reimbursement response. Many studies find a positive relationship across a wide range of procedures and time periods including nerve conduction studies, cesarean sections, bladder cancer care, and androgen-deprivation therapy from 2004 to 2015 (Callaghan et al., 2016; Gruber et al., 1999; O’Neil et al., 2016; Shahinian et al., 2010). A few studies looking at similar variation have estimated negative effects (or found no evidence of a positive effect) for intensity modulated radiation therapy, coronary artery bypass grafting, and lung cancer chemotherapy (Jacobson et al., 2010; Howard and Hockenberry, 2021; Yip, 1998).

My paper contributes to the literature on responses to Medicare reimbursement by analyzing a relatively recent exogenous change. My finding of a positive reimbursement-utilization relationship strengthens the growing consensus across countries, insurance providers, and types of care that providers respond positively to reimbursements. Another contribution is my finding that extensive margin provider participation is a mechanism in the response. That provider financial incentives shape volumes is well established, but within Medicare the typical mechanism is via service intensity—Clemens and Gottlieb (2014) find no response for the number of services or providers, and their entire positive results is driven by service intensity\(^7\). Though the extensive margin response that I find is not surprising given Clemens et al. (2021) finding that higher reimbursement rates cause providers to engage in capacity building.

The extensive margin mechanism relates my paper to the literature on access to care for Medicare patients (and given my setting, to rural access in particular). Cross-sectional evidence from Brunt and Jensen (2014, 2013) and Gillis and Lee (1997) show that higher Medicare rates are associated with more providers accepting new Medicare patients and Medicare patients reporting better

\(^6\)A related paper, Brunt and Hendrickson (2021a) uses reimbursement variation at the provider level to estimate a negative elasticity. This is more similar to the geographic level studies than it is to the procedure level studies since the variation is in the total amount that the physician’s Medicare income changes rather than only a subset. Brunt (2015) also finds a response driven by intensity, but his results have the opposite sign.
access to care. And in rural areas, these concerns are amplified by already poorer access—
Chan et al. (2006) finds that Medicare beneficiaries generally have to travel further for their care, and
Johnston et al. (2019) finds that rural Medicare patients lack access to specialists and that this is
associated with preventable hospitalizations and mortality. This paper contributes to that literature
by suggesting that reimbursements have an inverse relationship with provider participation.

My paper also contributes to Medicare policy by providing more recent evidence in favor of a
positive relationship between reimbursement and volumes. For many years the Centers for Medi-
care and Medicaid Services (CMS) assumed that reductions in Medicare payment rates would be
offset by increased volume. As of 2012, Medicare assumed that 30% of savings from reduced reim-
bursements would be offset by increased volume, but a panel of academics and actuaries reviewing
the Medicare Trustees’ assumptions suggested that the studies and data underlying their assumption
were old (Bertko et al. 2012). The 2017 review of the Medicare Trustees’ assumptions expresses
particular concern that Medicare access may decline as Medicare rates are projected to fall relative
to private rates (Meara et al., 2017). And recently Medicare has been particularly concerned
about access in rural areas in particular, due to their lower utilization rates (MedPAC, 2021). My
result suggests that to the extent that the response is symmetrical for reimbursement increases and
decreases that cutting reimbursement raises serious access concerns as supply falls in response.

The rest of my paper proceeds as follows. Section 2 discusses a theoretical framework for
how reimbursement affects utilization under administered rates. Section 3 describes background
information on the Physician Fee Schedule and the ACA policy change. Section 4 describes the
geographic area DID empirical strategy, data, and results. Section 5 presents the provider level
DID strategy and results. And Section 6 concludes.

2 Theoretical Framework

Theoretically the relationship between utilization and reimbursement could be either negative or
positive. The administratively set prices mean that the market may not be in equilibrium. A reim-
bursement change could cause the realized disequilibrium to move along an upward sloping supply
curve causing a positive utilization-reimbursement relationship. Or the realized disequilibrium may
move along either a downward sloping demand curve or a backward bending supply curve causing
a negative utilization-reimbursement relationship.

Provider supply is clearly a function of reimbursement rates. And patient demand is also a func-
tion of provider reimbursement rates. Medicare mechanically links patient out-of-pocket prices to
provider reimbursements via a deductible and a 20% coinsurance rate, when provider reimburse-

---

8MedPAC (2021) finds that rural Medicare enrollees have less specialist visits in 2018, for frontier counties utiliza-
tion was much lower though that may be because enrollees are healthier.
ment changes, so does patient price. The RAND Health Insurance Experiment tells us that patients indeed respond to their cost-sharing (Manning et al., 1987). Though it is likely that within Medicare demand is high or inelastic because many patients have supplemental insurance like Medigap. Supplemental insurance often breaks the mechanical link between patient and provider prices by replacing the coinsurance with a copayment that does not vary with provider reimbursement. Most patient have some form of supplemental insurance—as of 2014, in affected states only 15% to 24% of Medicare enrollees had no supplemental insurance. And inattentive patients or the general difficulty understanding health care prices would likewise make patient demand high or inelastic. While demand responses may be ex-ante unlikely, to the extent patient behavior affects utilization, the relationship to reimbursement would be negative.

Medicare administratively sets reimbursement rates, which implies that the market is likely not in equilibrium. Prices (reimbursement rates in this setting) cannot freely adjust to bring supply and demand to the same volume. The market can, in principle, operate on the demand curve with excess supply or on the supply curve with excess demand. Figure 1 shows examples of disequilibrium outcomes. Panel 1a shows two possible disequilibrium states depending on which price level Medicare sets. At $P_1$ there is excess demand, so presumably the market would be constrained by the available supply and end up at the disequilibrium indicated by point 1 on the graph with providers somehow rationing care. Then when Medicare increased prices, the disequilibrium would move up along the supply curve. If Medicare set a high enough reimbursement rate, there could be excess supply in the market as shown in panel 1a by reimbursement rate $P_2$—in this case a price increase would cause a decrease in volume as the disequilibrium moved along the demand curve. In this case, there would a negative utilization-reimbursement relationship driven by demand.

Because there is asymmetric information in health care markets, when there is excess supply, instead of the market operating on the demand curve, providers may induce demand (Chandra et al., 2011; McGuire and Pauly, 1991). In health care, providers both provide care and serve as agents to help patients decide what care they need or want. Thus, providers may change the information they provide to patients to shift the demand curve—Chandra et al. (2011) defines provider induced demand as “provision of care that a fully informed patient would not choose for himself.” Panel 1b of Figure 1 shows that when there is excess supply, providers may create a inflated demand curve by inducing demand so that the market operates on the supply curve and the relationship between reimbursements and volume is again positive. Here when Medicare increases price, providers induce additional demand and move along their supply curve.

---

9Based off the 2014 Survey of Income and Program Participation data—supplemental coverage includes Medigap, Medicare Advantage, retiree plans, or private plans.

10Patient demand is not necessarily the correct demand curve to use to think about welfare analysis because moral hazard caused by patients not paying the full marginal costs of their care may make it higher than is socially optimal.

11Ji (2021) suggests there are at least some segments within Medicare where demand is binding.
Another complexity with the healthcare market that adds further theoretical ambiguity to the elasticity of volumes with respect to reimbursement is that provider supply curves may slope backwards. Generally, substitution effects are assumed to dominate income effects in labor supply, but several aspects of healthcare make it plausible that income effects could dominate. The training to become a doctor is costly and long, so entry in response to increased reimbursements may only increase supply in the long term, while income effects cause current providers to work less. Fixed costs are also high (as are transaction costs to exit from provision or sell equipment for certain procedures). Thus when reimbursement rates decrease, in the short run, providers may increase care to cover fixed costs or meet behavioral income targets until exit or other substitution forces decrease provision. One final reason income effects are particularly plausible in response to Medicare reimbursement changes is that other payers often base their payments on Medicare’s rates so even rates for non-Medicare patients could be affected magnifying the shock to income (Clemens and Gottlieb, 2016). In Figure 2, I show examples of Medicare reimbursement rates associated with a backward bending supply curve—in this case there could either be excess demand (panel 2a) or provider induced demand (panel 2b) allowing the market to move along the downward sloping supply curve. In either case, there is a negative relationship between reimbursement and volume that is driven by providers.

The slope of the supply curve depends on whether substitution or income effects dominate. If the reimbursement rate for a procedure decreased, but that procedure were a small share of a provider’s income, or the decrease were offset by reimbursement increases for other procedures, one would not expect a sizable income effect. Large income changes generating large income effects could occur if a provider focuses on one or several procedures with reimbursement changes, or because the reimbursement change is geographic and applies to all Medicare procedures. In thinking about income effects, the types of reimbursement changes that cause them is only half the puzzle—given a shock to income providers could change their provision of any procedure to patients with any type of insurance, it does not have to be the payer and procedures whose price change induced the shock. McGuire and Pauly (1991) point out that the profit margins rather than the reimbursement rate will have a large impact on where income effects change volume. To make up lost income from a reimbursement decrease, it only makes sense for providers to increase volume for procedures or payers with a positive margin, and there will be more incentive to increase volumes for care with higher margins. Thus, income effects are symptomatic of large margins.

In this paper, I estimate the aggregate elasticity for all office-based professional care. Not all PFS care is one homogeneous market, it aggregates many smaller markets presumably defined by some combination of geography, provider specialty, procedure type, and diagnosis. In each of these separate markets, the disequilibrium outcomes may be different. Thus, the aggregate responsiveness is some sort of weighted average of these across the many markets. It is not entirely clear
what exactly a market is defined by, but a key point is that in any given market the administered reimbursement rates mean that presumably only one of providers or patients are making decisions affecting the volume response.

How the Framework Relates to the Findings of the Prior Literature  
The framework I lay out above suggests that the market operating on an upward sloping supply curve does not imply induced demand, as the administered reimbursements mean there could be excess demand. The older literature—both theoretical and empirical—looking at volume responses to demand assumed that increases in volume were suggestive of provider induced demand. Rice (1983) who studied a 1976 reimbursement change, in fact, has “Physician-Induced Demand” in the title—but the market he studied was more likely in equilibrium since at that point Medicare’s reimbursement was based off the local customary and prevailing rate. But if the market may be out of equilibrium, it take more than a positive (or for that matter, any supply driven response) to conclude induced demand. My framework with the market in disequilibrium also suggests excess demand as a possibly simpler explanation for why supply responses dominate. The literature often assumes demand effects to be non-existent or small arguing that supplemental insurance makes patients inelastic (Brunt and Hendrickson, 2021b; Clemens and Gottlieb, 2014), but with excess demand it is possible that quantity demanded changes a lot but has no impact on quantity due to rationing.

The slope of the supply curve depends on whether substitution or income effects dominate. In the short run it is plausible that income effects may dominate, but in the long run substitution effects are generally assumed to dominate (McGuire and Pauly, 1991). Entry and exit over time would relieve pressure to cover fixed costs and reach target incomes. And to the extent there are adjustment costs to substitute to other types of care that would delay substitution effects. For example, Clemens and Gottlieb (2014) argue that investment in new capital equipment explains much of the positive response they estimate. But it also seems like income effects may change practice styles and standards of care so may not phase out with entry and exit over time.

