What Constrains Medicare Fraud? The Role of Firm Culture in Nursing Homes[1]

John M. Griffin

Alex Priest

ABSTRACT

We uncover persistent variation in Medicare billing intensity that is not explained by sicker patients, patient selection, or coding completeness, but by systematic upcoding of specific comorbidities absent in the immediately preceding hospital stay. Substantial cross-sectional variation in upcoding is largely unrelated to factors emphasized in the literature such as billing regimes, enforcement, and competition. Upcoding appears driven by firm-wide practices that spread across facilities within skilled nursing facility (SNF) firms and through rapid acquisitions. For-profit and private equity firms have higher upcoding on average, but wide differences across firms of all ownership types underscore the importance of firm-level ethos.

[1]We appreciate helpful comments from Matthew Byrnes, Gonul Colak (discussant), Willie Fuchs, John Hatfield, Samuel Kruger, Liam Landrum, Riley, League, Jetson Leder-Luis, Ryan McDevitt (discussant), Ashvin Gandhi (discussant), Shan Ge, Christoph Herpfer (discussant), Joseph Kalmenovitz, Andrey Ordin, Marius Ring, Andrea Rossi (discussant), Jonathan Skinner, Rene Stulz, Billy Xu, and seminar participants at Conference on the Economics and Finance of Healthcare and Medicine at Washington University in Saint Louis, Conference on Financial Economics and Accounting, Eastern Finance Association, Financial Management Association 2025, Iowa State University, the NBER Economics of Aging, The Ohio State University, the University of Rochester, and the University of Texas-Austin. We further thank Integra Med Analytics for data assistance and research support. Integra Med Analytics is engaged in Medical Fraud detection and is a party to lawsuits against fraudulent health care parties including parties in skilled nursing. Integra Med Analytics also consults on issues in health care. Griffin is also the Owner/CEO of Integra Med Analytics, Integra FEC, and Integra Research Group. Outputs from this study and other data are aggregated at a facility level and available in an interactive map at https://www.care.watch/. Griffin is the James A. Elkins Centennial Chair in Finance at the McCombs School of Business, University of Texas, Austin. Email: john.griffin@utexas.edu. Priest is an Assistant Professor at the Simon Business School, University of Rochester and an academic collaborator with Integra Med Analytics. Email: apriest@simon.rochester.edu.

1. Introduction

In many developed economies, healthcare is publicly financed but delivered through private firms, raising questions about how firm incentives and organizational practices may hinder the efficient allocation of limited public funds across providers and patients. In the United States, Medicare spending from 2000 to 2024 increased at nearly twice the rate of inflation and now represents 14% of total federal spending.[2] While there are many potential explanations behind the spending increase, there is widespread concern that a sizeable proportion of Medicare spending is fraudulent or wasteful.[3] Understanding the economic forces that drive fraud is crucial for effective allocation of public resources, but fraud is often difficult to detect and accurately measure. We uncover and quantify widespread Medicare fraud in the nursing home industry and examine the determinants.

Healthcare fraud can arise from a combination of imperfect reimbursement rates and inadequate enforcement (Leder-Luis and Malani, 2025). When reimbursement rates are set too high, providers have an incentive to deliver medically unnecessary services. While enforcement can act as a constraining force, in the absence of credible punishment, providers may overtreat, undertreat, or misreport care, distorting the quantity and accuracy of healthcare services. In contrast, variation in the values and practices of owners and managers, often known as corporate culture (O'Reilly and Chatman, 1996; Guiso et al., 2015), could explain differences in billing behavior. Using precisely measured differences in facility billing practices over two Medicare reimbursement schemes, we assess whether imperfect reimbursement, lack of enforcement, firm norms, other factors including

[2]Spending increased from $5,844 per beneficiary in 2000 to $17,786 by 2024. https://www.cms.gov/oact/tr/2025, https://www.pgpf.org/budget-basics/medicare, https://www.gao.gov/products/gao-24-107487.

[3]Meta-studies have found that fraud and waste make up 25% of total healthcare spending (Shrank et al., 2019). Within the context of Medicare, at least 7.4% of payments fail to meet statutory or legal requirements and are considered improper (https://www.cms.gov/newsroom/fact-sheets/fiscal-year-2023-improper-payments-fact-sheet).

competition, firm sophistication, local cultural, or social norms explain cross-sectional variation in Medicare fraud.

From 2010 to September 2019, Medicare reimbursement for SNF care was determined by the Resource Utilization Group IV (RUG-IV), which tied payments to the amount of therapy minutes provided. Over time, therapy rates increased dramatically despite observably similar patient characteristics, raising substantial concerns that facilities were providing more rehab than medically necessary (Temkin-Greener et al., 2019). We find considerable differences in rehab therapy levels across facilities for observationally similar patients. In October 2019, RUG-IV was replaced with the Patient Driven Payment Model (PDPM), which was designed to redirect reimbursement from intensive therapy to clinically complex patients. We construct a measure of billing intensity based on the use of high-reimbursement codes beyond what can be explained by observable patient characteristics and diagnoses at the referring hospital. Widespread differences in PDPM billing intensity exist across counties, and even larger variation across facilities within a county. These differences are not explained by gender, age, or race, but rather past facility-level billing practices: those with the most aggressive rehab under RUG-IV also code at much higher rates under PDPM. The revenue gap is substantial: facilities with the highest PDPM coding capture an additional $10,400 per patient ($24,970 vs. $14,540).

Facilities might have higher levels of PDPM coding intensity because they treat sicker patients, selectively admit patients with more comorbidities, more accurately identify underlying patient comorbidities, or aggressively but incorrectly code patients. We distinguish between these possibilities using four main tests. First, facilities with the highest levels of RUG-IV rehab (from 2016 to September 2019) show a sharp increase in coding of the highest-compensating comorbidities precisely when

PDPM was enacted. Second, inconsistent with patient selection, these same facilities commonly code comorbidities that are not present at the referring hospital stay immediately preceding the SNF visit. Third, we utilize an instrumental variables approach to capture instances in which patients are likely to be admitted to a particular SNF because other local facilities are close to full occupancy. Instrumental variable estimates imply that patients who visit opportunistic facilities experience a 42% increase in coding intensity. Fourth, inconsistent with more complete coding practices, the facilities that previously engaged in excessive rehab practices only code higher on compensating diagnoses while recording slightly fewer non-compensating diagnoses. Overall, the evidence is consistent with the facilities that previously provided the highest levels of rehab under RUG-IV upcoding patients to extract additional revenue under PDPM.

Why do levels of upcoding vary so drastically across facilities? To investigate these questions, we examine the RUG-IV reimbursement era from October 2011 to September 2019 when payment was conditioned on rehab therapy, which is easily verifiable. Do much higher observed levels of therapy at certain facilities reflect medical necessity? Skilled nursing facilities providing the highest tercile of abnormal rehab in the RUG-IV regime are 6.65 times more likely to deliver weekly therapy in a ten-minute window just above the highest reimbursement threshold than in the entire 210-minute range directly below, strongly suggesting rehab provision based on financial incentive rather than medical necessity. If therapy overutilization arises from misaligned incentives from the reimbursement scheme, then cross-facility variation should reflect varying costs of providing therapy. Instead, the facilities that administered excessive amounts of ultra-high rehab also engaged in PDPM upcoding. These findings are contrary to CMS and industry projections, as the two regimes rewarded drastically different patient characteristics. The evidence highlights that eliminating incentive misalignment inherent in RUG-IV did not eliminate fraud but rather changed the type of overbilling from providing medically unnecessary care to upcoding.

What explains persistent facility-level differences in healthcare fraud across both reimbursement regimes? We first examine the role of enforcement by measuring DOJ enforcement actions. While DOJ actions vary geographically, we find no evidence that enforcement rates are related to cross-sectional differences in coding across markets, likely because penalties are too infrequent to deter fraudulent behavior. Fraudulent practices might be easier to sustain if a provider enjoys a local monopoly, but we find no evidence that this is the case. There is a statistically positive but economically weak correlation between facility size and upcoding levels. We further examine twelve variables that the literature has identified as related to crime, corruption, demographics, or social capital, and machine learning but find that they explain little of the variation in upcoding across facilities.

The coding rates at other facilities in the same SNF system explain more than 30% of the variation in facility-level billing practices. Effects are economically large with a one standard deviation increase in coding at other facilities within a system associated with a two-thirds of a standard deviation increase in coding intensity. Consistent with a causal channel, use of the highest-compensating patient codes increases significantly following acquisitions by opportunistic SNF systems, despite no increase in incoming patient sickness when recorded by the hospital. We further examine the spread of coding practices within a SNF system and find that a one-standard-deviation increase in upcoding at peer facilities increases expected next-quarter coding intensity at a facility by 0.23 standard deviations, even after controlling for the facility's past coding. This effect is highly persistent up to seven quarters in the future and further predicts coding variation across states and at distances more than 1,000 miles, which would be inconsistent with alternative explanations driven by shared local demand shocks, local enforcement, or local market structure. Firm norms or corporate culture appear to drive upcoding behavior, but a culture of integrity appears to be eroding over time as upcoding practices spread within systems and through acquisitions. The highest upcoding systems expand 78% faster than the lowest upcoding systems.

Is ownership type related to upcoding? Nonprofit and governmental facilities typically engage in less aggressive coding than traditional for-profit and private equity-backed facilities. Systems engaged in upcoding practices tend to be small- to medium-sized and privately held. Private equity-owned SNF systems code more aggressively on average but own only 7.02% of the facilities in the top tercile of opportunistic practices. While the majority of the most (least) aggressive coding facilities are controlled by traditional for-profit (non-profit) operators, there exists substantial variation in practices within each ownership type, highlighting the importance of heterogeneous firm culture.

There are several practical implications from our study. First, firm-wide norms appear to be the most important determinant of SNF healthcare fraud by far. The most opportunistic systems bill 36% more per patient than other facilities, generating excess Medicare costs of more than $2.10 billion annually. Second, although PDPM was designed to stop wasteful and unnecessary Medicare spending, skilled nursing facilities quickly gamed the new payment structure, indicating that reimbursement design changes should carefully consider how fraudulent actors might exploit such schemes. Third, historical DOJ enforcement does not appear to be a sufficiently binding constraint on fraud, with upcoding practices spreading rapidly within-system and across-system through acquisitions. Substantially more resources towards enforcement and criminal penalties appear necessary to change pervasive practices for certain SNF systems.

2. Literature Review and Skilled Nursing Facility Background

2.1. Related Research

Our paper builds on several strands of literature on fraud, incentives, and the allocation of public funds. In standard economics-of-crime models(Becker, 1968), individuals commit fraud when the expected gains outweigh the expected punishment, which is a function of the probability of being caught and the severity of punishment. In healthcare markets, severe informational frictions between payers, providers, and patients make these incentive problems particularly acute and create scope for moral hazard (Arrow, 1963). Leder-Luis and Malani (2025) develop a tractable model of healthcare fraud in which fraud occurs due to a misalignment of payment prices and a corresponding lack of enforcement. They define fraud as "any divergence between the care an insurer says a patient qualifies for, the care a provider provides, and the care a provider bills for." Medically unnecessary care occurs when a patient receives treatment at a level beyond what is deemed appropriate or necessary, and upcoding occurs when a provider bills for a service with a higher reimbursement rate than allowed or provided.[4] In the broader criminology literature, Sutherland's (1939; 1947) Differential Association Theory emphasizes that the technical knowledge of how to commit crime is necessary but not sufficient; one must also acquire the definitions of crime, including the norms, attitudes, and rationalizations. Organizations can shape these norms and sustain fraud levels (Matsueda, 1988).

Second, a broader literature emphasizes the importance of corporate or organizational culture (Kreps et al., 1990; O'Reilly and Chatman, 1996) more generally. O'Reilly and Chatman (1996) define organizational culture as "a system of shared values (that define what is important) and

[4] Leder-Luis and Malani (2025) highlight a legal debate over whether incentives (Feldman, 2001) or social norms (Hyman, 2001) drive healthcare fraud. A survey of the literature by Galizzi et al. (2023) finds that physicians typically act altruistically, though with considerable heterogeneity.

norms that define appropriate attitudes and behaviours for organizational members." Guiso et al. (2015) note that a firm's corporate culture, including integrity, can act as a constraint on activities and may relate to positive long-run firm performance.[5] Dupont and Karpoff (2025) model how 'social capital' in the form of honest behavior influenced by social and cultural norms can act as "a substitute and complement to legal and market forces to discipline opportunistic behavior and build trust." Differences in financial maleficence practices have been noted for financial advisor misconduct (Egan et al., 2019; Dimmock et al., 2018) and residential and commercial mortgage fraud (Piskorski et al., 2015; Griffin and Priest, 2023). Our results highlight the importance of substantial heterogeneous firm culture in the companies that commit Medicare fraud. Third, there is empirical work examining factors related to fraud and crime. O'Malley et al. (2021) find that home healthcare fraud spreads slowly over time through patient-sharing networks. Chen et al. (2018) examine characteristics of physicians excluded from Medicare for committing fraud, and Gertler and Kwan (2024) document that profit-motivated providers reported more false positive tests to sell unnecessary malaria drugs in Kenya. Leder-Luis and Malani (2025) note that the literature on incentives to commit healthcare fraud is sparse. More generally, Eide et al. (2006) survey the broader literature on crime determinants and conclude that the probability of being caught, and to a lesser extent, the severity of punishment deters crime; aside from unemployment, the evidence for other factors is less clear.[6] Utilizing a detailed measure for fraud in skilled nursing, we examine the role

[5] Li et al. (2021) find that corporate culture as measured in earnings call transcripts is related to a variety of firm outcomes including efficiency, risk-taking, and earnings management.

[6] Dills et al. (2008) discuss the literature on crime determinants and conclude that "economists know little about the empirically relevant determinants of crime." Crime and fraud has been linked to a multitude of factors such as the legal system, the presence of police (Di Tella and Schargrodsky, 2004), probability and severity of punishment (Witte, 1980), unemployment and wage levels (Gould et al., 2002), inequality (Fajnzylber et al., 2002), education (Lochner and Moretti, 2004), culture (Fisman and Miguel, 2007; Liu, 2016), social interactions within geographies (Glaeser et al., 1996), office co-workers (Dimmock et al., 2018), and media coverage by local newspapers (Heese et al., 2022). The influence of peers on committing criminal acts has been shown in a wide variety of contexts as surveyed by McGloin and Thomas (2019).

of competing forces related to widespread differences in fraudulent practices.

Fourth, there is a literature linking ownership structure and internal organization to healthcare billing strategies. Howell and Liu (2025) survey the literature on the role of private equity in healthcare and find that it increases prices in a variety of healthcare settings such as hospitals, nursing homes, and physician practices, which can be partially explained through superior bargaining ability of private equity (Liu, 2022). In the skilled nursing industry, Gupta et al. (2023) find that PE ownership of nursing decreases patient well-being and increases mortality. For-profit SNFs are more likely to provide high amounts of Ultra-High rehab therapy (Bowblis et al., 2016; Temkin-Greener et al., 2019) and for-profit hospices commit fraud (Gruber et al., 2025). Hospitals converting from non-profit to for-profit reduce spending on unprofitable activities such as emergency rooms (Herpfer et al., 2024). Internal firm organization and incentives (Gaynor et al., 2004; Clemens and Gottlieb, 2014) as well as vertical organizational structure with related entities (Geruso and Layton, 2020; Chen et al., 2016) and chain-ownership (Andreyeva et al., 2024) also affect healthcare billing. Provider incentives and ownership structures can substantially distort how medical care is allocated, generating large geographic variation in the use of profitable services that offer limited health benefits (Brownlee et al., 2017; Afendulis and Kessler, 2007). League (2022) finds that firms respond to increasing Medicare denial rates for physician services by consolidating into larger groups and increasing investment in billing technology. Ownership concentration has been linked to increased prices by hospitals (Gaynor and Town, 2011), health insurance providers (Dafny et al., 2012), and strategic supply reduction more generally at the costs of taxpayers (Doraszelski et al., 2024). Eliason et al. (2020) find that dialysis chains transfer billing and cost-cutting strategies through acquisitions. Our findings extend this literature by documenting substantial heterogeneity in fraudulent billing practices within common ownership types and the rapid expansion of fraudulent systems. Our results also suggest that the significant focus on private equity owners[7] might miss many potential fraudulent firms that do not appear to be linked to private equity.

Fifth, there is a literature examining potentially false billing practices and potential methods of preventing fraud and abuse. Hospital upcoding has been documented in a variety of settings (Silverman and Skinner, 2004; Dafny, 2005; Heese et al., 2015; Bastani et al., 2019; Joiner et al., 2024; Shekhar et al., 2023). Bowblis and Brunt (2014) find evidence that SNFs provide additional therapy to increase revenue.[8] We are the first paper to comprehensively examine potential skilled nursing fraud under the new PDPM regime.

Leder-Luis (2023) finds that privatization of enforcement activity may be an effective way to deter waste and abuse as False Claims Act cases generate deterrent effects that were approximately ten times larger than the direct settlement amounts. Howard and McCarthy (2021) also finds similarly large deterrent effects from DOJ settlements on unnecessary heart procedures. SNFs may shield themselves from legal liability by tunneling assets through related management and real estate entities (Gandhi and Olenski, 2024). Shi (2023) shows that Medicare's Recovery Audit Contractor (RAC) Program generates more than 24 times the cost in future savings. Through the staggered adoption of regulatory changes in ambulance reimbursement across states, Eliason et al. (2021) find that administrative changes are much more effective than enforcement. Our results suggest that regulators seeking to improve the allocational efficiency of public funds should carefully consider the importance of firm-level practices in enforcement actions, regulatory and billing approvals.

[7]For example, https://www.wsj.com/articles/antitrust-authorities-take-aim-at-privateequity-healthcare-deals-11655243804?mod=article_inline and https://www.wsj.com/articles/federal-agencies-probe-privateequity-profiteering-in-healthcare-2ba4c23a

[8]Skilled nursing facilities discriminate against patients with lower expected reimbursements either through selective admission (Gandhi, 2023) or discharge (Hackmann et al., 2024).

2.2. *Skilled Nursing Facilities*

Skilled nursing facilities (SNFs) provide nursing and rehabilitation services following an inpatient hospital stay. Medicare may pay for as many as 100 days in a SNF. To qualify, the SNF admission must occur within 30 days of a related hospital stay lasting at least three days. Skilled nursing facilities are reimbursed for provided services by Medicare at a flat fee-for-service per diem rate which depends on the specified reimbursement scheme.

Prior to October 1, 2019, skilled nursing facilities were reimbursed according to the Resource Utilization Group IV (RUG-IV) payment schedule, which was largely based on weekly therapy minutes.[9] Given that therapy could generally be provided at a substantially lower costs, this structure created strong incentives to increase therapy, and the proportion of patients billed for Ultra-High therapy rose from 17% in 2006 to over 53.7% by 2019.[10]

CMS has described "thresholding" just above payment cutoffs as evidence of care decisions driven by financial considerations rather than resident need.[11] In discussing its rationale for change, CMS stated, "we believe it is important to remove, to the extent possible, service-based metrics from the SNF PPS and derive payment from verifiable resident characteristics that are patient, and not facility, centered."

[9]The five therapy classification levels were Low (45-149 minutes), Medium (150-324 minutes), High (325-499 minutes), Very High (500-719 minutes), and Ultra-High (720+ minutes). In 2018, daily reimbursement rates for patients at a medium level of therapy ranged from $320.28-$616.13 per day depending on the services provided, but a patient receiving Ultra-High therapy would be eligible for reimbursements of $527.97—$832.89 per day (a $210 approximate difference per day). (https://www.govinfo.gov/content/pkg/FR-2018-08-08/pdf/2018-16570.pdf)

[10]https://oig.hhs.gov/oei/reports/oei-02-09-00202.pdf Substantial concerns were raised that RUG-IV reimbursement resulted in patients receiving unnecessary care and being "rehabbed to death" (Flint et al., 2019).

[11]https://www.govinfo.gov/content/pkg/FR-2018-05-08/html/2018-09015.htm and that the amount of therapy just over 720 minutes, "is too significant to be an accurate reflection of. . . population individualized needs." https://www.cms.gov/Outreach-and-Education/Outreach/NPC/Downloads/2018-12-11-PPS-Transcript.pdf

2.3. *Patient Driven Payment Model (PDPM)*

To reduce perverse incentives and medically unnecessary care, CMS replaced RUG-IV with the Patient Driven Payment Model (PDPM) on October 1, 2019.[12] PDPM bases reimbursement on five components: Nursing, Physical Therapy ("PT"), Occupational Therapy ("OT"), Speech-Language Pathology ("SLP"), and Non-Therapy Ancillary Services ("NTA"). Each component has case-mix groups determined by a patient's primary diagnosis, physical or cognitive status, and secondary conditions such as depression or other comorbidities as documented by the skilled nursing facility. Case-mixes requiring greater care receive higher reimbursement, potentially creating incentives to exaggerate patient conditions. Because reimbursement rates typically vary the most within the Nursing, SLP, and NTA components, our analysis largely focuses on these categories.

2.3.1. *Reimbursement Subcategories*

The nursing component reimbursement under PDPM is based on a patient's diagnosis, service intensity, physical function score, and depression status as shown in Exhibit 1.[13] In 2022, daily reimbursement for the nursing component ranged from $68.28 to $420.05. Within the nursing component, we focus on three categorizations that substantially increase SNF reimbursement—Special Care High, Depression, and Low Function. First, to qualify for Special Care High, a patient needs a physical function score below 14 and must suffer from a serious medical condition (such as septicemia, daily respiratory therapy, comatose, or fever with additional symptoms). Second, skilled nursing staff screen patients for depression. Finally, patients who can complete fewer tasks on their own are eligible for higher reimbursement. We refer to patients billed for the lowest category of

[12] https://www.cms.gov/medicare/medicare-fee-for-service-payment/snfpps/downloads/pdpm_faq_final_v5.zip

[13] Additional details can be found in Appendix 8.1.

physical function within each subgroup as "Low Function."

The Speech Language Pathology (SLP) component includes 12 case-mix indices, with daily reimbursement rates ranging from \$15.06 to \$93.25, as shown in Exhibit 2. Patients are screened for an Acute Neurologic (AN) primary diagnosis, additional SLP-related comorbidities, and cognitive impairment, with reimbursement increasing in the count of such conditions. Claims with at least two of the three conditions are classified as "SLP High". Patients are then assessed for requirement of a mechanically altered diet or swallowing disorders, and those with both conditions receive the highest reimbursement.[14]

We define the sum of these three binary nursing categories (Special Care High, Depression, and Low Function) and two binary SLP categories (SLP High and Dietary Restriction) as *Coding Intensity* and this forms our primary measure of PDPM billing intensity.[15]

Finally, the Non-Therapy Ancillary (NTA) component is designed to reimburse for patient conditions that generate costs not adequately compensated by the Nursing component. The NTA component features six possible case-mixes with daily reimbursement ranging from \$60.82 to \$273.67. The case-mix for this component is determined by the presence of 50 conditions and extensive services associated with increased non-therapy costs. NTA scores consider both the count and severity of various comorbidities. We examine the prevalence of diagnoses that qualify for additional NTA reimbursement.

[14] SLP case-mixes are described in further detail in Appendix 8.2.

[15] Additionally, in some tests, we examine two subcomponents of the SLP component (Acute Neurologic and SLP-Related Diagnoses), which have the advantage of being trackable prior to PDPM.

3. Data and Summary Statistics

The first primary dataset for this study is the Skilled Nursing Facility Limited Data Set, covering the universe of Medicare Skilled Nursing claims from January 1, 2016, to December 31, 2023. The second major dataset used is the Inpatient Limited Data Set, which contains inpatient hospital visits, diagnoses, and procedures. Both are compiled and distributed by CMS and share an anonymized beneficiary ID, allowing linkage of health history prior to an SNF stay. Key variables in the skilled nursing data include the patient's primary diagnosis and up to 24 secondary diagnoses, therapy level (RUG-IV) or case-mix (PDPM), age in five-year increments, race, gender, county, and number of treatment days. The inpatient data contains the same patient characteristics, as well as hospital-recorded diagnoses and procedures. The data covers a total of 15,911,862 skilled nursing stays from 7,962,416 unique patients.[16]

We also utilize several other data sources. Skilled Nursing Facility Enrollments from CMS Public Use Files (PUF) are used to classify facilities into systems. Facilities are considered part of the same system if they share an Affiliation ID. 9,682 of the 16,744 (58.9%) facilities, which account for 66.0% of patients, are identified as belonging to an SNF system. CMS affiliated entity status may not reflect every possible SNF system due to the complex ownership structures and reliance on accurate submission of ownership data.[17] Facilities without a CMS-reported affiliation are considered independent. The CMS Change of Ownership (CHOW) database identifies acquisitions. Additional facility-level information, including occupancy, staffing, and the proportion of assessments billed just above the Very-High (500-509 minutes) and Ultra-High (720-729 minutes) thresholds comes from

[16]Our dataset excludes long-stay residents which are typically paid for by Medicaid, and account for approximately 60% of all nursing home patient-days (Gupta et al., 2023).

[17]For technical methodology about how CMS identifies SNF affiliations, see https://data.cms.gov/resources/nursing-home-affiliated-entity-performance-measures-methodology.

facility-level CMS PUFs. Clinical Classification Software Refined (CCSR) from the Healthcare Cost and Utilization Project (HCUP) is used to classify similar diagnosis and procedure codes.

3.1. The Geography of Medicare Billing and Facility-Level Practices

Do billing practices of skilled nursing facilities vary across geographic markets? We first estimate a patient-level regression of Coding Intensity on basic demographics (gender, age, race) and health status (inpatient hospital diagnosis or time spent in hospital) to control for patient characteristics that might be correlated with coding intensity:

$$CodingIntensity_{ijt} = \alpha + \delta_t + \theta X_{it} + \epsilon_{ijt} \tag{1}$$

To examine the geography of billing practices, we aggregate residuals to the county level. If billing intensity beyond what would be expected given observable patient characteristics were largely idiosyncratic, then there might be no clear geographic pattern. Instead, billing practices vary substantially as shown in Figure 1. For example, the average patient in Brooklyn, New York is coded as having 0.95 more highly compensating codes than predicted, which is large relative to the unconditional mean of 1.04.[18] The geographic variation in Medicare billing is consistent with large differences in Medicare spending across markets (Newhouse et al., 2013), the majority of which is driven by post-acute care settings such as SNFs (Einav et al., 2025).

We also explore within-county variation in practices by aggregating residuals fromEquation 1 at the individual facility level. Figure 1, Panel B sorts counties on their level of coding intensity and shows that there is substantial within-county variation in coding practices. In fact, this variation is considerably larger than county-level variation in average practices. Moving from the 10th to the 90th percentile county shifts coding from –0.31 to 0.15, while the equivalent facility-level shift

[18] The specification uses the diagnosis from the immediately preceding hospital stay since SNFs may face incentives to manipulate reported diagnoses.

within a county ranges from –0.44 to 0.33. Facility-level variation does not simply reflect noise. At the facility level, coding intensity from October 1, 2019–December 31, 2021, is highly correlated (0.85) with coding intensity from January 1, 2022–December 31, 2023.[19] The wide variation in coding intensity across facilities provides a rich testing ground for potential explanations.

4. What Explains Variation in Coding Intensity?

Facilities might experience higher levels of coding intensity for several reasons. First, skilled nursing facilities due to reasons such as geographic location,[20] physical facility specification, or staff training may specialize in treating patients with higher compensating comorbidities. Second, higher levels of coding intensity may be observed if facilities can selectively admit patients with greater comorbidities. Third, certain facilities may code more completely, and the PDPM reimbursement design disproportionately rewards such thoroughness. Finally, facilities may overstate patient conditions to maximize revenue.

We now examine a series of tests to distinguish between these hypotheses. First, we examine the relation between RUG-IV and PDPM billing practices at the facility level. Second, we compare the prevalence of comorbidities across facilities prior to PDPM, using hospital diagnoses as an independent measure of patient conditions. Third, we utilize competitor occupancy constraints to generate quasi-random patient assignment in an instrumental variables framework. Fourth, we compare diagnosis rates for compensating versus non-compensating patient conditions. Fifth, we examine the diffusion of coding practices across facilities within the same ownership group. Finally, we analyze changes in coding intensity following facility acquisitions.

[19]Figure IA.12

[20]For example, a facility may be physically located next to hospitals with populations requiring more intensive care.

4.1. *Facility-Level Practices from RUG-IV to PDPM*

Under RUG-IV, reimbursement favored healthier patients with fewer comorbidities who could tolerate intensive therapy, whereas PDPM shifted payment toward sicker patients with clinically complex diagnoses. The change was intended to reallocate reimbursement from patients receiving the highest levels of therapy towards patients with more complex clinical needs. CMS and industry participants projected large revenue decreases for facilities with high shares of Ultra-High rehab residents.[21]

To assess the importance of individual facilities in explaining billing variation, we estimate patient-level regressions of Ultra-High Rehab (RUG-IV) and PDPM Coding Intensity on patient demographics (gender, age, race), health measures (inpatient diagnosis[22] or hospital days), geographic and time fixed effects using the following specification:

$$y_{ijt} = \alpha + \theta X_{it} + \delta_t + \gamma_j + \epsilon_{ijt} \tag{2}$$

We include a facility fixed effect, γ_j to measure facility-specific Ultra-High Rehab (RUG-IV) or Coding Intensity (PDPM) beyond what would be expected given patient characteristics. Similar to PDPM coding intensity, rehab provision under RUG-IV differs widely across facilities with estimated facility fixed effects, which we refer to as "Excess Rehab", ranging from –16.18 days to 31.81 days.[23]

We report the incremental R-squares from excluding each variable in Figure 2, Panel A. Patient demographics, geography, and time contribute little to billing variation, while diagnosis plays a

[21] CMS projected revenue reductions of 8.4% for residents receiving the highest therapy levels and average increases of 50.5% for patients receiving no therapy (https://www.govinfo.gov/content/pkg/FR-2018-05-08/pdf/2018-09015.pdf). Industry participants similarly anticipated revenue declines under PDPM for therapy-intensive facilities, given that such patients typically had fewer comorbidities (https://www.monterotherapyservices.com/articles/coming-soon-to-a-snf-near-you-new-payment-model/).

[22] Inpatient hospital rather than SNF diagnoses are used since SNF diagnoses directly influence SNF reimbursement.

[23] For reference, the median (mean) number of days of Ultra-High rehab a patient receives during RUG-IV is 12 (15.61) days, as shown in Table 1.

larger role. Facility fixed effects explain the most variation with incremental R-squares of 0.093 (RUG-IV) and 0.091 (PDPM). Strikingly, even under PDPM—which explicitly ties reimbursement to comorbidities—facility fixed effects explain more variation in billing than a patient's hospital diagnosis.

If specialization in providing intensive therapy were the main driver of the facility fixed effect observed in Figure 2, then facility-specific billing during RUG-IV might be largely uncorrelated, or even negatively correlated with billing during PDPM. However, facility-specific RUG-IV rehab and PDPM coding intensity are strongly positively correlated (0.58) (as shown in Figure 1, Panel A), with a one standard deviation increase in rehab associated with a 24% increase in PDPM coding intensity as shown in Table 2.[24] This strong and positive correlation contrasts with CMS and industry projections. As we will show below, most of the Ultra-high rehab billings for facilities in the top tercile of rehab levels bunches in ten-minute payment window just above the threshold. This type of targeting above thresholds was also part of the stated rationale for CMS moving away from the RUG-IV revenue system and appears more consistent with providing rehab to maximize revenue rather than customized rehab levels based on patient need. In subsequent analysis of PDPM, we classify facilities by their RUG-IV billing practices so that facilities are ranked based on prior practices in a different billing regime rather than contemporaneous PDPM practices. Nevertheless, we recognize that facilities change their practices over time.

4.2. *Differences in SNF and Hospital Coding of Patient Comorbidities*

Are high levels of coding intensity from skilled nursing facilities with high previous levels of rehab due to either specialization in treating or selection of patients with greater comorbidities?

[24] A positive association is found between RUG-IV rehab levels and each individual PDPM billing category considered indicating that the relation between RUG-IV rehab and coding levels is not driven by a particular type of care (Table IA.4)

As discussed in Section 2.3.1, under PDPM three groups of diagnoses—Acute Neurologic (AN), SLP-Related Comorbidities, and NTA-Related Comorbidities—trigger higher payments to SNFs through the SLP and NTA components.[25] Although a broader set of codes can be used to increase reimbursement, we focus on AN, SLP-Related, and NTA-Related diagnoses in this section because they were available on patient claims prior to PDPM. Before conducting a formal difference-in-difference analysis coding around PDPM, we first explore whether high-rehab facilities treated sicker patients under RUG-IV. Prior to PDPM, the prevalence of AN diagnoses was nearly identical across facility types—8.01% at low-rehab facilities versus 8.04% at high-rehab facilities. However, after PDPM was introduced, the shares increase to 20.68% high-rehab facilities versus 13.44% at low-rehab facilities (as shown in Figure IA.5). Similar patterns are observed for SLP and NTA-related diagnoses.

We next utilize a more formal difference-in-differences approach to estimate the relative likelihood of compensating diagnoses at high rehab facilities versus other facilities over time. Figure 3 plots coefficients from a dynamic difference-in-differences regression of the form:

$$Comorb_{ijt} = \alpha + \sum_{t \neq Q4\ 2018} \beta_t Quarter_t \times HighRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt} \tag{3}$$

where $Comorb_{ijt}$ is an indicator variable equal to one if patient i had a compensating diagnosis (AN, SLP, or NTA) recorded at facility j in quarter t. $HighRehab_j$ is an indicator equal to one if a facility was classified in the top tercile of RUG-IV rehab and zero otherwise.[26] Despite representing a wide range of unrelated conditions, all three diagnoses (AN, SLP, and NTA) sharply increase at high-

[25] A full list of these diagnoses and their mapping to billing is provided in the Internet Appendix. To qualify for additional compensation, an Acute Neurologic condition must be a patient's primary condition.

[26] We include a quarter × year fixed effect, δ_t, to control for aggregate time-series variation in comorbidity diagnoses. Q4 2018, the last quarter before the announcement of PDPM, serves as the reference and is omitted in the estimation. Facility fixed effects are included to control for time-invariant facility-level heterogeneity in the likelihood of compensating diagnoses.

rehab skilled nursing facilities following the implementation of PDPM and remain persistently high as shown by the blue line. Table 3 presents regression coefficient estimates with varying fixed effects that shows the increases in coding for formerly high-rehab facilities are both economically large and precisely estimated. Acute Neurologic diagnoses increase by 5.3 percentage points at facilities which previously had high levels of rehab, an increase of about 60% relative to the pre-PDPM mean.

We further investigate the cross-sectional relationship between a facility's billing practices during RUG-IV and each compensating diagnoses (in Figure IA.6) by plotting a binscatter of each compensating diagnoses against a facility's excess rehab separately for the RUG-IV and PDPM eras.[27] Facilities with high RUG-IV rehab demonstrate much higher levels of each of the compensating comorbidities, but only after PDPM is enacted.