The predictors of large income effects may explain why the studies based on individual procedures price changes are generally positive while many of those based on geographic variation are negative. The individual procedures are less likely to be a large enough share of a providers income to cause the income effect to dominate. And it is possible that the few negative relationships are in setting where the providers specialize highly, so income effects may be large for few procedures.

Geographic and procedure level variation in reimbursement rates allow the estimation of two subtly different elasticities. The key difference being what is held constant—for geographic vari-

---

12 For example, Johnson and Rehavi (2016) convincingly argues for induced demand by comparing treatment given to physicians and non-physicians. When physicians treat other physicians as patients, the information asymmetry is clearly lower, so there is less scope for the treating physicians to induce demand. They find that the positive reimbursement-volume relationship is weaker when the patients are themselves physicians.
ation, the relative reimbursement rates within Medicare is constant, while for procedure variation, all Medicare reimbursements for other procedures are constant. By holding relative reimbursement rates across procedures constant, geographic variation makes substitution forces between procedures less likely, though relative profit margins may vary so there may still be some substitution effects. The consequence is that a larger share of a provider’s revenue, all Medicare income, is affected, so income effects will be stronger. Conversely, procedure level variation holds all non-Medicare and other Medicare procedure reimbursements constant so a smaller share of income is affected and income effects are likely smaller. But by changing relative reimbursement rates, incentives to substitute between procedures are stronger. My setting is somewhere between the two extremes that have been studied, because my variation affects most Medicare procedures, but to different amounts.

3 Background: the Physician Fee Schedule and Reimbursement Rate Variation

The PFS covers professional care within Fee-for-Service Medicare and is a large subset of total Medicare spending. An ACA policy increased reimbursement in four states in 2011, but did so with differential impact across procedures and providers within those four states.

PFS spending accounts for 10% of total Medicare spending (MedPAC 2020). Medicare as a whole includes traditional Fee-for-Service health insurance, the privately provide Medicare Advantage health insurance, and Part D insurance for prescription drugs. Fee-for-Service Medicare covers institutional care through Part A and professional and other outpatient care through Part B. Part B is itself composed of many payment systems of which the PFS is a relatively large component; other examples of payment systems are the Outpatient Prospective Payment System and the Ambulatory Surgery Center Payment System.

The professional care covered by the PFS includes payments to physicians and other professionals. The types of services paid via the PFS range from office visits to surgical procedures to diagnostic and therapeutic services. These services can be performed in physician offices or in facilities like inpatient or outpatient hospitals, skilled nursing facility, or ambulatory surgical centers. In non-facilities (primarily offices) the PFS payment covers the full cost of care, while in facilities the PFS covers only the professional component and another Medicare program makes a separate payment to the facility.

The PFS sets a total reimbursement rate for each procedure, which is paid to the provider by the

---

13 Other professionals include physicians assistants, nurse practitioners, and physical therapists who can provide professional care, though the non-physicians only file a separate claim if they are providing care independently, most non-physician work is included in the payment to the physician.
patient’s cost-sharing (deductible and coinsurance) and Medicare’s payment. Patients have a 20% coinsurance rate and a deductible that applies to all Part B care—in 2011 the deductible was $162 and it increases slightly each year. The deductible and 20% coinsurance mechanically link patient and provider prices. Though patients may have supplemental insurance (for example Medigap) that covers this cost-sharing.

The PFS is a good setting for my study because it is large, it has frequent reimbursement variation, and it is paid at a particularly granular level. It is large enough and affects enough peoples’ health care that it is worth studying for its own sake. Its frequent reimbursement changes both make understanding the volume response an important policy question and generate variation to estimate the volume response. And finally, the PFS pays for care at the procedure level—these granular payments means that there is more scope for patients and providers to finely vary utilization than there is in other payment systems where reimbursements are bundled.

3.1 Physician Fee Schedule Reimbursement Rate Setting

The PFS pays for care at the procedure level and sets reimbursement for each procedure at the area, time level. It sets this reimbursement based on procedures’ input costs and adjusts for geographic variation in these costs.

The PFS pays separately for each individual service or procedure according to its Current Procedural Terminology (CPT) classification. With physician input, CMS sets relative value units (RVUs) which describe the input costs for each CPT code. This is done separately for three types of inputs: provider work, practice expenses (materials, employee costs, and office rental), and malpractice liability insurance. Thus, each procedure has three RVUs.

Each of a procedures three RVUs are separately adjusted for geographic variation in cost for that input by multiplying the RVUs by an area, input specific geographic practice cost index (GPCI). These GPCIs describe how the costs of provider work, practice expenses, and malpractice liability insurance vary across areas. The geographic area that CMS adjusts costs at are localities. There are 89 localities in the United States—some are states (eg, New Hampshire), some are metropolitan areas (eg, Boston Metropolitan Area), and some are the rest of states with metropolitan areas excluded (eg, Rest of Massachusetts). Medicare combines the RVUs and GPCIs with an annually set conversion factor that scales the RVU measure of relative cost into dollars.

A simplified version of the formula Medicare uses to combine the RVUs and GPCIs into procedure by area reimbursement rates is below. This omits malpractice liability insurance costs and a few other minor adjustments, which are not critical for understanding the reimbursement vacation my identification strategies use—see the appendix for a more complete discussion of how reimbursement rates are set.
Reimbursement_{pj} \approx ConversionFactor \times [RVU_p(\text{work}) \times GPCI_j(\text{work})
+ RVU_p(\text{practice expense}) \times GPCI_j(\text{practice expense})]

Where \(p\) indexes procedure and \(j\) indexes locality. The parameters change over time, so there is an implicit time subscript. The policy variation I exploit was an exogenous change in the practice expense GPCI in four states, so the important takeaway is that this only affects reimbursement in those four states and affects reimbursement more for procedures with larger practice expense RVUs.

**ACA Floor on Practice Expense Geographic Practice Cost Index:** The ACA mandated a permanent floor for practice expense (PE) GPCIs in Frontier States. Frontier States are defined as states where at least fifty percent of the counties have a population per square mile of less than six ([Federal Register, 2010](#)). The five Frontier States in 2011 were Montana, Nevada, North and South Dakota, and Wyoming, but the GPCI floor did not bind in Nevada. Hence the GPCI Floor increased reimbursement from 2011 on in Montana, North and South Dakota, and Wyoming. I call these four states the treated states.

Figure 3 shows the PE GPCI over time in the four treatment states and the surrounding states from which I select the control counties for my DID analysis—note that from 2011 on the treatment states have a GPCI of 1, which is much higher than their pre-2010 level. There was also a smaller increase in PE GPCI in the treatment and some of the control states in 2010 and 2011—this is a temporary ACA provision, which I treat as noise that just dilutes some of my variation. Figure 4 shows that the increase in PE GPCIs in the treated states translated into an increase in reimbursement per RVU, while there was a much smaller increase in control states over the same period.

The ACA included two other provisions intended to protect access to care in Frontier States. First, it established an area wage index floor for hospitals at 1.00, and second, it established an area wage adjustment factor for hospital outpatient department services at 1.00. ([111th Congress, 2010](#)) These are potential confounders. They do not affect office-based care directly, but to the extent that office-based care is a substitute or complement for hospital care, this could affect my results.

**Variation Induced by the 2011 PE GPCI Floor:** The PFS’s formula along with the PE GPCI floor creates differential reimbursement changes across procedures, areas, and providers.

---

(14) This was the result of a policy that only applied in 2010 and 2011 that limited the incorporation of below average costs into GPCIs—for localities with costs below average, it reduced the deviation from average by a constant percentage.
Since the PE GPCI multiplies only the PE RVU in the reimbursement formula the increase in reimbursement is larger for procedures with a larger share of RVUs from practice expenses. Comparing two procedures with very different PE shares exemplifies this distinction. Psychotherapy has a PE share of 0.23, while eye exams have a PE share of 0.75. For psychotherapy, a 0.15 increase in the PE GPCI would increase reimbursement by 4% since only 23% of the total RVUs are subject to the increased GPCI, while for eye exams the same 0.15 PE GPCI increase would increase reimbursement by almost 13% since 75% of the RVUs are subjected to the increased GPCI.

Table 1 shows the psychotherapy and eye exam example in panels A and B. Panel A lists the GPCIs for Montana in 2009 and 2011—there was a 0.153 increase in the PE GPCI. Panel B shows the RVUs for psychotherapy and eye exams in 2009 and 2011 in columns 1 through 4—three key points about the RVUs are that (1) psychotherapy is more intensive since the sum of its RVUs is higher, (2) psychotherapy has a lower PE share, and (3) eye exams had their PE RVUs updated between 2009 and 2011. Column 5 shows the GPCI-adjusted sum of RVUs in 2009—this is proportional to reimbursement, which is simply total adjusted RVUs multiplied by a conversion factor. Column 6 shows the 2011 GPCI-adjusted total RVUs, for psychotherapy this is just the same RVUs multiplied by the new 2011 GPCIs. But for eye exams, the PE RVUs also changed between 2009 and 2011 (compare columns (2) and (4))—thus the difference in GPCI-adjusted total RVUs incorporates both the change in GPCIs (which I am assuming was exogenous) and the endogenous change in PE RVUS. Column (7) shows the GPCI-adjusted total RVUs using the new 2011 GPCI rates, but the old 2009 RVUs, the difference between this and the 2009 GPCI-adjusted total RVUs is only due to the exogenous change in GPCI rates. Columns 8 and 9 show the percent changes in GPCI-adjusted total RVUs (and reimbursement) for the actual change and the change induced by the exogenous GPCI floor holding RVUs fixed.

The variation in reimbursement changes across procedures with different PE shares translates into variation in reimbursement changes across areas which provide different mixes of procedures. Areas with more high PE share procedures received larger average reimbursement changes than those where the average PE share was lower. Panel C of table 1 provides an example of this type of variation using two fictitious counties. In both counties the Medicare patients get 100 eye exams (column 2), but in the sad county the patients receive 100 psychotherapy visits while in the happy county they only require 10 (column 1). Multiplying these volumes by the GPCI-adjusted total RVUs yields the county aggregate adjusted RVUs (which is proportional to total Medicare revenue)—these numbers are shown in columns 3 and 4. Column 5 then shows that the percent change in aggregate adjusted RVUs is larger in the happy county where a great share of the revenue

\[ ^{15} \text{This example is simplified to only include work and PE inputs (i.e., it ignores liability insurance inputs—these are small and would increase complexity with no value added). In the empirical application insurance inputs are treated analogously to work input. Also note that there are multiple types of psychotherapy and eye exams, the listed values are for CPT codes 90806 and 92250 respectively.} \]
comes from the high PE eye exams. This example shows how the GPCI change caused differential reimbursement changes across counties.

A parallel logic applies to individual providers who provide different mixes of procedures. Obviously, there are probably few providers who provide both eye exams and psychotherapy, but there are family medicine physicians who provide more or fewer high PE share tests like EKGs and x-rays relative to the number of relatively low PE office visits they perform. I exploit this across provider variation in my second identification strategy.