4.2.1. Are Higher Diagnostic Rates Driven by Patient Selection?

Are the higher levels of compensating diagnoses at high-rehab facilities driven by selective admission policies? Although these facilities were no more likely to treat patients with AN, SLP or NTA-related conditions prior to PDPM, they may have increased admissions of these patients through targeted policies. To test this, we use an external measure of patient conditions by linking SNF stays to referring hospital records. If the rise in comorbidities at high-rehab facilities reflects selective admission rather than upcoding, then similar increases should be observed in hospital diagnoses immediately prior to the SNF stay.

We apply the difference-in-differences design from Equation 3 but use referring hospital diagnoses as the dependent variable. Figure 3 plots the estimated coefficients when qualifying diagnoses are measured using hospital data in orange. Unlike the diagnoses recorded by the SNF, there is no

[27] A facility's excess rehab is the facility fixed effect estimated in Equation 11, with the outcome variable being days of Ultra-High Rehab.

discernible increase in any of the compensating diagnoses (AN, SLP, or NTA) for patients admitted to high-rehab facilities following the adoption of PDPM. Increased prevalence of these diagnoses at high-rehab facilities comes from conditions coded at the SNF but not documented by the referring hospital.

4.3. Quasi-Random Patient Assignment

So far, we have examined the relationship between intensive rehab provision under RUG-IV and coding intensity under PDPM. A remaining concern is that unobserved differences in patient health might drive selection into these facilities. To address this, we construct an instrument that exploits variation in local facility occupancy as a source of quasi-random assignment. Occupancy constraints are common in skilled nursing—55% of facilities report sometimes turning away prospective patients[28]—and are amplified by strong geographic preferences, with patients typically entering facilities within five miles of home (Gupta et al., 2023).

We construct our instrument using occupancy levels of competing facilities within the same hospital service area (HSA).[29] The intuition of the instrument is that a patient is more likely to visit a given facility when competing facilities are close to maximum occupancy. This strategy builds on discrete-choice models of facility selection within local geographic markets used in setting such as hospitals (Geweke et al., 2003) or SNFs (Einav et al., 2025; Olenski and Sacher, 2024).[30] This approach yields a measure of facility rehab intensity that is plausibly unrelated to underlying patient health, which allows us to isolate the effect of facility practices from patient selection.

[28] https://www.mcknights.com/news/ahca-offers-wake-up-call-on-bed-and-facility-counts-446000-residents-may-be-displaced/

[29] The restriction to facilities within the same HSA is due to patients generally considering facilities within a close geographic proximity.

[30] Utilization of geographical preferences more generally have been used to instrument for facility selection in settings such as hospitals (McClellan et al., 1994; Card et al., 2023), nursing homes (Grabowski et al., 2013) and dialysis care centers (Wang et al., 2017).

To first explore the mechanism of occupancy constraints and to guide our empirical specification, we first examine the likelihood that a patient i is admitted to a particular facility j as a function of occupancy at other facilities ($j \neq k$) in the same HSA. Using a flexible local polynomial specification (Figure IA.14), we find that competitor occupancy strongly affects patient selection, but only when competing facilities are near full capacity. To construct our instrument, we first model facility selection by using a piecewise-linear OLS specification of the form:

$$\pi_{ijt} = \alpha + \beta_1 Occ_{j\neq k,t-1} + \beta_2 Constrained_{j,t-1} + \beta_3 Constrained_{j,t-1} \times Occ_{j\neq k,t-1} + \epsilon_{ijt} \quad (4)$$

where π_{ijt} is the probability that patient i visits facility j in month t. $Occ_{j\neq k,t-1}$ is the average occupancy at competing facilities in the same HSA, and $Constrained_{j,t-1}$ is an indicator variable equal to one if the average level of occupancy across competitors exceeded 93% in the prior month.[31]

We then scale the fitted probabilities from Equation 4 to ensure that the probability that a patient chooses a facility within the HSA sums to one. These scaled fitted probabilities $\widehat{\pi_{ijt}}$ are used as probability weights to construct our instrument, defined as a weighted average of past facility RUG-IV rehab levels according to:

$$ExpectedFacilityRehab_{mt} = \sum_{j=1}^{J} \widehat{\pi_{ijt}} \times RUGRehab_j \quad (5)$$

where $RUGRehab_j$ is the past facility-level excess rehab during RUG-IV period estimated according to Equation 11. In this case, the past excess rehab practices in the RUG-IV period proxies for the past facility billing behavior, but by instrumenting for these practices, we identify the level of facility practices that a patient is expected based on local occupancy-driven variation in facility assignment. The key identification assumption is that variation in facility choice induced by competitor occupancy is uncorrelated with unobserved patient health. Occupancy constraints among nearby SNFs are

[31] This 93% threshold choice is motivated by Figure IA.14, which shows occupancy constraints generally begin to bind around this level. Results are robust to alternative cutoffs of 90% or 95% (as shown in Tables IA.5 through IA.8).

plausibly exogenous to the health of any individual patient, and we provide indirect evidence of this using hospital diagnoses immediately prior to SNF admission. To account for geographic variation clustering of high-rehab facilities, we include HSA fixed effects in preferred specifications to ensure that variation in the expected facility rehab level is identified using only time-series changes in the relative occupancy within an HSA.

Finally, for the instrument to be valid, it must strongly predict the RUG-IV rehab practices of the facility where patients are admitted. This requires that the instrument generate meaningful variation in facility placement and that facilities exhibited sufficiently differing rehab practices during RUG-IV. We verify that the relevance condition holds in this setting by estimating a first-stage regression. The constructed instrument strongly predicts the observed facility rehab level (t statistic=7.24, F statistic=52.36).

We first visually examine the relationship between facility-level RUG-IV rehab and PDPM coding intensity by estimating a reduced-form specification in which we directly regress each PDPM coding intensity measure on the instrument according to:

$$y_{ijt} = \alpha + \beta \times ExpectedFacilityRehab_{mt} + \theta X_{it} + \delta_t + \epsilon_{ijt} \tag{6}$$

Where y_{ijt} is a billing condition that patient i receives at facility j in quarter t as a function of the expected excess rehab in market m at time t. Importantly, the constructed instrument is not a function of the specific facility j that a patient selects, but rather a function of the expected excess rehab practices that a patient might face based on the occupancy rates of facilities within a given HSA-month. We visualize this relationship in Figure 4. There is a strongly positive and statistically significant relationship between our instrument and each of the billing codes.

Building on this, we estimate a 2SLS framework to examine the relationship between facility RUG-IV rehab and PDPM coding intensity (presented in Table 4). The coefficient of 0.042 implies that a shift from the first to third tercile of excess rehab (10.5 days) predicts increased coding intensity of 0.441, or about 42% of the unconditional mean. Estimated coefficients are slightly larger than the OLS estimate of 0.0296 (Table 2), suggesting that selective patient admission is not leading to large discrepancies in billing levels.

For the instrument to have a causal interpretation, it must affect patient coding only through facility choice. This assumption might be threatened if time-varying, location-specific shocks correlate with the distribution of facility occupancy and patient health. Although this identification assumption cannot be directly tested, we conduct a variety of balance tests. We repeat the primary 2SLS regression but utilize the presence of compensating diagnoses at the preceding hospital stay as our primary measure of external patient health conditions. The instrument does not predict an increased likelihood of a patient receiving an Acute Neurologic, SLP-Related, or NTA-Related comorbidity documented at the hospital. Furthermore, the instrument is uncorrelated with patient gender, age or severity as proxied for by the length of hospitalization. For identification to be threatened, patient health would need to be systematically correlated only among unobservable and latent patient conditions, but not along a wide variety of observable patient health outcomes.

4.4. Can Higher Coding Rates Reflect More Thorough Coding Practices?

Higher coding intensity does not appear to be explained by selection on observable or unobservable characteristics. A remaining concern is that high levels of coding intensity could reflect more thorough coding by high-rehab facilities. Our earlier findings indicate that these high-rehab facilities would need to considerably improve coding practices precisely at the time of PDPM adoption for this explanation to be plausible.

If high-rehab facilities invested in diagnostic tools, they may record more thorough diagnoses after PDPM. To test this, we compare diagnosis rates separately for compensating and non-compensating diagnoses. Compensating diagnoses are those within the Acute Neurologic, SLP-Related, or NTA-related groups and increase revenue for the SNF. Non-compensating diagnoses, by contrast, are diagnoses that do not directly influence reimbursements for most claims.[32] The majority of potential diagnoses (1206 out of 1348) are non-compensating, while approximately 10.5% are compensating. Diagnoses naturally vary in prevalence, so we scale the likelihood of each by the national average, which we define as the diagnosis ratio.[33] A diagnosis ratio of more (less) than one means that a facility is more (less) likely to code a given condition than the national average. Are the higher rates of coding for compensating diagnoses at high-rehab facilities also observed for other diagnoses?

For the 10.5% of diagnoses classified as compensating, high-rehab facilities code at rates 55% above the national average, while for the remaining 89.5% of non-compensating diagnoses, they are 5% less likely to code a condition as shown in Panel A of Figure 5. Is the concentration of coding on a small subset of diagnoses likely to occur by chance? To test, we run a bootstrap simulation randomly assigning 10.5% of all diagnoses to be compensating and compute the average diagnostic ratio. Across 10,000 simulations, no difference approaches the magnitude observed in the data: the maximum simulated aggregate ratio is 1.22, compared with 1.55 observed in the data as shown in Panel B of Figure 5. This evidence points to deliberate upcoding of compensating diagnoses rather than a more general improvement in coding thoroughness.

[32]Due to complexities in the case-mix design, it is possible that some non-compensating codes may increase reimbursement in specific instances.

[33]We restrict attention to diagnoses that are recorded on at least 5,000 claims to ensure that results are not driven by outliers.

4.5. *Discussion*

So far, we have found that facilities with the highest rehab under RUG-IV begin coding comorbidities at higher rates under PDPM, but only for a narrow set of high-compensation diagnoses and without evidence of corresponding conditions in the preceding hospital stay. With little support for explanations based on patient composition, selection, or improved diagnostic practices, the evidence points to upcoding. We now turn to potential explanations for this behavior and consider the mechanisms that could explain these practices.

5. What are the Main Drivers of Fraud?

Why do some providers commit much higher levels of upcoding than others? We first consider whether cross-sectional differences in fraud primarily arise from incentive misalignment due to imperfect reimbursement. We then examine the role of twenty potential determinants including differences in enforcement intensity, competition, firm sophistication, demographics, corruption, religion, social capital, local social norms, and firm norms.

5.1. *Reimbursement Regime*

Healthcare fraud is thought to primarily arise due to a combination of imperfect reimbursement and inadequate penalties of fraud (Leder-Luis and Malani, 2025). Under PDPM, we focus on upcoding, which may be exacerbated by costly verification. To test the role of the reimbursement design in deterring fraud, we examine facility practices under the therapy-driven RUG-IV era, when payments were tied to a verifiable input.

There is a broad consensus that the marginal revenue from providing additional therapy minutes exceeded the cost, raising concerns among both CMS and industry participants about excessive rehab provision.[34] Excessive rehab provision constitutes medical necessity fraud, which can be deterred by reducing reimbursement for high therapy levels. If facility-level differences in cost structures alone account for differences in therapy utilization, there would be no reason to expect high-rehab facilities to also upcode under PDPM. However, if firm culture and tolerance are the driving factors, then the same facilities might engage in fraudulent behavior under both reimbursement designs. As previously discussed in Section 4, there were large differences in rehab provision across facilities that were strongly correlated with PDPM upcoding. Do these much higher levels of rehab therapy reflect medically necessary care or opportunistic behavior? To address this challenge, we exploit a unique feature of RUG-IV reimbursement: discrete payment thresholds. Medicare payments increase sharply at arbitrary cutoffs, though clinical need should not. Evidence of bunching just above cutoff thresholds points toward a provision of therapy driven by financial incentives rather than medical need—a concern also raised by CMS and OIG.

For facilities in the top decile of RUG-IV rehab, 49.9% of all days are billed within 720-729 minutes versus only 8.5% of days for facilities in the bottom decile (as shown in Figure IA.2). Patients at facilities in the highest decile of rehab therapy are 6.65 times more likely to receive between 720-729 minutes of weekly therapy than any time in the 210-minute window below.[35] Overall, most variation in facility rehab fixed effects arises from billing clustered just above the 720-cutoff.[36]

[34]https://www.wsj.com/articles/how-medicare-rewards-copious-nursing-home-therapy-1439778701.

[35]By comparison, patients at facilities in the lowest decile of ultra-high rehab were 37% less likely to receive therapy in the 720-729 minute threshold than the 210-minute range below. This does not appear to be driven by "round-hour" prescription heuristics—the 210-minute window immediately below this threshold contains alternative "round-hour" prescriptions for nine, ten, or eleven hours as well as any times between. Additionally, billing just above the Very High Rehab threshold (500-509 minutes) is strongly correlated with Ultra-High thresholding with a correlation coefficient of 0.812 (as shown in Figure IA.1, Panel B).

[36]The correlation between system excess rehab and the proportion of patient days billed between 720-729 minutes is 0.5457, whereas the correlation between system excess rehab and proportion of patient days billed any number of minutes above 730 is 0.1326 as shown in Panel A of Figure IA.1.

To further assess the medical necessity of high levels of rehab, we examine cases where additional rehab therapy is unlikely to be beneficial—specifically during the final days of a patient's life, when most patients are physically unable to perform such activities (Temkin-Greener et al., 2019). Among patients who die in the highest-rehab facilities, 45% received the most intensive Ultra-High rehab, over 100 minutes per day, up to the day of death.[37]

As a final test of the relationship between RUG-IV rehab and PDPM upcoding, we analyze patients within SNFs at the time of reimbursement change. Specifically, we examine instances in which patients who were previously receiving the highest levels of Ultra-High rehab under RUG-IV are billed as having either Low Function or Special Care High under PDPM,[38] categories that require impairment levels unlikely to be compatible with twelve hours of physical therapy per week. Yet, among the highest-rehab facilities, 42% of patients receiving the most intensive therapy on September 30, 2019 were reclassified the very next day as having these high impairment conditions under PDPM.[39] This drastic and systematic shift for many patients is difficult to reconcile with medical necessity and accurate reporting but is consistent with medical necessity fraud and upcoding.

5.1.1. Billing Practices and the Cost of Care?

Although high-rehab utilization facilities upcode at greater rates, was PDPM nonetheless effective at reducing fiscal waste? PDPM was intended to be budget neutral and allocate reimbursements away from facilities with high RUG-IV rehab levels. The average cost of care increased from $15,296 under RUG-IV to $16,483 under PDPM (as shown in Figure IA.8. Moreover, facilities that generated

[37] As shown in Figure IA.9. For comparison, only 8% of patients in the lowest-rehab facilities received such intensive therapy.

[38] Low Function patients fall in the lowest category of physical function, generally unable to perform mobility or self-care tasks on their own. Special Care High includes patients with reduced function and serious comorbidities such as septicemia, daily respiratory therapy, coma, or fever with additional complications.

[39] This occurred only 9% of the time for patients in the lowest-rehab facilities as shown in Figure IA.15.

the highest revenues under RUG-IV continue to maintain the highest levels of revenue under PDPM (as shown in Figure IA.8). Differences are substantial with a shift from tenth-percentile to ninetieth-percentile of RUG-IV rehab resulting in an cost increasing from \$14,540 to \$24,971 under PDPM. Overall, the transition appears ineffective at redirecting Medicare resources away from facilities exhibiting unusually high levels of RUG-IV rehab.

5.2. Does Enforcement Regime, Firm Sophistication, Local Culture, or Demographics Explain Fraud?

In total, we consider a set of twenty explanatory variables of which six variables relate to market-level or facility skilled nursing conditions, six relate to county-level demographics, three relate to county-level religion morality, or corruption, and five relate to social capital.[40] We first use OLS as an exploratory tool to highlight correlates of fraud, recognizing that omitted variables and reverse causality preclude causal interpretation. Specifically, we estimate a facility-level OLS regression of the form:

$$ResidCodeIntens_{jm} = \alpha + \theta X_{jm} + \epsilon_{jm} \tag{7}$$

Where $ResidCodeIntens_{jm}$ is the residualized facility-level coding intensity after controlling for observable patient conditions according to the facility fixed effect from Equation 11. The determinants of interest, X_{jm} are facility or market-level variables that might be correlated with fraudulent behavior, motivated by theoretical considerations we previously discussed. Regressions of facility upcoding on all candidate explanatory variables are presented in Table 5. For interpretability, all explanatory variables and the outcome variable are standardized to have mean zero and a standard

[40]Variables specifically related to skilled nursing include Medicare enforcement, competition, measures of facility or system size, billing practices of other facilities within a county, and finally billing practices of other facilities within a system. Demographic variables include the log population density, log median income, unemployment, poverty rate, percent college educated, and percent non-white. County-level social norms and culture related to religious affiliation, Ashley Madison usage, and public corruption. Measure of social capital come from Chetty et al. (2022) and include economic connectedness, clustering, support ratio, volunteer rates, and civic organization participation rates.

deviation of one.

5.3. *Enforcement*

In a standard Becker crime model, fraud occurs when expected benefits exceed expected costs, implying that weaker enforcement environments may exhibit higher levels of fraud. Enforcement practices may vary widely across DOJ offices depending on the aggressiveness of targeting medical fraud as well as staffing resources.[41] To test, we construct a geographic measure of Medicare enforcement at the U.S. District Attorney Office level using the universe of DOJ press releases following Leder-Luis and Malani (2025), but restrict focus to Medicare-related cases. Medicare enforcement rates do vary substantially across regions,[42] but there is no significant relationship between enforcement activity and SNF fraud as shown in Columns 1 and 6-10 in Table 5. Standard errors are small enough to rule out large effects.[43] Our findings suggest that the current levels of enforcement are probably too low to be a driving cross-sectional determinant in fraud as we further in Section 5.8.

5.4. *Competition and Sophistication*

Does market competition encourage or deter fraud? If fraud leads to deviations from patient-optimal care (e.g., excessive rehab) and facilities compete on quality, competition could improve practices and reduce fraud. Standard models of regulated markets with endogenous quality typically predict that competition enhances care quality (Gaynor et al., 2007). Conversely, competition may increase fraud if legitimate operators cannot remain profitable (Diwan et al., 2025). Estimating Equation 7, Table IA.3, Columns 2, 6-10 (and Figure IA.10) show that the coefficient estimate

[41] For example, certain DOJ offices like SDNY have developed a reputation of hiring top lawyers that wish to aggressively prosecuting fraud. https://www.businessinsider.com/why-the-southern-district-of-new-york-is-so-prestigious-2013-11

[42] Substantial geographic variation in Medicare enforcement actions per 100,000 is shown in Figure IA.7.

[43] Figure IA.10 plots regression coefficients and 95% confidence intervals for each SNF-specific explanatory variable.

on HHI is close to zero and statistically insignificant, suggesting that local SNF competition is unrelated to fraud. One potential explanation is that healthcare is a credence good, where patients face difficulties in perceiving care quality or may be entirely unaware of Medicare fraud.

A prerequisite for fraud is that facilities must possess the knowledge and expertise to carry it out. The PDPM reimbursement design was far more complex than RUG-IV with more than 14,000 valid case-mix combinations compared to only 66 under RUG-IV. Larger firms may better exploit such complexity through more sophisticated billing technologies, shared cost structures, economies of scale, or accumulated experience.[44] Facility size is generally positively correlated with coding, but neither facility nor firm size explain much of the variation in fraud rates as shown in Table IA.3, Columns 3, 6-10 (and Figure IA.10).

5.5. *Local Billing Practices Across SNF Systems*

Prior work has suggested that questionable practices can spread geographically, potentially through learning or norm-shifting (Chetty et al., 2013; O'Malley et al., 2021). To avoid confounding from common firm strategies, we compute the average coding intensity of all other facilities within a county not owned by the same parent firm according to:

$$CountyAverage_{jm} = \frac{1}{M} \sum_{m,j \notin J}^{M} CodingIntensity_j \tag{8}$$

A facility's fraud levels are positively associated with the fraudulent practices of neighboring facilities in the same county owned by other SNF systems after controlling for other skilled nursing controls or demographic variables.[45]

[44] For example, larger dialysis chains sustained greater profits through billing strategies not easily replicated by independent providers (Eliason et al., 2020).

[45] A one standard deviation increase in fraudulent practices for competitors is associated with a 9% standard deviation increase in a facility's own fraud, as shown in Table 5, Column 10.

5.6. *System Ownership Practices*

We next examine the role of ownership on billing practices. Many SNFs are owned by management groups that operate multiple facilities, but ownership links were historically difficult to establish.[46] However, in 2022 CMS began releasing data on SNF ownership, which we use to define all facilities operated by the same ultimate management group as belonging to an SNF "system" utilizing CMS-published affiliations.[47] We construct a variable equal to the facility leave-out-mean of unexplained Coding Intensity for all facilities belonging to the same system:

$$\mu_{j,s} = \frac{1}{N-1} \sum_{i \neq j} CodingIntens_{is}$$

We then regress facility-level practices on this system-level variable and find that average practices within a system are by far the strongest predictor of facility behavior. Including this variable alone increases the R-square by more than 30% after controlling for state fixed effects—an order of magnitude larger than previously considered determinants (as shown in Table 5, Columns 5-10 and Figure IA.10). In the most saturated specification, a one standard deviation increase in system-level billing intensity is associated with a 0.713 standard deviation increase in facility billing, roughly seven times the effect of any of the previously discussed SNF-specific variables.

5.6.1. *Geographic Culture, Demographics, and Social Capital*

Does skilled nursing fraud simply reflect other known correlates of crime? Prior studies have related fraud to religious affiliation, public corruption, and marital infidelity (Parsons et al., 2018; Grullon et al., 2009; Griffin et al., 2019). We examine whether these and typical crime correlates (demographics, education) or measures of social capital (Chetty et al., 2022) explain SNF fraud

[46] Relying on facility name alone misses many possible links since ownership groups often lack common branding.

[47] All ownership classifications—including CMS's—remain imperfect due to complex ownership structures, imperfect and incomplete ownership data and facilities with only partial overlapping ownership (Chen et al., 2024).

in Table IA.2. We add each variable category to the specification independently (Columns 7-9) and jointly (Column 10). Billing intensity levels are higher in counties that have greater population density, higher white population, and lower college education. Fraud is also slightly higher in counties with higher levels of public corruption and greater social clustering, but lower in areas with higher support ratios and civic organization participation. While many demographic and social capital measures are correlated with facility-level billing practices, they explain very little of the overall variation with an incremental R-squared of just 0.003 after controlling for SNF-specific variables.

5.7. Machine Learning Approach

OLS estimates may overlook important nonlinear relationships, including interactions among variables. To allow for such complexities, we employ gradient boosting decision trees (GBDTs), a machine learning algorithm well-suited for modeling nonlinearities and high-dimensional interactions. To assess the contribution of each factor within this framework, we utilize SHAP (Shapley Additive exPlanations) values. SHAP values provide a consistent measure of feature importance, quantifying how much each variable contributes to the model's predicted outcome. We standardize the outcome and all covariates.

Figure 6 Panel A shows that firm-specific variables remain the most influential in explaining variation in facility upcoding. The mean absolute SHAP value for upcoding at other facilities in the system is 0.544. This means that upcoding at other facilities moves the model expected coding intensity by 0.54 standard deviations on average, just over eight times more influential than next most important determinant, county-average level of coding. Measures of organizational sophistication, such as facility and system size, also contribute meaningfully, whereas demographic characteristics, broader skilled-nursing market variables, and proxies for social capital or morality play a substantially smaller role.

Figure 6 Panel B plots SHAP values at the facility level for the four most influential covariates. The first plot shows the effect of coding at other facilities within the system. The density axis demonstrates that that most of the observations are in SNF systems with low levels of upcoding. Some facilities have SHAP values of 3.0 for other facilities within the system, indicating extremely strong effects on within-system coding practices. For county-average coding, most effects are concentrated at extremes and within-county dispersion remains large. Overall, the machine learning results reinforce OLS findings: firm norms are the most important determinants of fraudulent billing behavior.

5.8. Interpretation

Some of our findings are initially surprising. Does the fact that the same SNF systems that delivered excessive levels of rehab begin upcoding under the new PDPM system indicate that billing regimes are irrelevant? There are at least two reasons that we do not believe this to be the case—first we study only one potential reimbursement change and second, the type of fraud changed from providing medically unnecessary services to upcoding. The health implications for patients from these two types of fraud may differ. It is possible that reimbursement regimes which are less easily gamed may be more effective at deterring fraud. Designing a system that rewards for patient comorbidities without having a rigorous verification mechanism did not lead to the intended outcome. Our findings highlight that policymakers should more carefully the gaming of future reimbursement mechanisms.

Similarly, our findings should not be interpreted as evidence that enforcement is irrelevant. Rather, the null results from empirical tests suggest that current enforcement patterns are not systematically related to observed cross-sectional variation in fraudulent practices. A reasonable interpretation is that the overall levels of enforcement may simply be too low to generate meaningful deterrent effects related to cross-sectional differences in DOJ prosecutions. Another interpretation consistent with our findings would be that SNF systems, which may have different ethical norms and legal risk tolerances, consider the deterrent effect of enforcement, but do so at an aggregate system rather than individual facility-level. Stronger or differentially targeted enforcement strategies particularly targeted at systems could play a critical role in constraining opportunistic behavior.

Additionally, reverse causality or omitted variables can result in biased estimates. However, a bounding exercise of the form of Oster (2019) indicates that selection on unobservable variables would need to be quite large to offset the estimated effect of system practices. Proportional selection on unobservable variables would reduce the coefficient moderately to 0.51. Relaxing the equal selection assumption, unobservable variables would need to have nearly 3.5 times the proportion impact of the 19 observable variables to decrease the coefficient on system practices to zero.[48] We examine causal channels and potential mechanisms below.

6. Mechanisms for Ownership Structure to Influence Billing

We next explore the mechanisms through which ownership could influence billing and assess the causal role of firm culture in shaping facility behavior. Specifically, we explore how coding practices change following acquisitions, spread within SNF systems, expand through acquisitions, and whether certain forms of ownership are more likely to drive upcoding. Finally, we provide back-of-the-envelope estimates of the fiscal costs associated with these-ownership-driven practices.

[48] Bias-adjusted coefficient of 0.51 assuming maximum R-square of 30% higher is computed [0.51=0.713-(0.576-0.443) × (0.755-0.713) / (0.443-0.416)]. Proportional impact of 3.45 times is computed [$3.45=\frac{0.713}{(\frac{0.576-0.443}{0.443-0.416})\times(0.755-0.713)}$] the proportional impact of observed control variables to decrease the coefficient to zero.

6.1. *Coding Activity After Mergers*

Although fraud levels are largely similar within a system, the most direct test of whether management drives fraud is to examine billing following changes in ownership. Acquisitions are often accompanied by shifts in management strategy that raise revenues and reduce costs (Eliason et al., 2020). If higher coding intensity at certain SNF systems reflect management-driven upcoding, then facilities acquired by these systems should increase use of highly-compensating codes. Using CMS Change of Ownership (CHOW) data, we identify 2,318 acquisitions between October 2019 and December 2022, of which 1,327 were by recognized SNF systems and 561 by systems in the top tercile of RUG-IV rehab. We compare changes in coding intensity at facilities acquired by high-rehab systems to facilities not acquired during the sample period using a staggered difference-in-differences design. Recognizing recent empirical concerns about staggered difference-in-difference designs (Baker et al., 2022; Roth et al., 2023), we use a stacked cohort design similar to Cengiz et al. (2019). Facility × cohort fixed effects control for time-invariant heterogeneity at the facility level, while quarter-year × cohort fixed effects capture aggregate time trends in coding intensity.[49] Figure 8 plots coefficients from a dynamic difference-in-differences regression:

$$CodingIntensity_{ijct} = \alpha + \sum_{t=1}^{T} \beta_t Period_t \times AcquiredHighRehab_j + \Gamma_{jc} + \delta_{tc} + \epsilon_{ijt} \qquad (9)$$

where $Period_t$ indicates the quarter relative to acquisition.[50]

Coding intensity increases in the first quarter following acquisition by 0.06 and continues increasing before stabilizing at 0.245, an increase of 23.3% relative to the unconditional mean of 1.04.

[49] Acquisitions by systems not in the highest tercile of RUG-IV rehab are excluded so that the comparison is to facilities with no ownership changes.

[50] Standard errors are clustered at the facility level. We find no evidence of differential pre-trends prior to acquisition, suggesting that billing practices at acquired facilities were similar to those at non-acquired facilities.

Results are precisely estimated, with coefficients for each treated quarter significant at the 1% level.[51] The systematic and sudden increase in coding intensity at hundreds of facilities following acquisition is consistent with management-driven upcoding.

6.2. *Diffusion of Practices Within a System*

If fraud depends on both technical know-how and the social normalization of misconduct (Sutherland, 1939), then facilities in the same system should spread their practices within the system at the direction of management across facilities, even when operating in different markets. If facility-level variation is driven by latent, idiosyncratic factors, we would not expect substantial convergence within a system over time. We test this by estimating:

$$\text{CodingIntensity}_{i,s,t+x} = \alpha + \beta\,\text{SystemAverage}_{-i,s,t} + \theta\,\text{FacilityCoding}_{i,t} + \delta_t + \gamma_i + \varepsilon_{i,s,t+x} \quad (10)$$

where $\text{SystemAverage}_{-i,s,t}$ is the system-quarter mean of PDPM billing intensity excluding facility i. Figure 7 Panel A shows that a one-standard-deviation increase in peer coding raises next-quarter coding intensity at a facility by 0.31 standard deviations after controlling for patient characteristics and including facility and time fixed effects. Adding an additional control for a facility's upcoding reduces the coefficient slightly to 0.23. Predictive effects dissipate each quarter but are still associated with a 0.12 standard deviation increase after four quarters and the effects remain statistically significant up to seven quarters in the future.

To further explore diffusion practices, we partition peers by geography and re-estimate the predictive regression in Equation 10 using (i) facilities at a certain distance away and (ii) only out-of-state peers. Because these facilities operate in distinct local markets, the predictive power would be difficult to attribute to shared local demand shocks, local enforcement, or local market structure.

[51] This increase is concentrated in four of the five individual potential upcoding categories (Figure IA.13).

Consistent with coding practices driven by centralized management, the coding of peer facilities within a system remains a positive, statistically significant predictor of a facility's future coding even for facilities further than 100, 250, 500, and 1,000 miles away.[52]

As a complementary test of within-system diffusion, we classify facilities as initially above or below their system's mean PDPM coding from October 1, 2019 to December 31, 2020, and then track outcomes for three subsequent years. To mitigate concerns about aggregate mean reversion, each below-average facility is matched to an above-average peer using nearest-neighbor matching on baseline coding (the first five PDPM quarters).[53] We then estimate a difference-in-differences with facility fixed effects and matched-pair by quarter fixed effects. Consistent with a diffusion of practices, Figure 7 Panel B shows substantially larger subsequent increases of 0.25 for facilities initially below the system average, an increase of 24% relative to the mean of 1.04.

6.3. Billing Practices and Acquisitions

Are billing practices of a SNF system related to firm growth? More profitable firms typically grow faster, but reputational concerns or regulator pressure might limit growth if regulatory licensure decisions consider billing fraud by an entity.[54]

Utilizing the CMS Change of Ownership (CHOW) database, we find that systems in the highest tercile of RUG-IV rehab acquired 996 facilities, 78% more than the number of facilities acquired by systems in the lowest tercile of rehab (559) as shown in Figure 8. Greater expansion of systems that engage in aggressive billing practices suggests upcoding may increase profitability and fuel growth.

[52]Furthermore, coefficients for far away facilities are economically close to the unrestricted sample indicating that local factors likely explain little of the predictive variation. For example, Figure IA.4 shows that the coefficient on other system coding for facilities at least 1,000 miles away is 0.227 versus 0.230 for the unrestricted sample.

[53]The comparison is thus between a facility with a below system-average PDPM coding intensity and a facility with above system-average coding, but with similar overall coding levels.

[54]CMS advises state regulators to consider the aggregate practices of affiliated facilities when reviewing acquisitions.https://data.cms.gov/sites/default/files/2023-06/FAQ_Affiliated_Entity_Performance.pdf

6.4. *Ownership Type and Facility Billing Practices*

Are excess billing practices related to facility ownership structure? Broad ownership categories of SNFs are for-profit, non-profit, and governmental, with private equity-backed facilities considered separately. Skilled nursing is dominated by for-profit facilities with 74.1% of patients, while non-profits and government operated facilities represent 22.5% and 3.4%, respectively. Private equity-backed systems, a subset of for-profits, treat 6.07% of patients.[55]

To examine how ownership type relates to billing, Panel A of Figure 9 sorts facilities into deciles based on excess rehab under RUG-IV and plot the distribution of PDPM coding intensity and the market share of each ownership type within each decile. Several broad patterns emerge: First, non-profits bill less intensively than for-profits, accounting for 50.7% of facilities in the lowest decile but only 8.61% in the highest. Second, the most aggressive facilities are predominately for-profit, but not private equity backed, with non-PE for-profits comprising 82.34% of the top decile. Third, coding intensity under PDPM is positively related to RUG-IV rehab for all ownership types, though the relationship is stronger for for-profit and private equity facilities. Finally, within each group, the average private equity operators tend to code slightly more aggressively. Overall, for-profit as well as private equity owned facilities code patients most aggressively, but there is considerable variation within ownership type indicating that ownership type alone is not a complete fraud explanation.

We further explore the importance of firm culture within various ownership types, by plotting SHAP values of upcoding at other facilities within a system against facility-upcoding rates in Figure 9 which shows that firm-specific upcoding practices explain upcoding rates within each of the major

[55]While we collect private equity data from major data providers including Pitchbook and hand-gathered data, we cannot ensure that every PE-backed facility is identified due to data quality issues as well as incomplete data. The 8.2% of for-profit patients being treated by private equity firms is similar to that estimated by Gupta et al. (2023) in 2015.

ownership group types. Traditional for-profit operated SNFs display both the most extreme system-level upcoding and firm-specific practices have the strongest impact on predicted upcoding in this ownership group.

6.5. Discussion of Upcoding Costs

We consider the potential costs arising from patient upcoding at facilities in opportunistic systems The average cost during PDPM per patient for a SNF system in the top tercile of excess rehab is \$20,319. Given the approximately 521,000 patients visiting these systems each year, if these patients were instead treated at a median cost facility (at \$16,296 per patient) the annual costs would be \$2.10 billion lower.[56] This amounts to approximately 7.2% of total SNF Medicare spending. There are, of course, limitations to this back-of-the envelope calculation. Though the general patterns of coding intensity are consistent with upcoding by opportunistic facilities, there could be some exceptions. Nevertheless, there are several reasons this approach likely understates the costs of fraud. First, this approach ignores any upcoding among facilities not previously in the top tercile of RUG-IV rehab. Panel A of Figure 9 suggests that there are some facilities with previously low RUG-IV rehab that upcode under PPDM. Second, we consider only a subset of PDPM codes and other fraudulent coding practices could be used to increase revenue.[57] Thus we expect that the above back-of-the-envelope estimates understate the true costs of fraud. The fact that within-market variation is almost as large as cross-market variation implies that policies which encourage patients to visit truthful operators might offer substantial benefit.