4 Measuring the Aggregate Volume Elasticity with Geographic Variation

I estimate the elasticity with which volume responds to reimbursement using a difference-in-difference (DID) model that compares areas where the 2011 Frontier State PE GPCI Floor increased reimbursements to a matched set of counties. I use Medicare claims data and CMS Physician Fee-for-Service RVU files to implement this.

4.1 Difference-in-Difference Regression

My baseline specification is a standard ordinary least squares DID regression:

\[ Y_{ct} = \beta \mathbb{1}(t \geq 2011) \Delta \text{Reimbursement}_c + \gamma_t + \gamma_c + \varepsilon_{ct} \]  

Where \( c \) represents counties and \( t \), year. \( \Delta \text{Reimbursement}_c \) is defined at the county level as the percent change in total reimbursements induced by the PE GPCI floor in Frontier States—as I discuss in detail in the data section, I generate this to isolate the exogenous change from the GPCI floor and exclude endogenous changes from RVU adjustment and changes in the mix of procedures provided. \( \gamma_t \) is a vector of year fixed effects and \( \gamma_c \) is a vector of county fixed effects. Since \( \Delta \text{Reimbursement}_c \) is a percent change in reimbursement, when the outcome is measured in logs, \( \beta \) is directly interpretable as the elasticity of the outcome with respect to reimbursement.

The DID depends on the parallel trends assumption that the treated and control areas would have been on parallel trends in the absence of the reimbursement change. This can to some extent be evaluated by checking for pre-trends in the event study version. To make the parallel trends assumption plausible, I select control counties from nearby states using a covariate matching algorithm that I discuss below.
Event Study to Look at Timing: To see the timing of the volume response to the reimbursement change and evaluate the pretrends, I run an event study version of the DID:

\[ Y_{ct} = \sum_{\tau} \beta_\tau \mathbb{1}(\tau = t) \Delta \text{Reimbursement}_c + \gamma_t + \gamma_c + \epsilon_{ct} \tag{2} \]

Where variables and indices are defined as in equation [1]. The series of \( \beta_\tau \) show the impact of the reimbursement change over time. Given the identification assumption, there should be no pre-trends, so the estimates of \( \beta_\tau \) should be small and insignificant before 2010 (a pre-trend in 2010 would not seem problematic because of the temporary smaller reimbursement increase in 2010). And their evolution after 2011, shows the timing of the effect.

Standard Errors I cluster my standard errors at the county level. This accounts for any serial correlation in county outcomes that is not absorbed by the county fixed effects. The GPCI floor was assigned at the state level, so clustering at the state level would potentially be more conservative (though it would run into the problem of small numbers of clusters). But the impact of the floor varies across counties based on the types of care provided in those counties; for example, counties that do more imaging, tests, and other high PE procedure have larger reimbursement increases, and this generates substantial variation within states.

Outcomes The general term “utilization” describes both the service count of procedures performed and the intensity of the procedures. For example, doubling the number of office visits would increase utilization, but holding the number of office visits constant and making all short office visits into long office visits would also be an increase in utilization. Measuring the impact on service counts simply entails running the main regression on log service counts. To incorporate intensity changes as well, I follow the standard practice of using log RVUs as an outcome. RVUs are designed by Medicare to measure the relative input costs across procedures, so to the extent Medicare succeeds in this endeavor, RVUs are an apples-to-apples comparison across procedures. This solves the problem that while short and long office visits have an obvious intensity ranking, for most procedures, even the ordinal ranking is not clear (for example, for a twisted ankle, is an x-ray and a physical therapy visit more intensive?)\(^{16}\)

Another outcome I explore is the number of providers providing any office-based care to Medicare patients. This examines the extensive participation margin. Though my focus on office-based

---

\(^{16}\)Increasing the intensity of any given procedure could be “upcoding” where the same procedure occurs but is coded differently so that reimbursement is higher, or it could be an actual change in the care provided. Daity (2005) suggests that in the inpatient setting can be purely a relabeling without affecting care provided. I remain agnostic to whether the increased intensity is real or nominal.
care leaves open the possibility that provider exits could be movements to facility-based care.

4.2 Selecting Control Areas: Covariate Matching

Since the DID specification depends on a parallel trends assumption between treated and control area, I use covariate matching to select a set of control counties likely to be similar to the treatment counties.

As shown in table 2, the treated states are far more rural and sparsely populated than the United States as a whole. The treated states lack large metropolitan areas—the one Frontier State (Nevada) with a large metropolitan area (Las Vegas) had a sufficiently high PE GPCI such that the floor did not bind. The fact that the treated states are such outliers makes it hard to find potentially comparable states or even to generate comparable “synthetic” states by aggregating entire state a la Abadie (2021) and Abadie et al. (2010). Thus, I instead look within states for counties that match treated counties so I can, for example, include rural northern Nevada without having to include Las Vegas.

The pool of possible control counties I consider is all the counties in the surrounding states as well as Nevada. To the extent that health care customs, economic conditions, and other factors influencing utilization choices change slowly with geography the surrounding states are likely to have counties on similar trends to the treatment counties. And Nevada is also considered a Frontier State, so many of its counties are likely similar to the counties in the treatment states.

I generate a propensity score for each county in treated and control states by regressing a treatment indicator on covariates and for each treated county select five matched control counties. The covariates used for matching are the 2010 poverty rate, SNAP recipiency rate, share of population living on farms, number of hospital beds per capita, and psychiatrists per capita. These predictors were selected by regressing the treatment indicator on a standardized index of many county characteristics and choosing the five predictors that were the most significant.

This method closely parallels the method used by Diamond et al. (2020).

For each treatment county I include five matched control counties. The five matched control counties are given a total weight equal to the weight of the treatment county and the weight is split among the control counties proportionately to their FFS population. Control counties can match multiple treatment counties, in which case I sum the weight they get from each treatment county. In

---

17 The full set of variables I initially considered were: 2010 population density; poverty rate, SNAP rate; share living on farms; housing density; rural population share; number of medical generalists, specialists, psychiatrists, and hospital beds per capita; and 2007-2010 trend in facility and office visits. I did not use variables that are related to absolute instead of per capita size since these tended separate treatment and control counties too well (intuition along the lines that an indicator for treatment would be a useless characteristic to base the propensity score on): eg FFS enrollees, population, number of facility or office visits. I also considered matching of the county level counterfactual GPCIs (as opposed to the real locality level ones), which is available in the 2011 County Data File for the 6th GPCI Update, but they do not vary much within states leading me to believe that they are based on many state level approximations (at least for the relevant regions).
all regressions, I weight the control counties as described above, and I weight treatment counties by their FFS enrollment. This weighting method has two benefits. First, matching each treated county to several control counties avoids placing too much weight on very small counties in control states which adds noise to my regression. And second, weighting control counties by the enrollment in their matched treatment county avoids placing too much weight on potentially large control counties that match only very small control counties (for example, the Twin Cities in Minnesota receive some weight, but not nearly as much weight as would be proportionate to their population).

Figure 5 shows that the counties in the treatment and control states have an overlapping distribution of propensity scores. As expected the treatment counties generally have higher propensity scores and there is a left tail of control counties that do not match any treatment counties. Figure 6 shows the weights placed on the treatment and control counties in the analysis—for treatment counties this is the FFS population, and for control counties this is the sum of the FFS population on the treatment counties they match. Note that the matching algorithm does not place too much weight on the big cities in the control states: Las Vegas has zero weight, while the counties containing Minneapolis, Saint Paul, Denver, and Salt Lake City respectively have weight 441, 321, 939, and 1611. Salt Lake City’s relatively high weight stems from the fact that it matches Missoula’s county, which has over 2000 FFS enrollees.

4.3 Data

For my analysis, I need data on the volume and types of procedures provided and the reimbursement rates. I use the Medicare Claims data to measure utilization and actual paid reimbursement rates and the Physician Fee-for-Service RVU files from CMS to measure statutory reimbursement. I also use supplemental data on area characteristics match control counties likely to fulfill the parallel trends assumption.

Restricting to Office-Based Care: I restrict to office-based care in my analysis. Office-based care generally faced larger reimbursement increases than facility based care this is because the practice expenses of operating a office are included in the reimbursements. For example, CPT “99213” the most common established patient office visits is 0.44 practice expenses in a private physician’s office, but 0.23 in a facility (so a 10% increase in PE GPCI would increase office reimbursement by 4.4% but facility reimbursement only by 2.3%). Office-based care is more likely to be discretionary, since professional payments in facilities include payments to emergency room doctors treating patients and the physicians caring for hospital inpatients. Providers may have more room to self-refer and induce volume if they so desire in office-based settings. And lastly, office-based care is more likely to be provided by independent providers or those in smaller group practices where the
provider’s income is more closely linked to the Medicare reimbursement their procedures accrue.

A down-side of restricting to office-based care is that it means that this analysis will not be able to measure any potential changes that the reimbursement rates changes may have had on where care is performed.

Measuring Reimbursement: I calculate the statutory reimbursement rates from parameters in the CMS Physician Fee-for-Service RVU files. These contain RVUs, GCPIs, and the annual conversion factor necessary to calculate the statutory reimbursements for each procedure in each locality each year. I select the most revised January publication each year, and run the inputs through the formula listed in annual documentation accounting for any special adjustments that apply in any given year to calculate procedure, locality, year level rates. Additionally, some modifier codes affect the payment, the ones that occur most often are modifier “26” and “TC” for the professional and technical component only, I account for these in my analysis. And some procedure’s have a payment cap at the level they would be paid were they paid via the Outpatient Prospective Payment System, which I account for.

These files also list the counties in each locality. For the states in my setting localities and states map one-to-one.

Some procedures can be done in either a facility (outpatient hospital, emergency room, or ambulatory surgery center) or non-facility setting (physician’s office), and these have different sets of RVUs that apply in the respective settings. Medicare pays different statutory rates for the same procedures depending on whether they are performed in facilities or non-facilities because they also pay a separate facility fee for care performed in facilities. I only consider the non-facility fees because I focus on office-based care in this paper. In the RVU file Medicare notes procedures that are only performed in facilities (and does not provide non-facility RVUs for these procedures)—I drop any procedure that Medicare does not provide a reimbursement rate for in non-facilities.

There are a few payment adjustmen that I do not account for. Probably the most important element that I ignore for simplicity is the reduction for multiple units of the same service to one patient on any given day. I also ignore the payment reduction for procedures performed by nurse practitioners and physicians assistants instead of physicians. And finally I ignore the payment increase for procedures performed in Health Provider Shortage Areas.

In 2010 I use the July publication because it includes a change to the January rates that applied retroactively, so the January rates were never the actual rates paid. Also, note that in some years there are slight adjustments to the formula—some year multiply some types of RVUs by budget neutrality adjusters (for example in 2008, work RVUs were paid at only 88.06 there normal rate), sometimes there are special adjustments for mental health care, and some procedures have their rates capped at a level based on the Outpatient Prospective Payment amount—I checked the annual formula adjustment and applied these.
Measuring Volume: To measure utilization paid through the PFS, I use Part B claims data for a 20% sample of Medicare beneficiaries. The PFS claims are a subset of the claims in the Carrier file—I assume that all claims in the Carrier file for procedures (defined by Current Procedure Terminology (CPT) codes) that the PFS covers were paid through the PFS.\footnote{I include any CPT codes that has a PFS status code of “A”, or “T” in the CMS Physician Fee-for-Service RVU files—these identify active codes paid separately under the PFS, restricted coverage codes that are carrier priced, and injections that are only paid if there are no other PFS services billed by the same provider that day.} I drop claims to people not enrolled in FFS Medicare that month because they are likely to be erroneous. I drop claims that were denied or assigned zero payment, assuming that these are duplicate or other erroneous claims. I restrict to office-based care by including only claims with place of service code “11”. I exclude procedures that CMS declares are never or rarely performed in non-facilities and does not provide a reimbursement rate for in non-facilities—these are ones with a “NA” in the PE RVU indicator field for that pos type. And I similarly drop claims where the listed provider zip code is not one of the counties included in the locality that the claim is listed as having being paid under (if these claims have two non-compatible locations listed I can’t be sure where they occurred).

For each of the claims I record the procedure type (CPT code and payment relevant modifiers), the number of units of service provided, the actual reimbursement made, and the provider zip code. I crosswalk the provider zip code to the provider county in order to run the analysis at the county level since reimbursement is defined based on the provider’s location. I also crosswalk on the total RVUs to use as a measure of service intensity.

Defining County Level $\Delta$Reimbursement: The RVU files provide the locality, year, procedure level reimbursement rates. These change over time as RVUs, GPCIs, and other adjustment change. The frontier state PE GPCI floor plausibly provides an exogenous change in some reimbursement rates. To isolate the exogenous change in reimbursement rates from the GPCI floor, I calculate the counterfactual change in reimbursement rates that would have occurred had only the GPCIs changed, but the RVUs and procedure mix had remained constant. This follows the logic described in the background section and demonstrated in table[1] with psychotherapy and eye exams—I extend that logic to all procedures and calculate the change induced only by the exogenous GPCI changes.

I define counterfactual reimbursement rates for each procedure based on 2009 RVUs and 2011 GPCIs. Then I apply the true 2009 rates and the counterfactual rates to the same set of pre-period claims, aggregate the total amounts to the county level, and take the percentage difference between the total based on the two rates as the county level percent change in reimbursement rate. For the pre-period claims, I use all available claims from 2008 and 2009 in order to have a less noisy measure since some of the counties are small and I only have a 20% sample of claims. I do not use 2010 because there was a small increase in GPCIs in the treated and some control states that may have impacted quantities. This definition of the reimbursement change only includes variation
from the GPCI floor so is as exogenous as the policy. It also takes advantage of variation within states in the mix of procedures typically offered in a county, because the change in GPCI caused a reimbursement increase that varied from around 3% to around 15% depending on the procedure. Thus, within treatment states there is substantial variation in the average reimbursement rate change. Figure 7 shows that within control states reimbursement increases varied from around 0.5% to 4.5% while in treated states, reimbursement rate changes varied from 5% to 12%.

**Supplemental Data:** I get county characteristics for my propensity score matching algorithm from the Agency for Healthcare Research and Quality. This has county level data on county demographics and health care provider density.

Finally, I use the Master Beneficiary Summary File to measure fee-for-service enrollment in each county.

### 4.4 Results on the Aggregate Elasticity

My analysis finds that the aggregate volume elasticity for office-based care is positive. Over the first few years after the price increase (where the parallel trends assumption is most likely to hold) the elasticity point estimates range from 2.6 for service count, to 2.4 for RVUs, and 1.1 for number of providers.

The main event study results from equation 2 are shown in figure 8. There are no significant pre-trends and the point-estimates before 2011 are small. Immediately, in 2011 when the reimbursement increases, the service count, the number of RVUs, and the number of providers rise. The estimates can be directly interpreted as elasticities since the outcomes are in logs and the regressor is a percent change. The pre/post DID version of this from equation 1 is shown in table 3 in column 1—the point estimate for the elasticity using years 2008 to 2013 is around 2.6 for service counts, 2.4 for RVUs, and 1.1 for the number of providers. The RVU elasticity is slightly higher than the elasticity of 1.5 that [Clemens and Gottlieb (2014)](#) estimate, while [Clemens and Gottlieb (2014)](#) actually estimate no effect for service counts or the number of providers.

It is interesting that the effect is driven by service count changes (and providers). The fact that RVUs move in parallel to service counts means that the RVUs (intensity per service) stays constant. Papers looking at the volume response to outpatient reimbursement typically explore both the service count and intensity margins, but the typical finding is that the intensity margin drives the results. [Clemens and Gottlieb (2014)](#) for example evaluates the impact on both service counts and RVUS and find a positive RVU response but no impact on service counts, which implies the result is driven by making procedures more intensive. [Brunt (2015)](#) looks directly at how reimbursement

---

20 The underlying data for these DID regressions is shown in binscatters in the appendix (figure A.1) — there is no discernible trend break in treated areas at 2011.
variation affects the intensity level at which office visits and other procedures with obvious rankings and similarly finds that the response is driven by changes in coding intensity.\textsuperscript{21}

4.5 Robustness:

To probe the robustness of the positive result, I try a few variations on the sample of counties included and the weights for the regression. The variations I try are all ex-ante less powerful than my baseline specification or have an ex-ante more poorly matched control group. The variations all yield very noisy estimates that are not suggestive of a positive response.

I explore two alternative samples that it is reasonable to believe would be on parallel trends. First, I try using only the counties along the border between the treated and control states—this is less powerful than my main specification because it only uses a fraction of the data. To the extent that factors influencing healthcare change smoothly with distance, these counties would be on parallel trends. The critique that healthcare provision changes discretely with regulation at state borders, would only be problematic here if there were another change in state regulations at 2011 since the DID uses the across time change. This yields a very noisy estimate not suggestive of a positive response, as shown in column 2 of table \textsuperscript{3}. Second, I try to take advantage of the variation in reimbursement change among only the treated counties (as shown in figure \textsuperscript{7} this ranges from 4.5\% to over 12\%). Running the baseline specification only among treated states is not powerful enough since it throws out over half the data. Results shown in column 3 of table \textsuperscript{3}. The event study versions of these are all shown in the appendix in figure \textsuperscript{A.2}.

Next, I explore sensitivity to different weighting of the matched counties—I include the same matched set of control counties and weight them by their FFS enrollment (column 4 of table \textsuperscript{3}). This is ex-ante worse than the baseline specification because it places large weights on large control counties that match only small treated counties. In column 5 of table \textsuperscript{3} I try weighting the control counties by the minimum of their matched weight and their FFS enrollment—there are a few small counties that have a matched weight greater than their own enrollment, so it is plausible that they add noise, but rather they seem like particularly good controls for the treated counties since their down-weighting causes pre-trends to appear. Lastly, I try simply using all counties in the states I considered as donors for the control counties (column 6 of table \textsuperscript{3}). This is ex-ante worse than the baseline specification because Denver and the Twin Cities seem like bad controls for the treated states.

Finally, I explore how clustering the standard errors at the state level rather than the county...

\textsuperscript{21}Office visits for established patient can be billed with 5 different CPT codes (99211 to 99215) where the lower codes are for short and simple office visits and the higher codes are for longer and complex visits. Thus, for office visits, the next more intensive version is clearly the longer more complex version, while many procedures do not have such obvious next more intensive procedures.
level affects the interpretation. As shown in brackets for the baseline specification, this increases the standard errors. The p-values increase to 0.07 for log service count and 0.09 for log RVUs. But these standard errors are subject to the complaint that the number of clusters (11 states total, only 4 treated) is small.

5 Provider Level Analysis

To address concerns about the parallel trends assumption underlying the area DID, I run a complementary provider level analysis that depends on a very different identifying assumption. I run a provider level DID that compares more and less affected providers only within treated states—I determine the policy change’s impact on each provider based on the mix of procedures they provided in the pre-period similarly to how I determine the county level reimbursement changes. I then compare providers within treatment states who were more or less affected by the policy. As a falsification to this analysis, I compare providers in control states who did not actually face a reimbursement rate change and thus should not react to one. This provider level DID corroborates the positive response the area analysis finds. It also suggests that the response is driven not by intensive margin increases in provider volumes, but by extensive margin provider participation.

5.1 Family Medicine Case Study

I introduce the provider level DID analysis with a case study of family medicine physicians. Family medicine (FM) physicians are the most common specialty in the treated states—they make up 24% of the physicians appearing in treated state claims and account for 11% of total office-based Medicare spending (including non-MD spending) so they are an important subset to consider (Table 4 lists the specialties and their characteristics).

The baseline specification is a DID regression comparing family medicine physicians who received smaller and larger changes in reimbursement:

\[ Y_{it} = \sum_{\tau} \beta_{\tau} \mathbb{1}(\tau = t) \Delta \text{Reimbursement}_i + \gamma_t + \gamma_i + \varepsilon_{it} \]  

(3)

Where \( i \) represents providers, and \( t, year \). \( \Delta \text{Reimbursement}_i \) is the provider level reimbursement change, \( \gamma_i \) are provider fixed effect, and \( \gamma_t \) are year fixed effects. I cluster my standard errors at the provider level. And I weight the providers by their total spending in 2008 and 2009—I use both 2008 and 2009 to reduced noise caused by the fact that I only have data for 20% of Medicare patients so see only a fraction of each provider’s claims.
**Provider Level Reimbursement Changes**  To define the impact of the exogenous PE GPCI floor on each provider (and exclude endogenous variation from RVU changes or endogenous responses to the policy), I follow a procedure exactly analogous to that I use for counties for the area DID above. For each provider (defined by National Provider Identifiers), I take the bundle of procedures they provided in 2008 and 2009 and look at how the reimbursement they would have received for these procedures changes between 2009 and 2011. Taking a fixed bundle of claims from the pre-period avoids including endogenous responses in provision, and using two years of claims reduces noise for lower volume providers. For 2011 reimbursement, I use the counterfactual rates based on 2009 RVUs and 2011 GPCIs to avoid including reimbursement changes caused by endogenous RVU changes. I assign each provider a reimbursement rate change equal to the percentage difference between the total reimbursement for the fixed pre-period bundle of claims using 2009 and counterfactual 2011 rates.

Panel (a) of Figure 9 shows that there are a few percentage points of variation in the rate changes that FM physicians in treated states face. The 10th percentile reimbursement increase was 7% and the 90th percentile was 9.7%. This variation not huge, but it is reassuring that FM providers are a relatively homogeneous group. The identifying assumption is that FM physicians with different reimbursement rate changes would have had parallel trends, so it is important that there are not wildly different type of FM physicians. The variation is driven by the fact that some FM physicians do mostly low PE office visits, while other FM physicians take more x-rays and electrocardiograms and other high PE share tests and procedures.

**Identification Assumption and Falsification**  The identification assumption is that FM physicians with higher and lower reimbursement rate changes would have been on parallel trends had they been treated identically by the policy change. This is subtly different than the standard DID identification assumption that they would have been on parallel trends in the absence of the policy change because the “control group” in my analysis is not untreated by the policy change, but rather is treated to a lessor degree.