[56] 521,000 × (\$20,319-16,296)=\$2.10 billion.

[57] For example, facilities might also increase revenue through other fraudulent means that are more difficult to measure on a large-scale such as illegitimate use of waivers for the three-day hospitalization requirement following COVID, steering or referral of clients from related party ownership of hospitals, patient referrals reflecting kickbacks, or billing for services not actually provided.

7. Conclusion

This paper documents widespread skilled nursing overbilling that persists across fundamentally different reimbursement regimes, contrary to the expectations of policymakers. Higher billing intensity is not explained by sicker patients, patient selection, or more complete coding, but by systematic upcoding of specific comorbidities absent in the immediately preceding hospital stay.

Using this precisely measured upcoding, we examine a range of potential fraud determinants. Facility upcoding appears undeterred by a change in billing regime or by regional variation in DOJ enforcement levels. Nor is upcoding related to regional crime, local social norms, or local competition. Facility upcoding is strongly related to upcoding by other facilities in the system (firm) and to a lesser extent the upcoding practices of other systems in the same geography. Fraudulent SNF systems spread billing practices within system and across systems through acquisition with systems in the top tercile of fraudulent billing acquiring facilities 78% faster than other systems. Overall, within-firm norms appear to be the primary determinant of fraud, but virtuous norms are quickly being eroded over time as aggressive upcoding practices expand within and across firms.

Our findings suggest several areas for potential reform and further research. First, given the importance of firm-wide practices, enforcement and policy should consider the practices of all facilities operated by an affiliation. CMS should continue to expand initiatives to increase transparency regarding facility ownership. Second, given that these practices appear to be mostly driven by top ownership decisions, more criminal charges (Coffee Jr, 1980; Polinsky and Shavell, 1993; Admati et al., 2025) and substantially larger monetary penalties may be needed, especially in settings where liability may be shielded (Gandhi and Olenski, 2024). Third, more research is needed to further detect other forms of fraud and its externalities. Griffin and Priest (2025) build on this paper by showing that the fraudulent SNF systems provide inferior care to patients, resulting in bed sores, urinary tract infections, and higher mortality, outcomes that are partly obscured by underreporting, suggesting that the costs of fraud extend beyond measurable economic transfers from taxpayers. Fourth, given the widespread and growing fraudulent practices we document, monitoring tools with proven deterrent effects such as Medicare's Recovery Audit Contractor (RAC) Program (Shi, 2023), prior authorization Eliason et al. (2021), whistleblower and DOJ suits (Leder-Luis, 2023; Howard and McCarthy, 2021) may have context-specific effectiveness and should be more widely applied. Fifth, SNF mergers and acquisitions should only be approved for systems that adhere to appropriate billing levels.

Given imperfect information, legal, enforcement, and regulatory hurdles, administration inefficiencies, and likely substantial legal and political capital spent by fraudulent SNFs, radical reforms appear warranted to disincentivize Medicare fraud and encourage the growth of the many honest SNF systems not engaged in fraudulent practices.

References

Admati, A. R., Pfleiderer, P., and Atkinson, N. (2025). Profitable misconduct, corporate governance, and law enforcement. *Corporate Governance, and Law Enforcement (June 06, 2025)*.

Afendulis, C. C. and Kessler, D. P. (2007). Tradeoffs from integrating diagnosis and treatment in markets for health care. *American Economic Review*, 97(3):1013–1020.

Andreyeva, E., Gupta, A., Ishitani, C., Sylwestrzak, M., and Ukert, B. (2024). The corporatization of independent hospitals. *Journal of Political Economy Microeconomics*, 2(3):602–663.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5):941–973.

Baker, A. C., Larcker, D. F., and Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2):370–395.

Bastani, H., Goh, J., and Bayati, M. (2019). Evidence of upcoding in pay-for-performance programs. *Management Science*, 65(3):1042–1060.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of political economy*, 76(2):169–217.

Bowblis, J. R. and Brunt, C. S. (2014). Medicare skilled nursing facility reimbursement and upcoding. *Health economics*, 23(7):821–840.

Bowblis, J. R., Brunt, C. S., and Grabowski, D. C. (2016). Competitive spillovers and regulatory exploitation by skilled nursing facilities. In *Forum for Health Economics and Policy*, volume 19, pages 45–70. De Gruyter.

Brownlee, S., Chalkidou, K., Doust, J., Elshaug, A. G., Glasziou, P., Heath, I., Nagpal, S., Saini, V., Srivastava, D., Chalmers, K., et al. (2017). Evidence for overuse of medical services around the world. *The Lancet*, 390(10090):156–168.

Card, D., Fenizia, A., and Silver, D. (2023). The health impacts of hospital delivery practices. *American Economic Journal: Economic Policy*, 15(2):42–81.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.

Chen, A., Blumenthal, D. M., and Jena, A. B. (2018). Characteristics of physicians excluded from us medicare and state public insurance programs for fraud, health crimes, or unlawful prescribing of controlled substances. *JAMA network open*, 1(8):e185805–e185805.

Chen, A. C., Skinner, R. J., Braun, R. T., Konetzka, R. T., Stevenson, D. G., and Grabowski, D. C. (2024). New cms nursing home ownership data: Major gaps and discrepancies: Study examines completeness of nursing home ownership data. *Health Affairs*, 43(3):318–326.

Chen, B. K., Gertler, P. J., and Yang, C.-Y. (2016). Physician ownership of complementary medical services. *Journal of Public Economics*, 144:27–39.

Chetty, R., Friedman, J. N., and Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review*, 103(7):2683–2721.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022). Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921):108–121.

Clemens, J. and Gottlieb, J. D. (2014). Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review*, 104(4):1320–1349.

Coffee Jr, J. C. (1980). Making the punishment fit the corporation: The problems of finding an optimal corporation criminal sanction. *N. Ill. UL Rev.*, 1:3.

Dafny, L., Duggan, M., and Ramanarayanan, S. (2012). Paying a premium on your premium? consolidation in the us health insurance industry. *American Economic Review*, 102(2):1161–1185.

Dafny, L. S. (2005). How do hospitals respond to price changes? *American Economic Review*, 95(5):1525–1547.

Di Tella, R. and Schargrodsky, E. (2004). Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, 94(1):115–133.

Dills, A. K., Miron, J. A., and Summers, G. (2008). What do economists know about crime?

Dimmock, S. G., Gerken, W. C., and Graham, N. P. (2018). Is fraud contagious? coworker influence on misconduct by financial advisors. *The Journal of Finance*, 73(3):1417–1450.

Diwan, R., Eliason, P., League, R., Leder-Luis, J., McDevitt, R. C., and Roberts, J. W. (2025). Competition and fraud in health care. *Work in Progress*.

Doraszelski, U., Seim, K., Sinkinson, M., and Wang, P. (2024). Ownership concentration and strategic supply reduction. Technical report, American Economic Review.

Dupont, Q. and Karpoff, J. M. (2025). A theory of social capital and trust. *HKU Jockey Club Enterprise Sustainability Global Research Institute Paper*, (2025/074).

Egan, M., Matvos, G., and Seru, A. (2019). The market for financial adviser misconduct. *Journal of Political Economy*, 127(1):233–295.

Eide, E., Rubin, P. H., Shepherd, J. M., et al. (2006). Economics of crime. *Foundations and Trends® in Microeconomics*, 2(3):205–279.

Einav, L., Finkelstein, A., and Mahoney, N. (2025). Producing health: measuring value added of nursing homes. *Econometrica*.

Eliason, P. J., Heebsh, B., McDevitt, R. C., and Roberts, J. W. (2020). How acquisitions affect firm behavior and performance: Evidence from the dialysis industry. *The Quarterly Journal of Economics*, 135(1):221–267.

Eliason, P. J., League, R. J., Leder-Luis, J., McDevitt, R. C., and Roberts, J. W. (2021). Ambulance taxis: The impact of regulation and litigation on health care fraud.

Fajnzylber, P., Lederman, D., and Loayza, N. (2002). Inequality and violent crime. *The journal of Law and Economics*, 45(1):1–39.

Feldman, R. (2001). An economic explanation for fraud and abuse in public medical care programs. *The Journal of Legal Studies*, 30(S2):569–577.

Fisman, R. and Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political economy*, 115(6):1020–1048.

Flint, L. A., David, D. J., and Smith, A. K. (2019). Rehabbed to death. *N Engl J Med*, 380(5):408–409.

Galizzi, M. M., Godager, G., Li, J., Linnosmaa, I., Tammi, T., and Wiesen, D. (2023). Economics of healthcare provider altruism. In *Handbook of labor, human resources and population economics*, pages 1–23. Springer.

Gandhi, A. (2023). Picking your patients: Selective admissions in the nursing home industry. *Available at SSRN*, 3613950.

Gandhi, A. and Olenski, A. (2024). Tunneling and hidden profits in health care. *Working Paper*.

Gaynor, M. et al. (2007). Competition and quality in health care markets. *Foundations and Trends® in Microeconomics*, 2(6):441–508.

Gaynor, M., Rebitzer, J. B., and Taylor, L. J. (2004). Physician incentives in health maintenance organizations. *Journal of Political Economy*, 112(4):915–931.

Gaynor, M. and Town, R. J. (2011). Competition in health care markets. *Handbook of health economics*, 2:499–637.

Gertler, P. and Kwan, A. (2024). The essential role of altruism in medical decision making. Technical report, National Bureau of Economic Research.

Geruso, M. and Layton, T. (2020). Upcoding: evidence from medicare on squishy risk adjustment. *Journal of Political Economy*, 128(3):984–1026.

Geweke, J., Gowrisankaran, G., and Town, R. J. (2003). Bayesian inference for hospital quality in a selection model. *Econometrica*, 71(4):1215–1238.

Glaeser, E. L., Sacerdote, B., and Scheinkman, J. A. (1996). Crime and social interactions. *The Quarterly journal of economics*, 111(2):507–548.

Gould, E. D., Weinberg, B. A., and Mustard, D. B. (2002). Crime rates and local labor market opportunities in the united states: 1979–1997. *Review of Economics and statistics*, 84(1):45–61.

Grabowski, D. C., Feng, Z., Hirth, R., Rahman, M., and Mor, V. (2013). Effect of nursing home ownership on the quality of post-acute care: an instrumental variables approach. *Journal of health economics*, 32(1):12–21.

Griffin, J. M., Kruger, S., and Mahajan, P. (2025). Is fraud contagious? social connections and the looting of covid relief programs. *Review of Financial Studies*.

Griffin, J. M., Kruger, S., and Maturana, G. (2019). Personal infidelity and professional conduct in 4 settings. *Proceedings of the National Academy of Sciences*, 116(33):16268–16273.

Griffin, J. M. and Priest, A. (2023). Is covid revealing a virus in cmbs 2.0? *The Journal of Finance*.

Griffin, J. M. and Priest, A. (2025). Overbilling and killing? an examination of the skilled nursing industry. *Working Paper*.

Gruber, J., Howard, D. H., Leder-Luis, J., and Caputi, T. L. (2025). Dying or lying? for-profit hospices and end of life care. *American Economic Review*.

Grullon, G., Kanatas, G., and Weston, J. (2009). Religion and corporate (mis) behavior. *Available at SSRN 1472118*.

Guiso, L., Sapienza, P., and Zingales, L. (2015). The value of corporate culture. *Journal of financial economics*, 117(1):60–76.

Gupta, A., Howell, S. T., Yannelis, C., and Gupta, A. (2023). Owner incentives and performance in healthcare: Private equity investment in nursing homes. *The Review of Financial Studies*, page hhad082.

Hackmann, M. B., Pohl, R. V., and Ziebarth, N. R. (2024). Patient versus provider incentives in long-term care. *American Economic Journal: Applied Economics*, 16(3):178–218.

Heese, J., Krishnan, R., and Moers, F. (2015). Regulator leniency and mispricing in beneficent nonprofits. 2015(1):11998.

Heese, J., Pérez-Cavazos, G., and Peter, C. D. (2022). When the local newspaper leaves town: The effects of local newspaper closures on corporate misconduct. *Journal of Financial Economics*, 145(2):445–463.

Herpfer, C., Lin, J., and Maturana, G. (2024). Corporate behavior when running the firm for stakeholders: Evidence from hospitals.

Howard, D. H. and McCarthy, I. (2021). Deterrence effects of antifraud and abuse enforcement in health care. *Journal of Health Economics*, 75:102405.

Howell, S. T. and Liu, T. (2025). Private equity in healthcare. In *The Palgrave Encyclopedia of Private Equity*, pages 913–919. Springer.

Hyman, D. A. (2001). Health care fraud and abuse: market change, social norms, and the trust "reposed in the workmen". *The Journal of Legal Studies*, 30(S2):531–567.

Joiner, K. A., Lin, J., and Pantano, J. (2024). Upcoding in medicare: where does it matter most? *Health Economics Review*, 14(1):1.

Kreps, D. M. et al. (1990). Corporate culture and economic theory. *Perspectives on positive political economy*, 90(109-110):8.

League, R. (2022). Administrative burden and consolidation in health care: Evidence from medicare contractor transitions.

Leder-Luis, J. (2023). Can whistleblowers root out public expenditure fraud? evidence from medicare. *Review of Economics and Statistics*, pages 1–49.

Leder-Luis, J. and Malani, A. (2025). The economics of healthcare fraud. Technical report, National Bureau of Economic Research.

Li, K., Mai, F., Shen, R., and Yan, X. (2021). Measuring corporate culture using machine learning. *The Review of Financial Studies*, 34(7):3265–3315.

Liu, T. (2022). Bargaining with private equity: Implications for hospital prices and patient welfare. *Available at SSRN 3896410*.

Liu, X. (2016). Corruption culture and corporate misconduct. *Journal of Financial Economics*, 122(2):307–327.

Lochner, L. and Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American economic review*, 94(1):155–189.

Matsueda, R. L. (1988). The current state of differential association theory. *Crime & Delinquency*, 34(3):277–306.

McClellan, M., McNeil, B. J., and Newhouse, J. P. (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality?: analysis using instrumental variables. *Jama*, 272(11):859–866.

McGloin, J. M. and Thomas, K. J. (2019). Peer influence and delinquency. *Annual Review of Criminology*, 2(1):241–264.

Newhouse, J., Garber, A., Graham, R., McCoy, M., Mancher, M., and Kibria, A. (2013). Variation in health care spending: Target decision making. *Not Geography, Washington, DC: Institute of Medicine*.

Olenski, A. and Sacher, S. (2024). Estimating nursing home quality with selection. *Review of Economics and Statistics*, pages 1–31.

O'Malley, A. J., Bubolz, T. A., and Skinner, J. S. (2021). The diffusion of health care fraud: A network analysis.

O'Reilly, C. A. and Chatman, J. A. (1996). Culture as social control: Corporations, cults, and commitment.

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

Parsons, C. A., Sulaeman, J., and Titman, S. (2018). The geography of financial misconduct. *The Journal of Finance*, 73(5):2087–2137.

Piskorski, T., Seru, A., and Witkin, J. (2015). Asset quality misrepresentation by financial intermediaries: Evidence from the rmbs market. *The Journal of Finance*, 70(6):2635–2678.

Polinsky, A. M. and Shavell, S. (1993). Should employees be subject to fines and imprisonment given the existence of corporate liability? *International Review of Law and Economics*, 13(3):239–257.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*.

Shekhar, S., Leder-Luis, J., and Akoglu, L. (2023). Unsupervised machine learning for explainable health care fraud detection.

Shi, M. (2023). Monitoring for waste: Evidence from medicare audits.

Shrank, W. H., Rogstad, T. L., and Parekh, N. (2019). Waste in the us health care system: estimated costs and potential for savings. *Jama*, 322(15):1501–1509.

Silverman, E. and Skinner, J. (2004). Medicare upcoding and hospital ownership. *Journal of health economics*, 23(2):369–389.

Sutherland, E. H. (1939). *Principles of Criminology*.

Sutherland, E. H. (1947). Differential association theory. *Criminology Theory: Selected Classic Readings*, pages 225–227.

Temkin-Greener, H., Lee, T., Caprio, T., and Cai, S. (2019). Rehabilitation therapy for nursing home residents at the end-of-life. *Journal of the American Medical Directors Association*, 20(4):476–480.