The easiest model of physician behavior that satisfies the identifying assumption is that physicians have a homogeneous response to each unit of reimbursement rate change, then the more treated physicians would have looked just like the less treated physicians if they had been treated to a lessor degree. More flexibly, one could assume that the treatment effect of each unit of reimbursement rate increase is not the same, but that all providers reacted identically to the first 7% increase so that had the physicians with larger increases only faced a 7% increase, they would have behaved similarly to the physicians who only received that increase.

This is basically assuming something along the lines of that the treatment effect is not correlated with the assignment of treatment level.
A possible violation of my identifying assumption that exemplifies how it is different from the standard DID identifying assumption when the “control group” is untreated would be if treatment effects and treatment levels were correlated. For example, physicians with lower PE shares (and lower reimbursement rate changes) could be more elastic to reimbursement rate changes than physicians with higher PE shares (and higher reimbursement rate changes). This would mean that in the absence of any changes the two groups would be on parallel trends. But if reimbursement rates went up the same amount for both groups, the lower PE share providers would respond more strongly, and if reimbursement rates went up more for higher PE share physicians (as they did) the different response could be due to both a different elasticity and a different reimbursement rate change., which could in principle even yield a spurious negative elasticity.

To some extent the parallel trends assumption can be tested by running a falsification analysis on providers in areas unaffected by the GPCI floor. For providers in the matched control counties I used for the area DID, I define the counterfactual reimbursement rate changes they would have faced if they had practiced in the treated states—these counterfactual reimbursement rate changes are shown in panel (a) of figure 9 and are quite similar to the types of changes the treated physicians faced suggesting that the groups are comparable. One concern is that there were smaller reimbursement increases in the control states but these were not correlated with the PE share, so the actual and counterfactual reimbursement rate changes were not correlated—see figure A.5 in the appendix.

I then run exactly the same DID specification on these control county physicians in the unaffected states as a falsification. If the providers with different PE shares had different trends in control states in the absence of any policy change, it would not be reassuring that the providers in the treated states would have had parallel trends. This falsification check only tests that the providers would have had parallel trends in the absence of the policy change. It does not test the additional more subtle assumption that the treatment effects for the lower and higher PE share providers were similar and thus that the higher PE share physicians would have had parallel trends to the lower PE physicians if they had had similar reimbursement rate changes.

**Sample Selection** I select family medicine providers based on HCFA specialty codes, and include codes “1”, “8”, and “38.”

I drop outlier providers. Very small providers will have noisy estimates of the reimbursement rate change. Very large providers are unusual in some respect and would have a large impact on the results because I weight by provider revenue. I drop physicians whose Medicare revenue in 2008 and 2009 is below the 5th percentile or above the 95th percentile of all FM physicians in treated states. Dropping 5% of outliers is a rather large amount, but the physicians at the bottom of the distribution are very low volume indeed, and I check for robustness to dropping only the top 1% of outliers. For the falsification regressions I drop providers whose total 2008 and 2009 Medicare
revenue was outside the middle 90% of physicians in treated states to get a more comparable sample.

I also restrict to physicians who were primarily office-based by including only those who received at least 60% of their revenue from office-base care. Though I also test for robustness to including all physicians.

**Family Medicine Results** The main results for the FM physician DID are that the reimbursement increase causes an extensive margin participation response so that providers with larger increases are more likely to continue to provider office-based care and that the increase causes intensive margin response for those who continue providing care (though that result is noisier). The extensive margin response is somewhat offset by decreasing in facility based care (these results are quite noisy), but overall, there is a relatively tight null result on the provision of Medicare-paid care in any place of service.

The main event study results for the FM physician DID from equation \[3\] are shown in figure \[9\] in panels (b) through (f). The blue squares are the DID estimates in the treated states, while the red squares are the falsification DID estimates from the untreated states. The results are also shown in Table 5, which has the pre-post DID results.

Panel (b) has the extensive margin effect on provision of any office-based care. Pre-trends are minimal (the precise 0 pretrend for the extensive margin in 2008 is mechanical because I include only providers who provided care in both 2008 and 2009) and the falsification test provides no evidence that the parallel trends assumption is suspect. Immediately following the reimbursement increase the probability of a provider having non-zero Medicare claims goes up for the providers who received larger reimbursement increases. The coefficients can be interpreted as the percentage point change in probability of nonzero claims if reimbursement increased by 100%. The DID coefficient for a the first two years together is 2.7 (table 5, first panel, column 1). Since this is estimated by comparing more and less affected providers the way to scale this coefficient is based on the differential change between more and less affect providers. Moving a provider from the 10th percentile increase to the 90th percentile increase raises they reimbursement increase from 7% to 9.8%. This 2.8 percentage point relative increase in reimbursement translates into a 0.07 increase in the probability of continuing to provide office-based care. For the subsequent 3 years, 3 to 5 years after the reimbursement increase, the impact is larger, the point estimate is 8.5, which means that moving a provider from the 10th to the 90th percentile reimbursement rate change would increase his probability of continuing to provide care by 0.24. This is a rather large impact, but it is driven by some providers with the lowest quartile increases switching to facilities—in 2013, the probability the lowest quartile of physicians still provided office-based care fell from above 80% to less than 60%, while the probability of providing facility-based care increased from almost 65% to almost 75% (see figures A.3 and A.4 in the appendix).
I explore the possibility that the increase in office-based provision is offset by decreased facility-provision by including indicators for facility-based and any claims as outcomes (shown in panels (c) and (d) of figure 3). The estimates for any facility-based care area are noisy, but the point estimates are negative suggesting there is some offsetting. Interestingly, when I weight each provider equally, rather than by their Medicare revenue, the extensive margin result for any facility-based care becomes significant (see figure A.6 in the appendix), which means that smaller volume providers who received relatively small reimbursement increases are shifting from office to facility-based care. Looking at Medicare participation in any place of service (panel (d)) shows that there is a relatively precise null effect on any participation. Again the falsification tests do not provide any evidence for failed parallel trends assumptions.

Finally I explore how physicians change their intensive margin quantity conditional on continuing to provide office-based care to Medicare patients (panels (e) and (f)). The estimates are noisy and the point estimates are quite large but it is suggestive that in the long run there is a positive response of volume to reimbursement. On the intensive margin, however, the falsification test fails—in control areas the remaining providers who had higher PE shares decreased their volumes relative to those with lower PE share.

**Family Medicine Result Robustness** I test how these results respond to variations in the provider weights and the sample included. I try weighting providers equally rather than by their Medicare revenue. I exclude only the top 1% of outliers, and I include all physicians including ones who are not majority office-based. And the results in general do not change much.

### 5.2 Aggregate Analysis at the Provider Level

I expand the analysis to all providers to find the aggregate response across specialties. A nuance with this is that I compare trends only within specialty. The identification assumption for this DID is that providers who received larger reimbursement increases from the PE GPCI floor would have been on parallel trends with the providers who received smaller reimbursement increases in the absence of the floor. To extent that psychiatrists tend to provide less PE intensive care and hence have lower reimbursement increases than cardiologists, this assumption is implausible, but within specialty it is more plausible. Again there is concern that cardiologists, for example, that provide more PE intensive care and hence have higher reimbursement changes may not have been on parallel trends, but this can to some extent be evaluated by checking whether there are pre-trends or whether the falsification test in control areas shows a response.

The specification for the provider level analysis is the following:
\[ Y_{ist} = \sum_{\tau} \beta_{\tau} \mathbb{1}(\tau = t) \Delta Reimbursement_i + \gamma_{ist} + \gamma_i + \varepsilon_{ist} \] (4)

Where notation is defined as before with the addition that \( s \) indexes provider specialty, and \( \gamma_{ist} \) are specialty by year fixed effects. The individual provider fixed effects implicitly include specialty since that is a provider level attribute. I cluster my standard errors at the provider level. And I weight the providers by their total spending in 2008 and 2009.

**Provider Classification and Sample**  The sample included in the aggregate provider analysis includes all medical doctor specialties. The distribution of types of specialties and the number of providers and their Medicare revenue is shown in table 4. The claims data include HCFA specialty codes which are very precise. I group all medical doctors into primary care, medical specialties, surgical specialties, and other specialties following [Machado et al. (2021)](Machado et al. 2021), and within these categories define narrower groups to match the list of main specialties that [Chan and Dickstein (2019)](Chan and Dickstein 2019) uses in their paper. Appendix table C.1 lists how I group HCFA code for my analysis. I drop all providers whose HCFA code does not match one of the main specialties.

There are also many claims to non-physicians, for example, physical therapists, optometrists, and chiropractors. I do not include these in the main analysis.

Within each of the broad categories I drop the top and bottom 5% of providers by their Medicare revenue in 2008 and 2009 together. And I restrict to providers who receive at least 60% of their Medicare revenue for care in office-based settings.

**Results for All Physicians**  The aggregate results for all physicians suggest that providers with larger reimbursement increases have both extensive and intensive increases in healthcare provision. The main event study results from equation 4 are shown in figure 10. The blue squares are the DID estimates in the treated states, and the red circles show the falsification results.

I investigate the extensive margin response by looking at outcomes of indicators for office-based, facility-based, any place of service claims. Panel (a) shows that physicians with higher reimbursement increases were more likely to continue providing care (though the same pattern holds to a smaller degree in untreated states). Panel (b) shows that there is a relatively tight null result for the impact on substitution to facility-based care. And panel (c) shows that the increase in office-based care with no offset onto facility-based care yields more physicians filing claims of any sort.

The intensive margin response among physicians who continue to provide office-based Medicare services is positive in the long run. Panels (e) and (f) show that the in the short run, physicians
do not respond to the reimbursement increase, but in the longer run (3 to 5 years later) the number of services and number of RVUs increases for physicians with larger reimbursement increases.

Robustness: I test robustness to the same variations as I do for the FM case study: weighting providers equally rather than by their Medicare revenue, exclude only the top 1% of outliers, and including all physicians including ones who are not majority office-based. Again, the results in general do not change much.

5.3 Provider Level Analysis By Specialty

I estimate the provider level analysis separately for the most common specialties to see how homogeneous the main results are across specialties. I group physicians broadly into primary care, medical specialists, surgical specialists, and other specialists following Machado et al. (2021), and I separately evaluate responses among non-physicians.

Primary care specialties include family practice, internal medicine, and pediatrics (which are quite rare in Medicare). Unsurprisingly given that FM physicians make up almost 70% of primary care physicians, the aggregate primary care results are quite similar to the FM results. There is a positive extensive margin participation response for office-based care that is somewhat offset by decreased facility-based care (though the facility-response is noisy). And conditional on continuing to provide office-based care, increased reimbursement causes physicians to provide more. These results are shown in figure A.7 in the appendix since they are quite similar to the FM results.

For medical specialists the results are rather noisy, but overall, it seems that there are minimal if any impacts on extensive margin participation, and the results for intensive marginal responses are too noisy to say much. Medical specialties include dermatology, cardiology, medical oncology, neurology, and other smaller specialties listed in the appendix. The results are shown in figure 11 and the fourth panel of table 5. Since the standard errors are so large, it is impossible to rule out small positive effects.