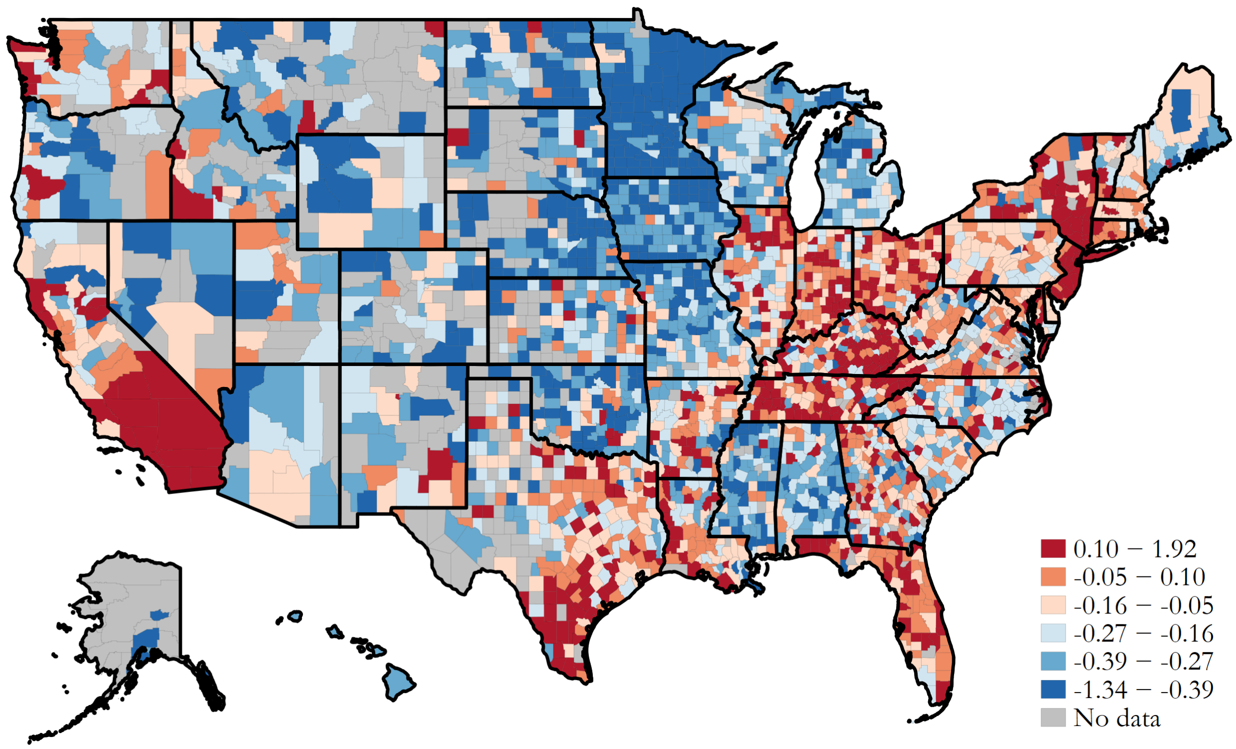
Wang, V., Maciejewski, M. L., Coffman, C. J., Sanders, L. L., Lee, S.-Y. D., Hirth, R., and Messana, J. (2017). Impacts of geographic distance on peritoneal dialysis utilization: refining models of treatment selection. *Health Services Research*, 52(1):35–55.

Witte, A. D. (1980). Estimating the economic model of crime with individual data. *The quarterly journal of economics*, 94(1):57–84.

Figure 1. Geographic Variation in Billing Practices

This figure shows geographic variation in PDPM Coding Intensity. Coding Intensity is the sum of five binary billing categories that increase revenue and is residualized on patient characteristics according to Equation 1 and measures the level of coding beyond what would be expected given observable patient characteristics. For ease of interpretation, the Coding Intensity is standardize to have mean zero and variance of one. Panel A shows the average coding intensity across counties, Panel B displays within county variation in coding intensity across providers. In Panel A, counties are colored according to their average coding intensity and counties with fewer than 500 patients are excluded. In Panel B, the average coding intensity for an individual facility is on the y-axis and the average county-level coding intensity is shown on the x-axis.

Panel A. Coding Intensity by County



Panel B: Within-County Variation in Coding Intensity

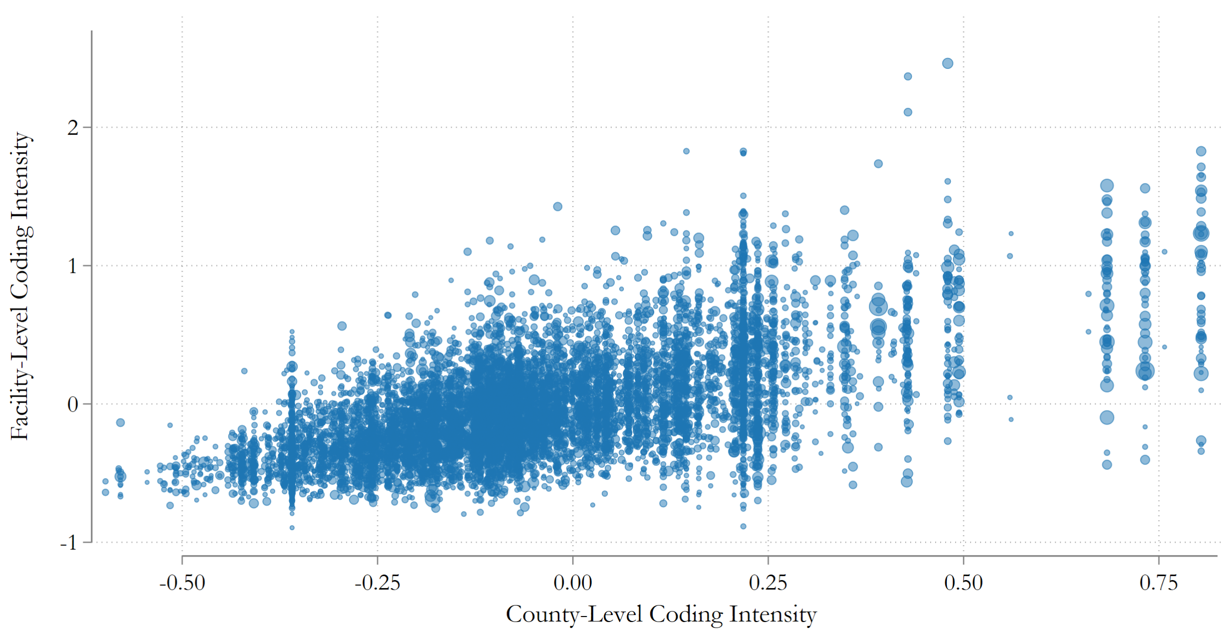
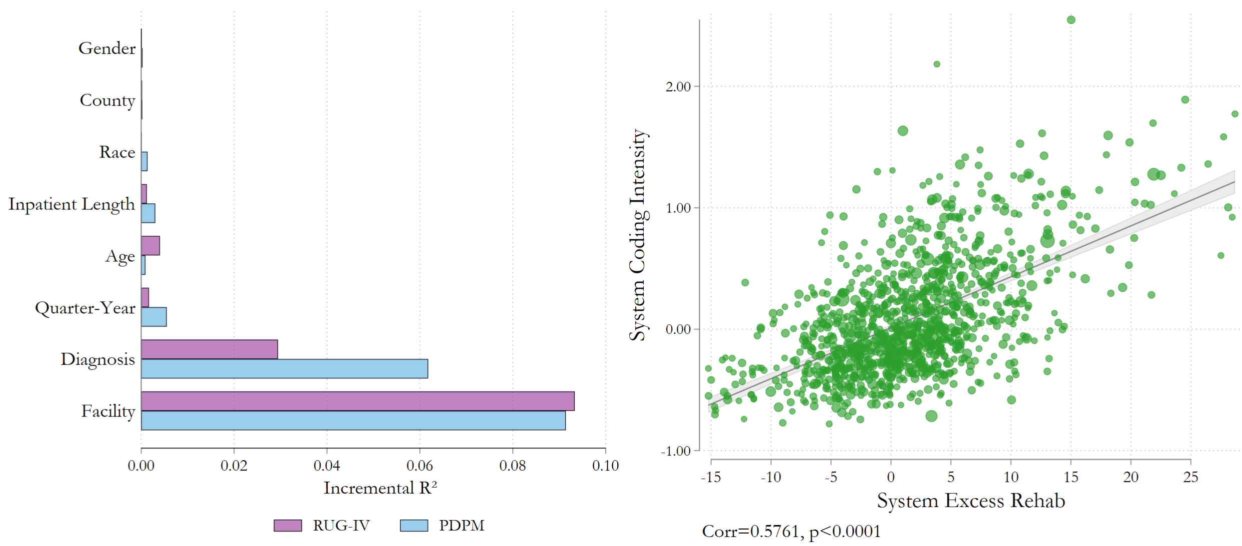


Figure 2. Importance of Facility in Excess Rehab and Coding Intensity

This figure explores the importance of the facility in explaining variation in either Rehab Utilization (RUG-IV) or Coding Intensity (PDPM). Patient characteristics include Gender, Race, and Age. Diagnosis is a patient's inpatient diagnosis at the CCSR-level. The left subgraph denotes the incremental Adjusted R-Square from each explanatory variable. Purple bars denote RUG-IV and the light blue bars denote Coding Intensity. The right subgraph is a scatterplot with a facility's Excess Rehab as determined by Equation 11 on the x-axis and the PDPM Coding Intensity of the y-axis. The size of the bubbles denotes the number of patients at each system. Panel B plots the estimated coefficient from a regression of standardized PDPM coding intensity on standardized facility excess rehab for each demographic subgroup. The dashed grey line denotes the estimated coefficient for the full sample.

Panel A: Importance of Facility to Billing Practices



Panel B: Stability of Coefficient by Patient Subsample

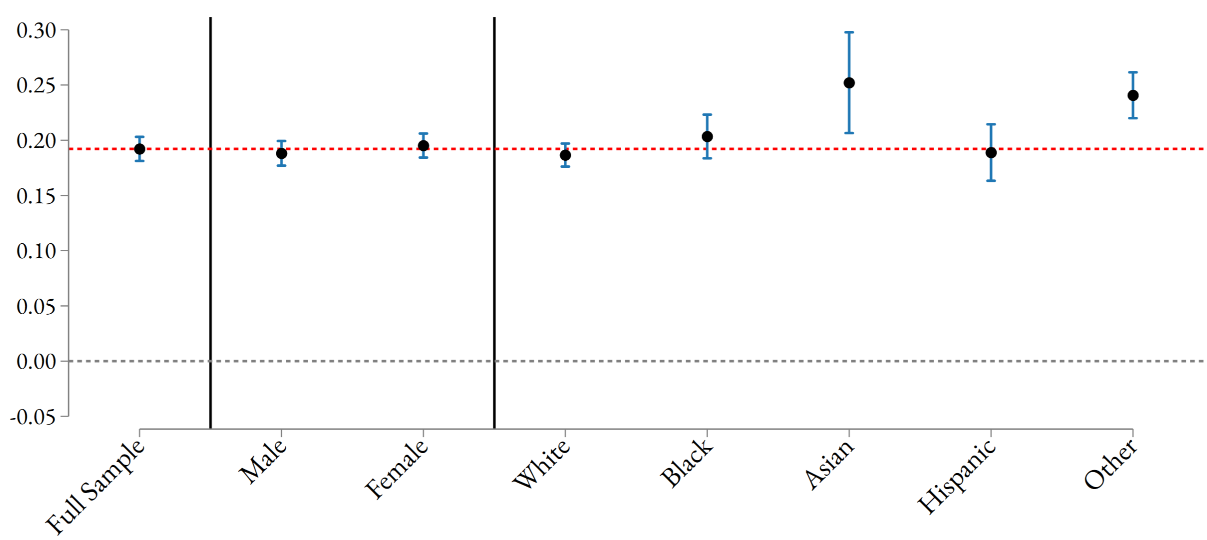
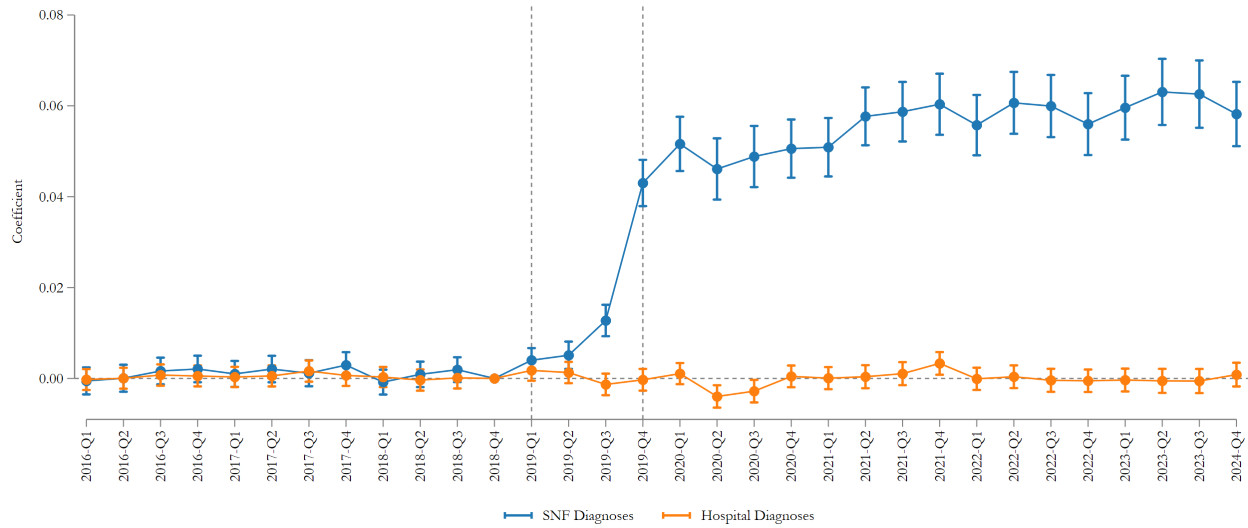


Figure 3. Diagnoses of Comorbidities Around Regime Change

Blue denotes comorbidities as recorded at the skilled nursing facility while orange denotes comorbidities as recorded at the preceding inpatient hospital stay. Coefficients are estimated using Equation 3, but the outcome variable is an indicator equal to one if any diagnoses are SLP- related (Left Subgraph) or NTA-related (Right Subgraph). 95% confidence intervals for each coefficient are displayed.

Panel A: Acute Neurologic Primary Diagnoses



Panel B: Incidence of SLP-Related and NTA-Related Comorbidities

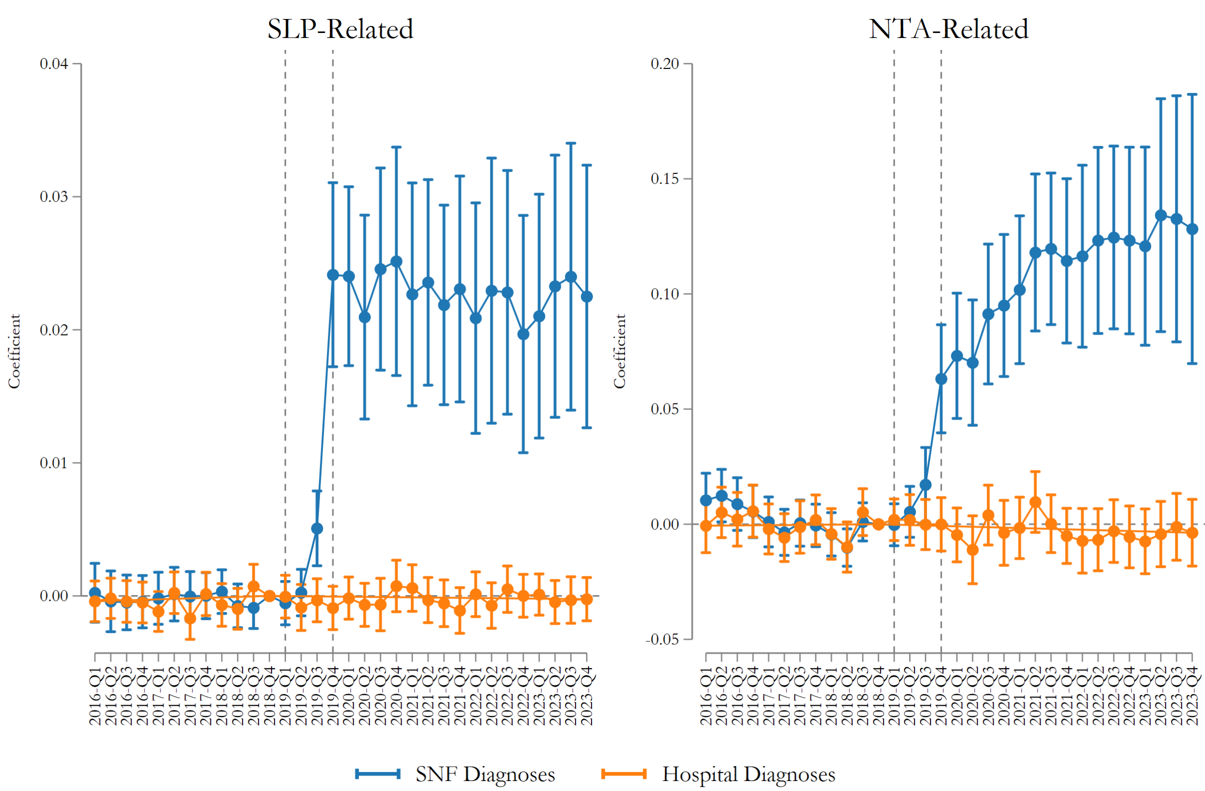


Figure 4. Reduced Form Relation with Coding Intensity

This figure explores the relationship between billing practices in the RUG-IV era and coding intensity in the PDPM era by instrumenting for *ExcessRehab* using Equation 6. This figure plots a reduced form specification in which indicators for each individual billing category are regressed on *ExpectedFacilityRehab*. Standard errors are clustered at the facility level.