For surgical specialists, it seems like the increased reimbursements increased both extensive margin participation and intensive margin quantities. These results are shown in figure 12 and the fifth panel of table 5. Surgical specialties include ophthalmology, urology, obstetrics and gynecology, and all specialties with “surgery” in the name. On all three margins of participation (office-based care, facility-based care, and any care) the physicians with larger reimbursement increases have increased participation—though the falsification tests in control areas show smaller yet similarly positive patterns. And for intensive margin quantities conditional on providing any office based care, reimbursement increases do cause an increase in provision.

I also look at the utilization response for other specialties: pathology, radiology, psychiatry, and
other specialties listed in the appendix that don’t fit into either medical or surgical specialties. The results for these specialties are quite noisy—possibly because they are a heterogeneous residual rather than a group with any similarities (the results are shown in the appendix in figure A.8). I similarly investigate the response for non-physicians (physical therapists, optometrists, chiropractors, etc listed in the appendix)—the intensive margin results are noisy and the falsification tests fail, but there is some evidence that there is a positive intensive margin response.

6 Conclusion

This paper evaluates how utilization responds to Medicare reimbursement changes using two DID strategies and a plausibly exogenous reimbursement change caused by the ACA. Utilization responds positively to reimbursement across both the area and the provider level specification. I show that the primary mechanism driving this response is the extensive margin choice of providers to provide any office-based services to Medicare patients. My focus on office-based care raises the question of substitution among places of service; for example, whether there is an offsetting increase in facility-based, or whether the physicians who exit from the office-based care market begin to affiliate with institutional providers. The positive result and extensive margin mechanism raise the concern that Medicare must consider access implications when it decreases reimbursement rates.

The positive elasticity has implications for Medicare’s budgeting process. If Medicare raised reimbursements by 1% and there were no endogenous quantity responses, the spending would increase by 1%. But with a elasticity of 2.4 (the point estimate from the area level specification for RVUs which aggregates service count and service intensity changes), the spending increase would include the 1% mechanical increase as well as the approximately 2.4% increase from the endogenous quantity response. To the extent the increase in office-based care is offset by decreases in facility-based care, that would offset this budgetary impact—in fact it could completely reverse it if the substitution were large enough because facility-based care is generally more costly. The extensive margin response also has implications for access as Medicare has generally been decreasing professional reimbursements. Though the implications vary depending on the extent to which facility-based care is substituting for office-based care—this means understanding spillovers onto facility-based care is a key follow up question.

References


Federal Register (2010). Medicare Program; Payment Policies Under the Physician Fee Schedule and Other Revisions to Part B for CY 2011; Final Rule.


7 Figures

Figure 1: Disequilibrium Outcomes

(a) Excess Supply or Excess Demand

(b) Physician Induced Demand

Note: This figure shows some possible disequilibrium outcomes. Panel (a) shows two possible disequilibriums with excess demand and excess supply where the volume transacted is limited by the short side of the market at an administratively set price—when the price increases, the disequilibrium outcome moves along an upward sloping supply curve or a downward sloping demand curve. Panel (b) shows a market with excess supply that still operates on the supply curve because providers induce demand—when the price increases, providers induce more demand (shifting the artificial induced demand curve out) and the disequilibrium outcome moves along the supply curve.
Figure 2: Backward Bending Supply

Note: This figure shows examples of disequilibrium outcomes on a backward bending supply curve with either excess demand or provider induced demand. Panel (a) shows excess demand and a backward bending supply curve where a price increase causes the disequilibrium outcomes to move along the backwarding bending supply curve. Panel (b) has exactly the same supply curve and response to the price change, but the providers must induce demand to operate on their supply curve; in this case, the induced demand curve also shifts out, but I do not show this to avoid crowding the graph.
Figure 3: Practice Expense GPCIs in Treated and Control States

(a) Treatment States

(b) States Providing Control Counties

Note: The left panel shows the practice expense GPCIs in the treated states where the 2011 GPCI floor was binding, in these states, the floor increased the GPCIs by about 10 to 15% relative to 2009. The right panel shows the practice expense GPCIs for neighboring states that I consider as control areas. Note that in many control states and all the treated states, there was a small increase in PE GPCI in 2010—this was the result of a temporary ACA provision.

Figure 4: Actual Reimbursements per RVU in Treated and Control States

(a) Treatment States

(b) States Providing Control Counties

Note: This shows the actual reimbursement rates Medicare paid per RVU for office-based care in the states affected by the ACA PE GPCI floor and in the set of states I use as potential controls. In treated states there was an approximately 15% increase, while in the control states there were much smaller increases.
Figure 5: Propensity Scores of Treatment and Control Counties

Note: This figure shows the histogram of propensity scores for each county in the treatment and control states, with counties weighted by population. The propensity score is the fitted value from the regression: $\mathbb{I}(treated_c) = X_c + \varepsilon_c$ where $c$ indexes county and $X_c$ includes the 2010 poverty rate, SNAP recipiency rate, share of population living on farms, number of hospital beds per capita, and psychiatrists per capita. The key take-away is that the propensity score have substantial overlap.

Figure 6: Map of Treatment and Matched Control Counties

Note: The treatment states are outlined in blue. The colors show the weight on each county in the treatment and control states. In the treatment states, the weight is the FFS population. In the control states the weight is the FFS population of the matched treatment counties. The key take-away is that the control counties do not place too much weight on a few small counties (which would cause noisy estimates) and do not place too much weight on the big cities in the control states (which would indicate that the matching algorithm did not do well).
Figure 7: Reimbursement Changes for Treatment and Matched Control Counties

Note: This shows the percent increase in reimbursement rate in treated and control counties caused by the GPCI floor; that is, the change in reimbursement that results from fixing procedure RVUs and procedure case mix at the 2009 levels and then assigning statutory reimbursement rates based on either 2009 or 2011 GPCIs. This isolates the most plausibly exogenous change in reimbursement and does not include endogenous changes in reimbursement rates caused by updates to RVUs or changes in the types of care provided in a county.
Figure 8: Area DID Results: Event Study

(a) Log Service Counts

(b) Log RVUs

(c) Log Number Providers

Note: This figure has the results from equation 2, the DID event study comparing counties based on the percent increase in reimbursement rates—this regresses the outcome on the county level reimbursement rate change interacted with year dummies and year and county fixed effects. The outcomes are log service counts, log RVUs, and log providers. Treated counties are weighted by their FFS enrollment and control counties by the FFS enrollment of their matched treatment county. Standard errors are clustered at the county level.
Note: Panel (a) is the histogram of percent reimbursement changes for family medicine physicians in treated and matched control counties caused by the GPCI floor. Note that the control providers did not actually see this reimbursement change. Each provider is weighted equally. Panels (b) through (f) have the results from equation 3, the DID model comparing family medicine physicians based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.
Figure 10: Provider Level Event Study (All Physicians)

Panel (a) is the histogram of percent reimbursement changes for physicians in treated and matched control counties caused by the GPCI floor. Note that the control providers did not actually see this reimbursement change. Each provider is weighted equally. Panels (b) through (f) have the results from equation the DID model comparing providers based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year by specialty fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.

Note: Panel (a) is the histogram of percent reimbursement changes for physicians in treated and matched control counties caused by the GPCI floor. Note that the control providers did not actually see this reimbursement change. Each provider is weighted equally. Panels (b) through (f) have the results from equation the DID model comparing providers based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year by specialty fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.
Figure 11: Medical Specialist Provider Level Regressions

(a) Histogram of Reimbursement Rate Changes
(b) Indicator for any Office-Based Claims
(c) Indicator for any Facility-Based Claims
(d) Indicator for any Medicare Claims
(e) Log Service Counts
(f) Log RVUs

Note: Panel (a) is the histogram of percent reimbursement changes for medical specialist physicians (cardiology, dermatology, gastroenterology, medical oncology, nephrology, neurology, rheumatology, urology, and a few smaller subspecialties, see the appendix). See note on figure 10 for more details. Panels (b) through (f) have the results from equation 4, the DID model comparing providers based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year by specialty fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.
Figure 12: Surgical Specialist Provider Level Regressions

(a) Histogram of Reimbursement Rate Changes
(b) Indicator for any Office-Based Claims
(c) Indicator for any Facility-Based Claims
(d) Indicator for any Medicare Claims
(e) Log Service Counts
(f) Log RVUs

Note: Panel (a) is the histogram of percent reimbursement changes for surgical specialist physicians (cardiac surgery, general surgery, neurosurgery, surgical oncology, orthopedic surgery, plastic surgery, and a few smaller subspecialties, see the appendix). See note on figure 10 for more details. Panels (b) through (f) have the results from equation 4, the DID model comparing providers based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year by specialty fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.
Table 1: Example Reimbursement Variation

Panel A: Montana GPCIs

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PE</td>
<td>0.847</td>
<td>1</td>
</tr>
</tbody>
</table>

Panel B: RVUs and Reimbursement Changes for Example Procedures

<table>
<thead>
<tr>
<th>Procedure</th>
<th>RVUs</th>
<th>GPCI-Adjusted Total RVUs</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2009</td>
<td>2011</td>
<td>(× Reimbursement)</td>
</tr>
<tr>
<td>Psychotherapy</td>
<td>Work (1)</td>
<td>1.86</td>
<td>PE (2)</td>
</tr>
<tr>
<td>Eye Exams</td>
<td>0.44</td>
<td>1.35</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Panel C: Example Area Reimbursement Changes

<table>
<thead>
<tr>
<th></th>
<th>Service Volume</th>
<th>Aggregate Adjusted RVUs</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychotherapy</td>
<td>(1)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Eye Exams</td>
<td>(2)</td>
<td>392.6</td>
<td>422</td>
</tr>
</tbody>
</table>

Note: Panel A shows Montana work and PE GPCIs for 2009 and 2011. Panel B shows RVUs for two procedures (psychotherapy and eye exams) in columns 1 through 4. Panel B shows how the RVUs and GPCIs translate into GPCI-adjusted total RVUs, which is proportional to reimbursement. Adjusted Total RVUs = RVU(work) * GPCI(work) + RVU(PE) * GPCI(PE) and are based off the Montana GPCI levels in Panel A. Columns 5 and 6 of Panel B have the 2009 and 2011 adjusted total RVUs. Column 7 has the adjusted total RVUs based on 2009 RVUs and 2011 GPCIs, this shows how the total would have changed if there had been no endogenous RVU changes. Column 8 has the actual percent change in adjusted total RVUs—the percentage difference between columns 5 and 6. Column 9 has the percent change in adjusted total RVUs caused by the exogenous GPCI floor—the percentage difference between columns 5 and 7, which does not include the RVU updates. Panel C shows how there procedures level changes translate into county level variation in reimbursement for two fictitious counties—the sad county that requires a lot of psychotherapy and has a smaller total change in aggregate adjusted RVUs since psychotherapy rates do not change as much as eye exam rates. The RVUs for psychotherapy are for current procedural terminology (CPT) code 90806, and eye exam RVUs are for CPT 92250.
<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FFS Enrollees</strong></td>
<td>686.8</td>
<td>1338.1</td>
</tr>
<tr>
<td><strong>Percent Persons in Poverty 2010</strong></td>
<td>14.55</td>
<td>13.51</td>
</tr>
<tr>
<td><strong>Food Stamp/SNAP Recipients 2010 per capita</strong></td>
<td>0.103</td>
<td>0.101</td>
</tr>
<tr>
<td><strong>Hospital Beds per Capita</strong></td>
<td>5.786</td>
<td>5.622</td>
</tr>
<tr>
<td><strong>Psychiatrists per Capita</strong></td>
<td>0.0323</td>
<td>0.0677</td>
</tr>
<tr>
<td><strong>Share living on Farms</strong></td>
<td>12.73</td>
<td>6.344</td>
</tr>
<tr>
<td><strong>Annual Office-Based Services per Enrollee</strong></td>
<td>16.92</td>
<td>32.50</td>
</tr>
</tbody>
</table>