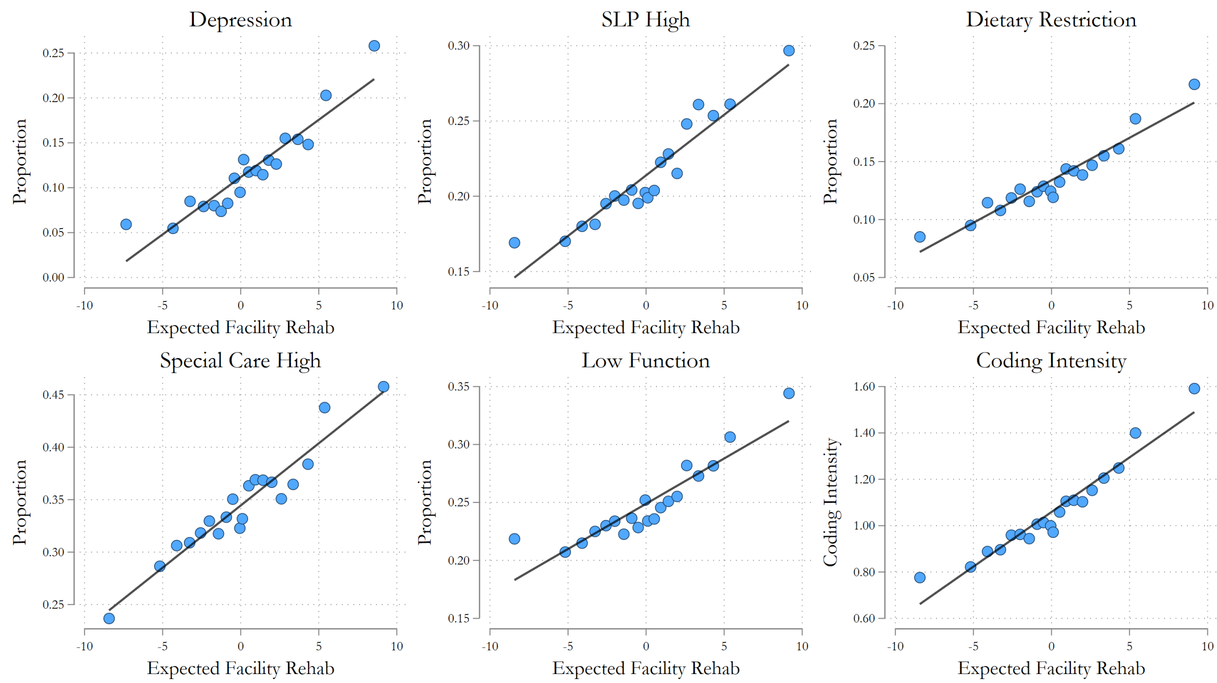
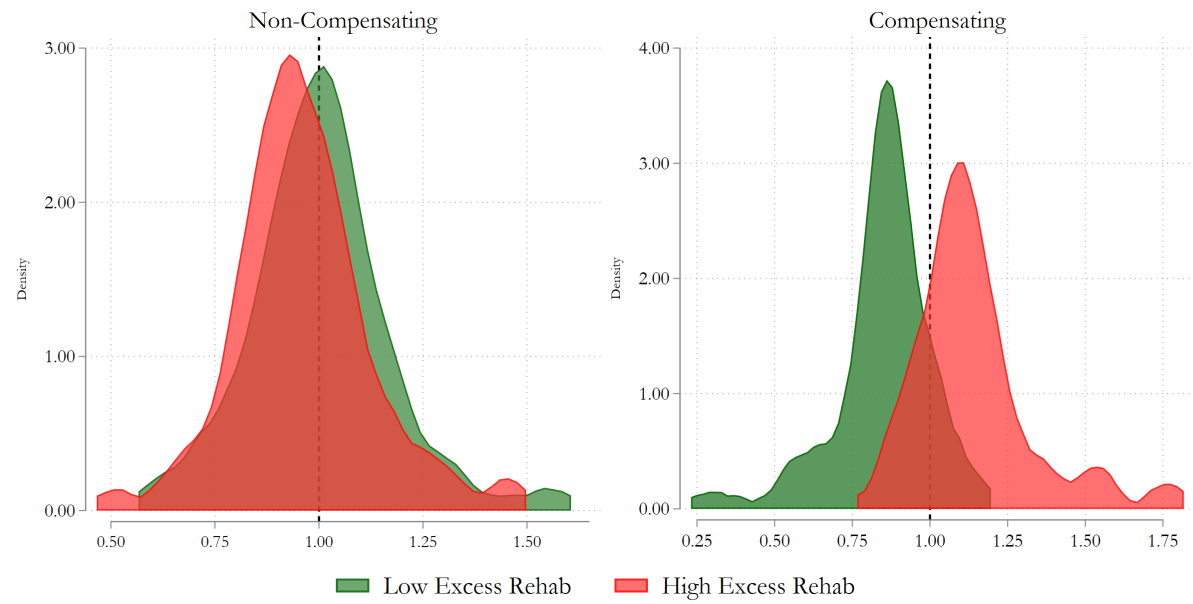


Figure 5. Intensity of Coding by Diagnoses Type

This figure explores the intensity of coding by displaying the ratio of patients receiving each diagnoses at opportunistic versus other facilities. Panel A displays a histogram where the observation unit is an ICD-10 diagnoses code. Diagnoses codes with at least 5,000 observations are included. Acute Neurologic, NTA-Related and SLP-Related Diagnoses are included among the compensating diagnoses while all other diagnoses are considered non-compensating. The red density plot denotes the coding rates at facilities in the highest tercile of excess rehab whil the green denotes facilities in the lowest tercile. Panel B explores distribution of higher diagnosis rates for compensating diagnoses using a bootstrap approach which randomly assigns an indicator for compensating. The simulation is repeated 10,000 times and each observation is an average diagnosis ratio.

Panel A: Distribution of Coding Intensity by Diagnoses Type



Panel B: Bootstrap Simulation

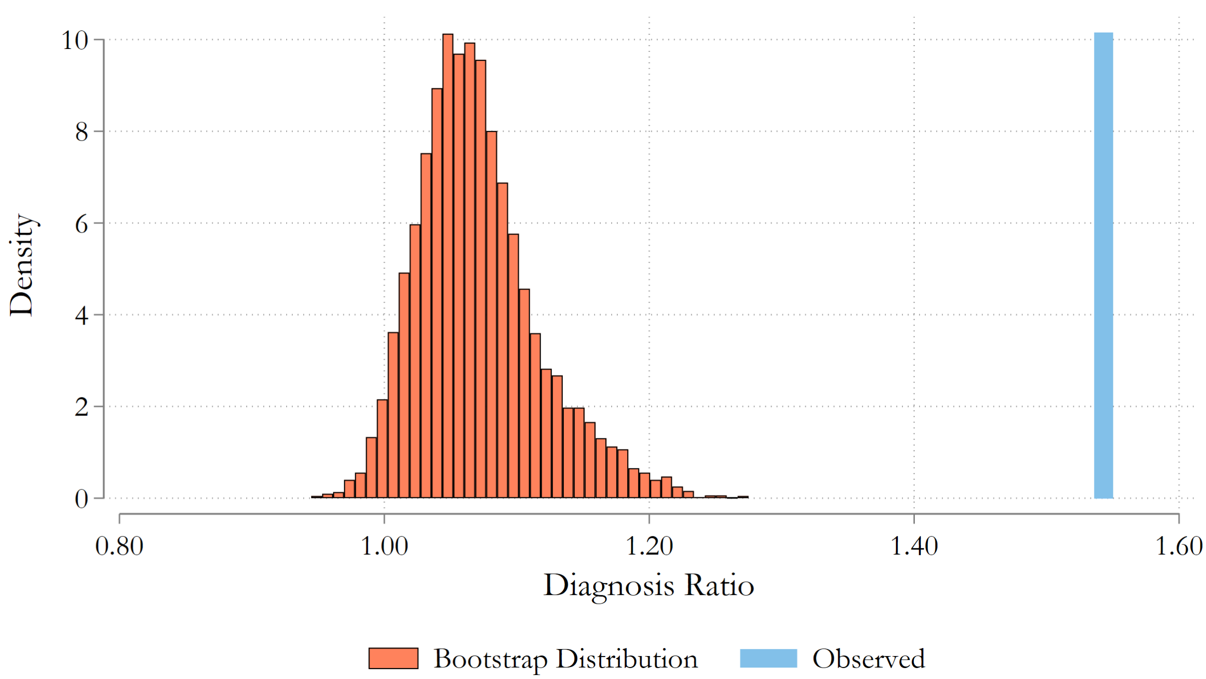
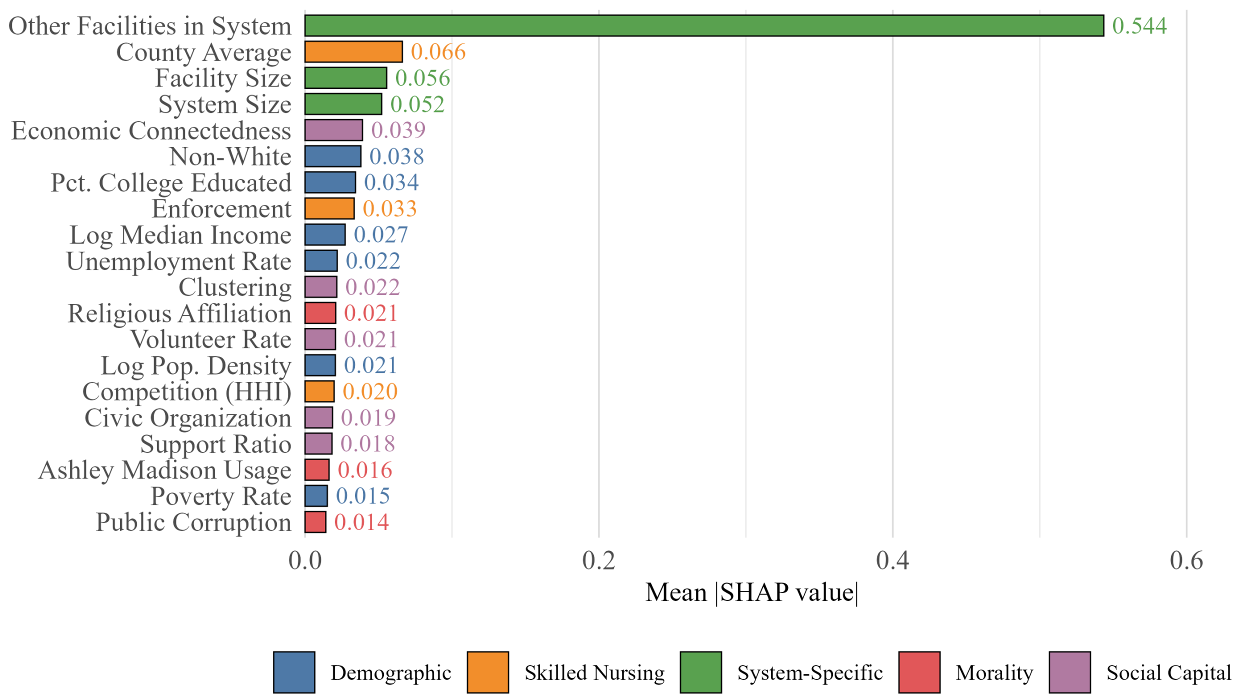


Figure 6. Determinants of Facility-Level Fraud

Panel A shows the average absolute SHAP values for a variable, which is the estimated contribution of that variable to a facility's upcoding rates. Predicted upcoding is estimated using the variables listed below and with default parameters of the gradient boosting decision tree model LightGBM. The bar of each variable is colored according to the type of variable being presented. Panel B plots how the top four explanatory variables affect facility upcoding in the gradient boosting decision tree models. The SHAP value for a variable is the estimated contribution of that variable. The x-axis denotes the value of each feature. The y-axis denotes the SHAP value, which is how much the feature changes the baseline prediction.

Panel A: Mean Absolute SHAP Value for Models Predicting Upcoding



Panel B: Dependence Plots for Top Four Features

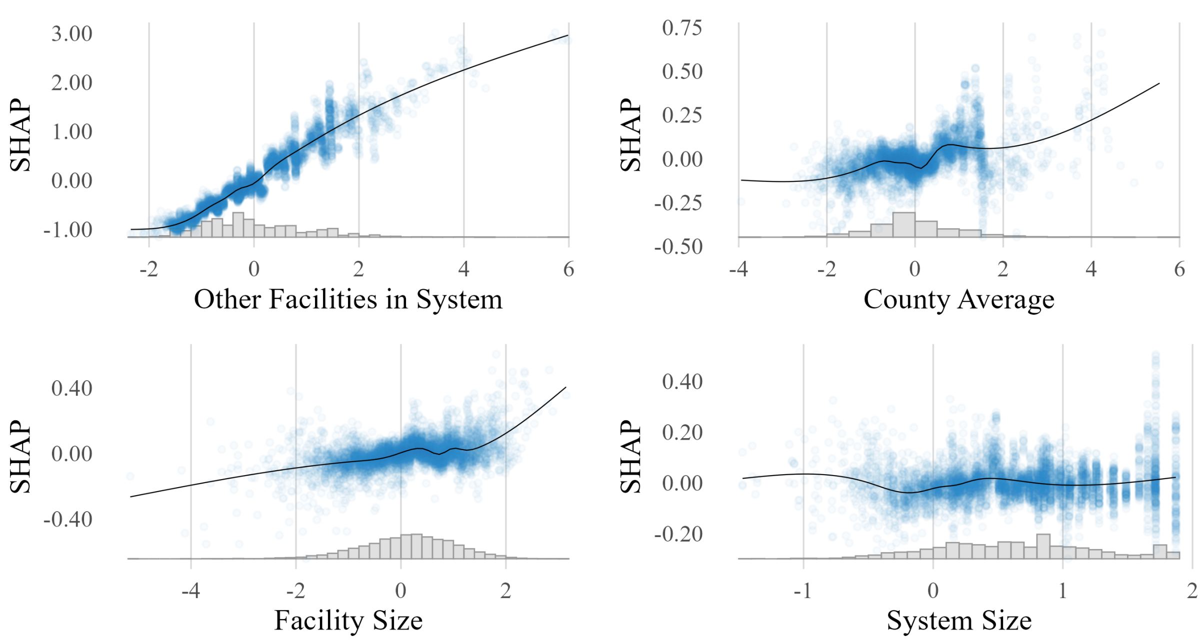
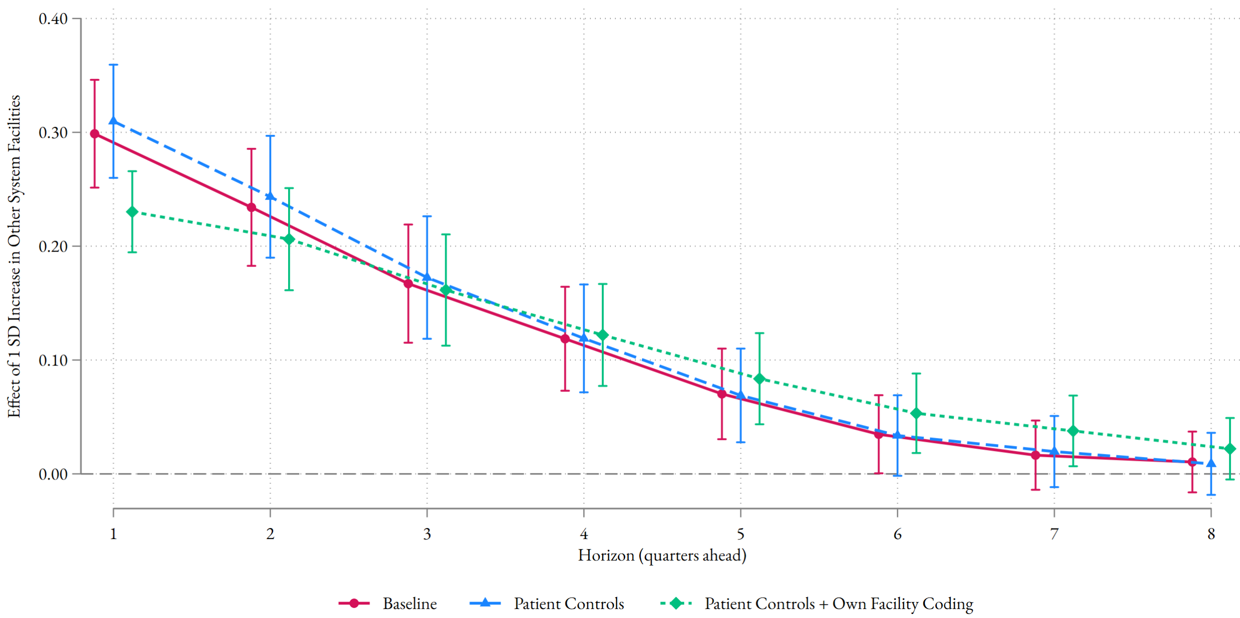


Figure 7. Within-System Diffusion of Coding Practices

This figure explores within-system diffusion of coding practices. Panel A explores the predicted effect on future upcoding for a one standard deviation increase in contemporaneous upcoding at other facilities within an SNF system up to eight quarters out. Standard errors denote 95% confidence intervals. Within a SNF system facilites are sorted based on whether PDPM coding intensity was below or above the system-level average from October 1, 2019 through December 31, 2020. To ensure that differences are not driven by aggregate mean reversion, we match each below-system average facility with an above system average facility based on pre-2020 PDPM coding levels. Difference-in-difference event study are presented in Panel B.