*Note:* This table provides basic summary statistics for the matching variables used to match the treatment and control counties. The treated counties are weighted by their FFS enrollment, and the control counties by the FFS enrollment of their matched treatment county. I also include the average number of office-based services per enrollee to show that it is higher in the treated states.
<table>
<thead>
<tr>
<th>Table 3: Area DID Results: 2008 to 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Log Service Count</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Log RVUs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Log N Providers</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>N Counties</td>
</tr>
<tr>
<td>Control Weights</td>
</tr>
</tbody>
</table>

*Note:* Standard errors with county clustering in parentheses. Standard errors with state clustering in brackets. Based off county clustered standard errors, *$p < 0.05$, **$p < 0.01$. This table has the results from equation the DID model comparing counties based on the percent increase in reimbursement rates—this regresses the outcome on the county level reimbursement rate change interacted with an indicator for post-2011 and year and county fixed effects. The outcomes are log service counts, log RVUs, and log providers. Treated counties are weighted by their FFS enrollment and control counties by the FFS enrollment of their matched treatment county. The regressions include observations from 2008 to 2013. Column 1 shows the results for the baseline specification. Columns 2 and 3 have two versions that are plausibly unbiased but less powerful—column 2 restricts to counties on the border between treated and control areas, weighting counties by their own FFS enrollment, and column 3 restricts to only treated states and depends on variation within them. Columns 4 to 6 have likely biased variations—column 4 weights control counties by their own FFS enrollment, column 5 weights each control county by the minimum of its FFS enrollment and matching weight, and column 6 includes all counties in the states considered as potential controls weighting them by their own FFS enrollment.
Table 4: Providers in Treated States

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Revenue</th>
<th>PE Share</th>
<th>Reimbursement Change (%)</th>
<th>Mean</th>
<th>10%-ile</th>
<th>90%-ile</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MDs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary Care</td>
<td>1172</td>
<td>23796916</td>
<td>0.49</td>
<td></td>
<td>8.3</td>
<td>6.9</td>
<td>9.8</td>
</tr>
<tr>
<td>Family medicine</td>
<td>810</td>
<td>13169594</td>
<td>0.50</td>
<td></td>
<td>8.3</td>
<td>7.0</td>
<td>9.7</td>
</tr>
<tr>
<td>Internal medicine</td>
<td>335</td>
<td>10423591</td>
<td>0.49</td>
<td></td>
<td>8.2</td>
<td>6.7</td>
<td>9.8</td>
</tr>
<tr>
<td><strong>Medical Specialties</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dermatology</td>
<td>70</td>
<td>8560338</td>
<td>0.59</td>
<td></td>
<td>10.0</td>
<td>8.7</td>
<td>11.0</td>
</tr>
<tr>
<td>Cardiology</td>
<td>143</td>
<td>7022143</td>
<td>0.63</td>
<td></td>
<td>10.9</td>
<td>7.7</td>
<td>13.4</td>
</tr>
<tr>
<td>Oncology</td>
<td>68</td>
<td>3764917</td>
<td>0.64</td>
<td></td>
<td>11.3</td>
<td>8.8</td>
<td>13.1</td>
</tr>
<tr>
<td>Neurology</td>
<td>74</td>
<td>2137619</td>
<td>0.54</td>
<td></td>
<td>9.4</td>
<td>7.3</td>
<td>13.7</td>
</tr>
<tr>
<td>Rheumatology</td>
<td>45</td>
<td>1269061</td>
<td>0.59</td>
<td></td>
<td>10.3</td>
<td>7.5</td>
<td>13.5</td>
</tr>
<tr>
<td>Pulmonary medicine</td>
<td>47</td>
<td>1024615</td>
<td>0.48</td>
<td></td>
<td>8.2</td>
<td>6.1</td>
<td>11.7</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>55</td>
<td>736539</td>
<td>0.53</td>
<td></td>
<td>9.3</td>
<td>6.9</td>
<td>11.1</td>
</tr>
<tr>
<td>Nephrology</td>
<td>41</td>
<td>601635</td>
<td>0.42</td>
<td></td>
<td>6.9</td>
<td>5.4</td>
<td>7.9</td>
</tr>
<tr>
<td><strong>Surgical Specialties</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ophthalmology</td>
<td>127</td>
<td>9552386</td>
<td>0.52</td>
<td></td>
<td>8.8</td>
<td>7.7</td>
<td>9.7</td>
</tr>
<tr>
<td>Orthopedic surgery</td>
<td>265</td>
<td>4679936</td>
<td>0.56</td>
<td></td>
<td>9.8</td>
<td>7.8</td>
<td>11.9</td>
</tr>
<tr>
<td>Urology</td>
<td>82</td>
<td>3775670</td>
<td>0.57</td>
<td></td>
<td>9.7</td>
<td>8.5</td>
<td>10.8</td>
</tr>
<tr>
<td>General surgery</td>
<td>195</td>
<td>1858322</td>
<td>0.56</td>
<td></td>
<td>9.6</td>
<td>7.3</td>
<td>13.2</td>
</tr>
<tr>
<td>Otolaryngology</td>
<td>78</td>
<td>1292345</td>
<td>0.55</td>
<td></td>
<td>9.3</td>
<td>7.7</td>
<td>10.9</td>
</tr>
<tr>
<td>Obstetrics and gynecology</td>
<td>199</td>
<td>835990</td>
<td>0.54</td>
<td></td>
<td>9.0</td>
<td>7.3</td>
<td>10.2</td>
</tr>
<tr>
<td>Neurosurgery</td>
<td>50</td>
<td>521984</td>
<td>0.57</td>
<td></td>
<td>10.0</td>
<td>6.3</td>
<td>17.3</td>
</tr>
<tr>
<td>Plastic surgery</td>
<td>40</td>
<td>279686</td>
<td>0.52</td>
<td></td>
<td>9.3</td>
<td>7.6</td>
<td>10.4</td>
</tr>
<tr>
<td><strong>Other Specialties</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiology</td>
<td>213</td>
<td>6235088</td>
<td>0.70</td>
<td></td>
<td>12.8</td>
<td>5.1</td>
<td>17.4</td>
</tr>
<tr>
<td>Pathology</td>
<td>57</td>
<td>1409916</td>
<td>0.51</td>
<td></td>
<td>9.0</td>
<td>3.5</td>
<td>13.5</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>75</td>
<td>1103918</td>
<td>0.35</td>
<td></td>
<td>5.9</td>
<td>4.2</td>
<td>7.5</td>
</tr>
<tr>
<td><strong>Non-MDs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Therapy</td>
<td>494</td>
<td>6732906</td>
<td>0.40</td>
<td></td>
<td>6.9</td>
<td>5.7</td>
<td>7.4</td>
</tr>
<tr>
<td>Optomitrist</td>
<td>469</td>
<td>6349887</td>
<td>0.52</td>
<td></td>
<td>8.7</td>
<td>7.6</td>
<td>9.9</td>
</tr>
<tr>
<td>Chiropractic</td>
<td>825</td>
<td>5499305</td>
<td>0.31</td>
<td></td>
<td>4.9</td>
<td>4.3</td>
<td>5.5</td>
</tr>
<tr>
<td>PA</td>
<td>452</td>
<td>3320930</td>
<td>0.51</td>
<td></td>
<td>8.7</td>
<td>6.9</td>
<td>10.7</td>
</tr>
<tr>
<td>NP</td>
<td>414</td>
<td>2845860</td>
<td>0.49</td>
<td></td>
<td>8.4</td>
<td>6.5</td>
<td>11.4</td>
</tr>
<tr>
<td>Podiatry</td>
<td>94</td>
<td>2356226</td>
<td>0.49</td>
<td></td>
<td>8.7</td>
<td>7.4</td>
<td>9.7</td>
</tr>
<tr>
<td>Psychology</td>
<td>205</td>
<td>1289369</td>
<td>0.27</td>
<td></td>
<td>4.6</td>
<td>3.3</td>
<td>7.1</td>
</tr>
<tr>
<td>Clinical Social Worker</td>
<td>176</td>
<td>846761</td>
<td>0.24</td>
<td></td>
<td>4.1</td>
<td>3.5</td>
<td>4.4</td>
</tr>
<tr>
<td>Occupational Therapy</td>
<td>66</td>
<td>402283</td>
<td>0.41</td>
<td></td>
<td>7.0</td>
<td>6.7</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Note: This lists the common specialties nested into larger categories (the summary statistic for larger categories includes information on smaller specialties that I omit individual information on). N is the number of distinct providers with office-based care in both 2008 and 2009—this does not restrict to primarily office-based providers as I do in my analysis. Revenue is the sum of all office-based Medicare reimbursements for the 20% sample of patients in 2008 and 2009. The reimbursement changes include only the changes induced by the exogenous GPCI floor—I calculate these by applying 2009 and 2011 GPCIs to 2009 RVUs and case mix.
Table 5: Provider DID Results by Specialty

<table>
<thead>
<tr>
<th>Specialty</th>
<th>Treated</th>
<th>Falsification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (Office)  (2) (Facility)  (3) (Any) Ln(Count)</td>
<td>(4)</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>All MDs</td>
<td>0.815 (0.535)               -0.0537 (0.653) 0.855* (0.412) -1.987 (2.048)</td>
<td>0.522* (0.259) 0.168 (0.368) 0.600** (0.228) -0.675*** (1.036)</td>
</tr>
<tr>
<td></td>
<td>2.171 (0.740)               0.117 (0.765) 1.229* (0.590) 6.031* (2.805)</td>
<td>1.302*** (0.375) 0.344 (0.413) 1.047** (0.328) -0.420 (1.864)</td>
</tr>
<tr>
<td>Primary Care</td>
<td>2.081 (1.160)               -0.939 (1.018) 0.452 (0.824) 10.48 (4.310)</td>
<td>0.436 (0.474) 0.844 (0.604) 1.016 (0.365) -6.799* (1.707)</td>
</tr>
<tr>
<td>Medical Specialists</td>
<td>0.412 (1.071)               1.258 (0.820) 1.223 (0.640) 9.573 (4.064)</td>
<td>0.520 (0.396) -0.514 (0.486) 0.773* (0.380) -5.154** (1.580)</td>
</tr>
<tr>
<td>Surgical Specialists</td>
<td>2.011 (1.102)               2.126 (1.244) 1.640 (1.067) 3.510 (2.621)</td>
<td>1.396** (0.456) 1.197* (0.564) 1.077* (0.444) 0.898 (1.357)</td>
</tr>
<tr>
<td>Other MDs</td>
<td>0.145 (0.823)               -2.022 (1.536) 0.290 (0.793) -4.874 (3.667)</td>
<td>-0.0873 (0.626) -0.192 (1.005) 0.302 (0.524) 5.485 (2.866)</td>
</tr>
</tbody>
</table>

Note: This table has the results from equation 3 in the first panel and from equation 4 for various groups in subsequent rows. Columns 1 though 3 have the extensive margin results from the outcomes of indicators for any office-based claims, and facility-based claims, and any claims at and column 4 has the intensive margin results for the log service count. Columns 5 through 8 have the analogous results from similar regression run on providers in control states who did not receive reimbursement increases from the GPCI Floor. I split each DID regression into a short run (1 to 2 year) and long run (3 to 5 year) effect and show coefficients for each of those.

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors in parentheses.
A Extra Tables and Figures

Figure A.1: Area DID: Underlying Binscatters

(a) Log Service Counts

(b) Log RVUs

(c) Log Number Providers

Note: This figure has binscatters that summarize the underlying data in the area DID event studies shown in figure 8.
Figure A.2: Robustness of the Main Result

Note: This figure has the results from equation 2, DID model comparing counties based on the percent increase in reimbursement rates—this regresses the outcome on the county level reimbursement rate change interacted with year dummies and year and county fixed effects. The outcomes are log service counts. Standard errors are clustered at the county level. Panel (a) shows the results for the baseline specification—this repeats the result in figure 8 for comparison. Panels (b) though (f) show that the main result is not robust to several variations: panel (b) weights control counties by their own FFS enrollment, panel (c) weights control counties by the minimum of their own FFS enrollment and their share of the FFS enrollment of their matched treatment counties, panel (d) restricts to only treated states and depends on variation within them (the change in reimbursement varies from 5 to about 12%), panel (e) includes all counties in the states considered as potential controls weighting them by their own FFS enrollment, and panel (e) restricts to counties on the border between treated and control weighting counties by their own FFS enrollment.
Figure A.3: Family Medicine Extensive Margin Binscatters

(a) Treated States

(b) Falsification

Note: This figure displays binscatters of the share of providers providing any office-based care by their quartile of reimbursement rate change (or counterfactual change for the falsification). Each provider is weighted by their total Medicare revenue from 2008 and 2009. The takeaway is that in treated states, the quartiles seem to be fanning out after 2011 more than they do in control states.
Figure A.4: Family Medicine Facility-Based Care Binscatters (Spending Weighted)

(a) Treated States

(b) Falsification

Note: See note for figure A.3, the only difference is that the outcome is facility-based care. The takeaway is that the lowest quartile of providers in treatment states seem to be substituting to facility-based care.
Figure A.5: Family Medicine Falsification Providers Reimbursement Changes

Note: This figure is a binscatters of the reimbursement rate change that providers in the control states faced actually faced due to GPCI changes between 2009 and 2011 in their state and the reimbursement rate change they would have faced had they faced the average change in GPCIs that the treated states faced. Each provider is weighted by the total Medicare revenue from 2008 and 2009.
Figure A.6: Family Medicine Specialty Provider Level Regressions—Provider Weighted

(a) Histogram of Reimbursement Rate Changes

(b) Indicator for any Office-Based Claims

(c) Indicator for any Facility-Based Claims

(d) Indicator for any Medicare Claims

(e) Log Service Counts

(f) Log RVUs

Note: See figure 9, the only difference is that here I weight each physician equally.
Figure A.7: Primary Care Provider Level Regressions

(a) Histogram of Reimbursement Rate Changes
(b) Indicator for any Office-Based Claims
(c) Indicator for any Facility-Based Claims
(d) Indicator for any Medicare Claims
(e) Log Service Counts
(f) Log RVUs

Note: Panel (a) is the histogram of percent reimbursement changes for generalist physicians (family medicine and internal medicine). See note on figure 10 for more details. Panels (b) through (f) have the results from equation 4, the DID model comparing providers based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year by specialty fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.
Figure A.8: Provider Level Event Study (Other Specialty Physicians)

(a) Histogram of Reimbursement Rate Changes

(b) Indicator for any Office-Based Claims

(c) Indicator for any Facility-Based Claims

(d) Indicator for any Medicare Claims

(e) Log Service Counts

(f) Log RVUs

Note: Panel (a) is the histogram of percent reimbursement changes for non-physicians in treated and matched control counties caused by the GPCI floor. Note that the control providers did not actually see this reimbursement change. Each provider is weighted equally. Panels (b) through (f) have the results from equation 4, the DID model comparing providers based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year by specialty fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.
Figure A.9: Provider Level Event Study (All Non-Physicians)

(a) Histogram of Reimbursement Rate Changes
(b) Indicator for any Office-Based Claims
(c) Indicator for any Facility-Based Claims
(d) Indicator for any Medicare Claims
(e) Log Service Counts
(f) Log RVUs

Note: Panel (a) is the histogram of percent reimbursement changes for non-physicians in treated and matched control counties caused by the GPCI floor. Note that the control providers did not actually see this reimbursement change. Each provider is weighted equally. Panels (b) through (f) have the results from equation 4, the DID model comparing providers based on the percent increase in reimbursement rates—this regresses the outcome on the provider level reimbursement rate change interacted with year dummies and provider and year by specialty fixed effects. The outcomes are indicated in the captions. Providers are weighted by their FFS revenue in 2008 and 2009. Standard errors are clustered at the provider level.
B More Complete Discussion of the PFS Reimbursement Formula

A more complete version of formula that Medicare uses to set reimbursement rates is below—this includes the Malpractice liability insurance components.

\[
Reimbursement_{pj} = ConversionFactor \times \left[ RVU_{p}(\text{work}) \times GPCI_{j}(\text{work}) \\
+ RVU_{p}(\text{practice expense}) \times GPCI_{j}(\text{practice expense}) \\
+ RVU_{p}(\text{insurance}) \times GPCI_{j}(\text{insurance}) \right]
\]

Where \( p \) indexes procedure and \( j \) indexes locality. The parameters change over time, so there is an implicit time subscript. There are some adjustments to the formula that factor into reimbursement: place of service (facility and non-facility rates vary), modifier codes, multiple services on the same day.

C Data Notes

C.1 Specialty Definitions

I group HCFA specialty codes that are provided in the carrier claims into slightly broader specialties. I categorize these groups to match those in Chan and Dickstein (2019) with the addition of physical therapy, occupational therapy, chiropractic, and psychology. These categorization I made is shown in table C.1. This is the categorization I use for the provider summary statistics in table 4.

I then group these larger specialty groups into primary care, medical specialties, surgical specialties, and other specialties following Machado et al. (2021). This grouping is described in table C.2.
<table>
<thead>
<tr>
<th>Grouped Specialties</th>
<th>HCFA Codes included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anesthesiology</td>
<td>5, 9, 72</td>
</tr>
<tr>
<td>Cardiac Surgery</td>
<td>78</td>
</tr>
<tr>
<td>Cardiology</td>
<td>C3, 6, 21, 23, 31, 76</td>
</tr>
<tr>
<td>Chiropractic</td>
<td>35</td>
</tr>
<tr>
<td>Clinical Social Worker</td>
<td>80</td>
</tr>
<tr>
<td>Colorectal surgery</td>
<td>28</td>
</tr>
<tr>
<td>Dermatology</td>
<td>7</td>
</tr>
<tr>
<td>Emergency medicine</td>
<td>93</td>
</tr>
<tr>
<td>Family medicine</td>
<td>1, 8, 38</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>10</td>
</tr>
<tr>
<td>General surgery</td>
<td>2</td>
</tr>
<tr>
<td>Infectious disease</td>
<td>44</td>
</tr>
<tr>
<td>Internal medicine</td>
<td>11</td>
</tr>
<tr>
<td>Internal medicine subspecialties</td>
<td>46, 82</td>
</tr>
<tr>
<td>Nephrology</td>
<td>39</td>
</tr>
<tr>
<td>Neurology</td>
<td>13, 86</td>
</tr>
<tr>
<td>Neurosurgery</td>
<td>14</td>
</tr>
<tr>
<td>Nurse Practicioner</td>
<td>50</td>
</tr>
<tr>
<td>Obstetrics and gynecology</td>
<td>16</td>
</tr>
<tr>
<td>Occupational Therapy</td>
<td>67</td>
</tr>
<tr>
<td>Oncology</td>
<td>83, 90, 91, 98</td>
</tr>
<tr>
<td>Ophthalmology</td>
<td>18</td>
</tr>
<tr>
<td>Optometrist</td>
<td>41</td>
</tr>
<tr>
<td>Orthopedic surgery</td>
<td>20, 40</td>
</tr>
<tr>
<td>Osteopathic manipulative therapy</td>
<td>12</td>
</tr>
<tr>
<td>Other</td>
<td>C0, 17, 25, 84, 79, 81</td>
</tr>
<tr>
<td>Otolaryngology</td>
<td>4</td>
</tr>
<tr>
<td>Pathology</td>
<td>22</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>37</td>
</tr>
<tr>
<td>Physical Therapy</td>
<td>65</td>
</tr>
<tr>
<td>Physicians Assistant</td>
<td>97</td>
</tr>
<tr>
<td>Plastic surgery</td>
<td>24, 85</td>
</tr>
<tr>
<td>Podiatry</td>
<td>48</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>26, 27</td>
</tr>
<tr>
<td>Psychology</td>
<td>62, 68</td>
</tr>
<tr>
<td>Pulmonary medicine</td>
<td>29</td>
</tr>
<tr>
<td>Radiation oncology</td>
<td>92</td>
</tr>
<tr>
<td>Radiology</td>
<td>30, 36, 94</td>
</tr>
<tr>
<td>Rheumatology</td>
<td>3, 66</td>
</tr>
<tr>
<td>Thoracic surgery</td>
<td>33</td>
</tr>
<tr>
<td>Urology</td>
<td>34</td>
</tr>
<tr>
<td>Vascular surgery</td>
<td>77</td>
</tr>
</tbody>
</table>
### Table C.2: Categorizing Specialty Codes to Broad Groups

<table>
<thead>
<tr>
<th>Category</th>
<th>Specialties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Care</td>
<td>Family Medicine, Internal Medicine, Pediatrics</td>
</tr>
<tr>
<td>Medical Specialties</td>
<td>Cardiology, Dermatology, Gastroenterology, Endocrinology, Hematology, Nephrology, Neurology, Medical Oncology, Pulmonary Medicine, Rheumatology</td>
</tr>
<tr>
<td>Other Specialties</td>
<td>Anesthesiology, Emergency Medicine, Infectious Disease, Neuropsychiatry, Hospice and Palliative Care, Sports Medicine, Physical Medicine, Preventive Medicine, Critical Care, Sleep Medicine, Pathology, Psychiatry, Psychology, Radiation oncology, Radiology</td>
</tr>
<tr>
<td>Non Physicians</td>
<td>Chiropractic, Clinical Social Worker, Nurse Practitioner, Occupational Therapy, Optometrist, Osteopathic Manipulative Therapy, Physician Assistant, Physical Therapy, Podiatry</td>
</tr>
</tbody>
</table>

*Note: I group specialties together following [Machado et al., 2021](#), and using common sense to include non-physicians and specialties that they do not include.*