Panel A: Predictive Regressions of System Practices on Future Upcoding



Panel B: Diffusion of Practices Matched Difference-in-Differences

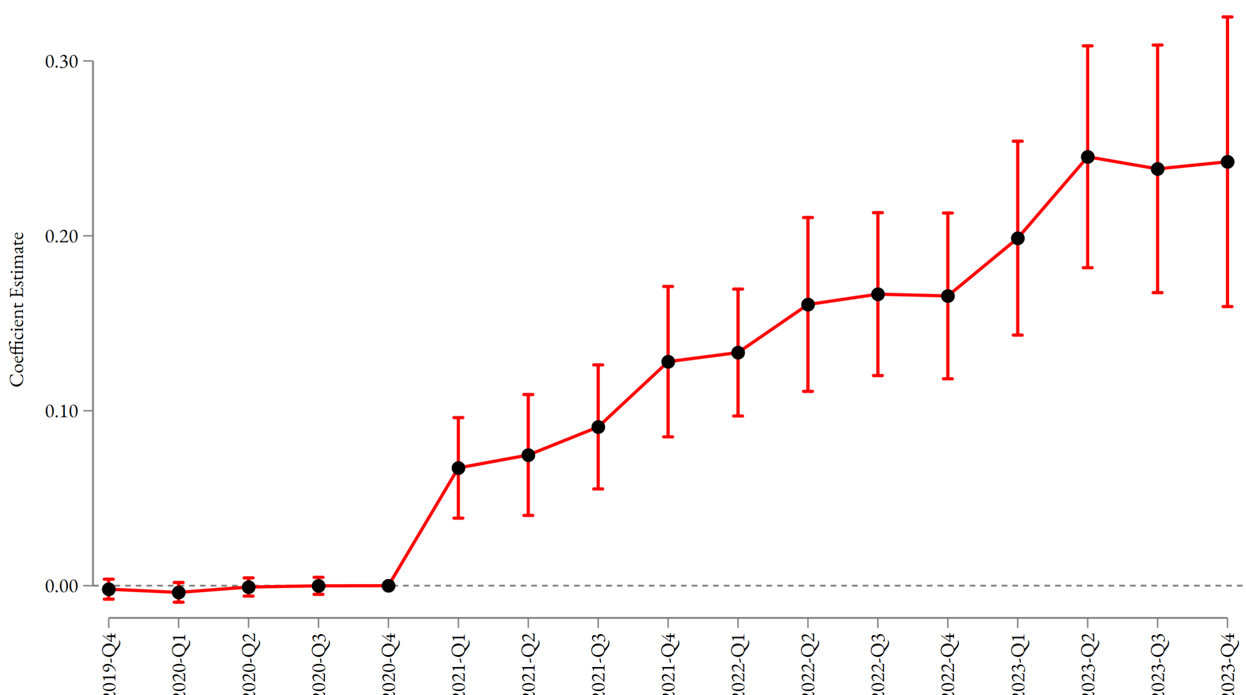
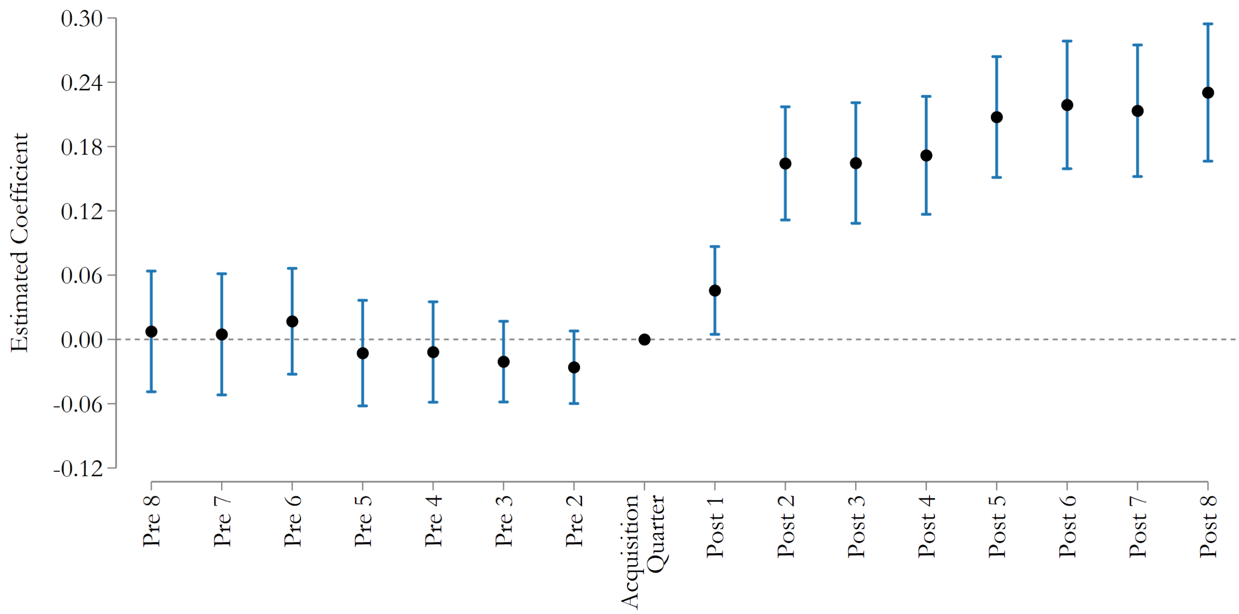


Figure 8. Change in Coding Intensity around Acquisition

This figure plots estimated coefficients from Equation 9. Robust standard errors are clustered at the facility level. 95% confidence intervals for each coefficient are displayed. Panel A studies the spread of fraudulent billing practices across SNF system after a facility is acquired by a system with high excess rehab. Panel B plots the spread of fraudulent billing practices by plotting the cumulative number of acquisitions by SNF systems according to their rehab billing practices under RUG-IV. SNF systems are sorted into terciles based on their excess rehab levels, estimated as the facility fixed effects according to Equation 11

Panel A: Across System Variation in Coding through Acquisition



Panel B: Acquisitions by System Type

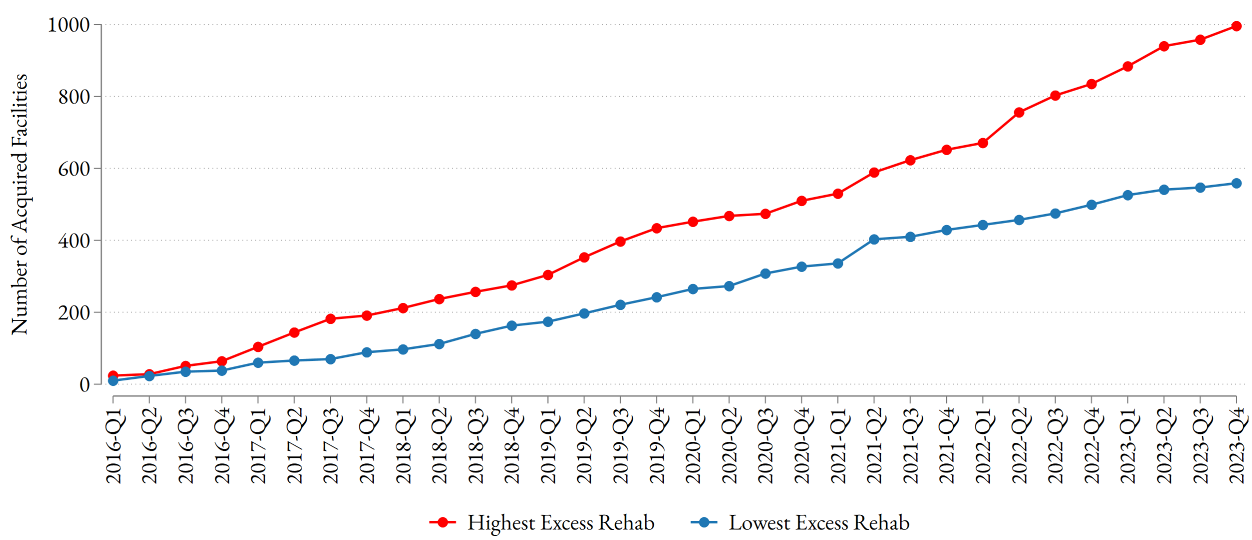
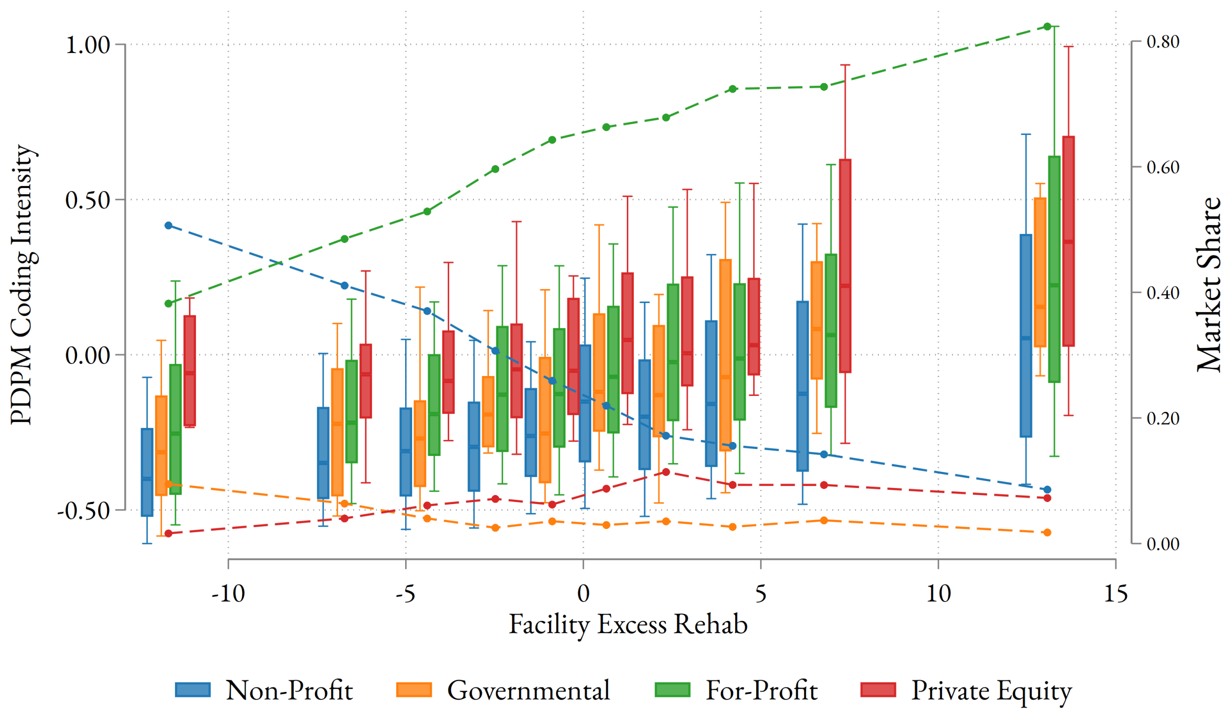


Figure 9. Ownership Type and Billing Practices

Panel A of this figure displays the 10th to 90th percentile of coding intensity for 20 quanitles of facility excess rehab. The interquartile ranges are denoted by the boxes, the median by the horizontal line and finally the mean is marked by the diamond. On the right axis of Panel A is the market share of each ownership type for the corresponding quantile of excess rehab. Market shares are denoted by the lines and color corresponds to ownership type. For-Profit excludes identified facilities with private equity backing. Panel B explores the composition of ownership by plotting the 250 lowest and highest coding facilities.The y-axis of each subgraph denotes the facility-level coding intensity. The color corresponds to onwership type. To be included in the figure, we require that a facility has at least 500 patients. Finally, Panel C displays the distribution of coding intensity for each ownership type.

Panel A: Ownership Type and Billing Practices



Panel B: Importance of System Level Practices by Ownership Type

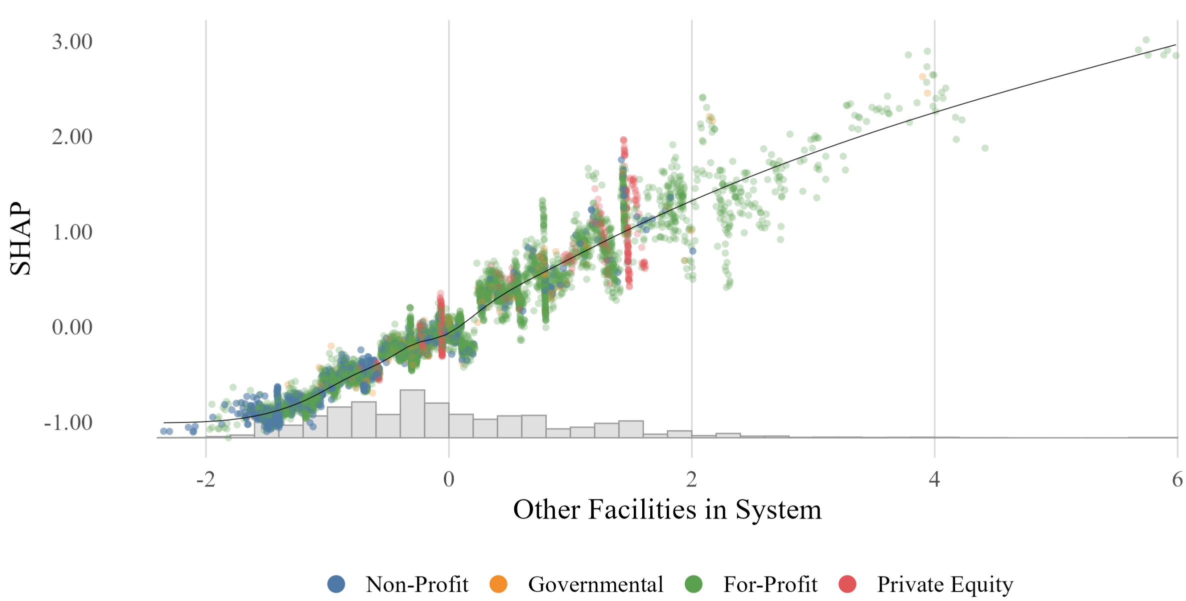


Table 1. Summary Statistics

This table presents summary statistics for our sample. The sample includes Medicare claims for skilled nursing facilities from January 1, 2016 to December 31, 2023. Data comes from the Skilled Nursing Facility Limited Data Set, which covers the universe of Medicare claims for skilled nursing. RUG-IV is an indicator variable equal to one for claims billed under the Resource Utilization Group IV era from January 1, 2016 to September 30, 2019. PDPM is an indicator equal to one for claims billed under the Patient Driven Payment Model from October 1, 2019 to December 31, 2023. "Unexplained" denotes that a variable was residualized on fixed effects for patient gender, age, race and diagnosis at the CCSR-level as recorded at the referring hospital. Ultra rehab days billed is the number of days for which therapy is billed at the highest level under RUG-IV. Coding intensity is computed only for PDPM era claims.

Panel A: Patient-Level Summary Statistics

	Mean	S.D.	P5	P10	P25	P50	P75	P90	P95
PDPM	0.48	0.50	0.00	0.00	0.00	0.00	1.00	1.00	1.00
RUG_IV	0.53	0.50	0.00	0.00	0.00	1.00	1.00	1.00	1.00
Total Days Billed	26.85	22.11	3.00	6.00	11.00	20.00	35.00	58.00	76.00
Ultra Rehab Days	15.61	18.95	0.00	0.00	0.00	12.00	24.00	40.00	55.00
Coding Intensity	1.05	1.09	0.00	0.00	0.00	1.00	2.00	3.00	3.00
Unexplained Ultra Days	0.00	18.54	-19.47	-16.87	-13.69	-4.06	8.24	23.58	37.68
Unexplained Cod. Inten.	0.00	0.86	-1.02	-0.86	-0.58	-0.26	0.45	1.22	1.67
Dietary Restriction	0.13	0.34	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Special Care High	0.34	0.47	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Low Function	0.25	0.43	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Depression	0.12	0.32	0.00	0.00	0.00	0.00	0.00	1.00	1.00
SLP High	0.21	0.41	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Observations	15,911,862								

Panel B: Facility-Level Summary Statistics

	Mean	S.D.	P5	P10	P25	P50	P75	P90	P95
Total Days Billed	27.66	7.65	16.34	19.45	23.00	27.02	31.88	37.27	40.80
Ultra Rehab Days	14.05	8.08	0.16	2.34	8.65	14.13	19.19	24.08	27.40
Coding Intensity	0.95	0.40	0.41	0.50	0.67	0.89	1.16	1.46	1.67
Unexplained Ultra Days	-1.49	8.00	-14.81	-12.97	-6.84	-1.37	3.63	8.52	11.66
Unexplained Cod. Inten.	0.00	0.50	-0.63	-0.51	-0.32	-0.08	0.24	0.61	0.91
Dietary Restriction	0.11	0.09	0.01	0.02	0.05	0.09	0.16	0.23	0.28
Special Care High	0.28	0.16	0.06	0.09	0.15	0.26	0.38	0.51	0.58
Low Function	0.25	0.11	0.10	0.13	0.18	0.24	0.31	0.38	0.43
Depression	0.09	0.13	0.00	0.00	0.01	0.05	0.11	0.24	0.37
SLP High	0.21	0.11	0.08	0.10	0.14	0.19	0.26	0.35	0.41
Observations	16,203								

Table 2. Excess Rehab and PDPM Billing

In this table, we examine the relationship between SNF system-level excess rehab during the Resource Utilization Group IV (RUG-IV) era (January 1, 2016-September 30, 2019) with billing intensity during the Patient Driven Payment Model (PDPM) era (October 1, 2019-December 31, 2023). We estimate an OLS regression of the form:

$$CodingIntensity_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt} \tag{11}$$

where $CodingIntensity_{ijt}$ is the sum of indicators for whether a patient is classified as Low Function, Special Care High, Depression, SLP High or Dietary Restriction and ranges from zero (least intense) to five (most intense). $ExcessRehab_j$ is the facility fixed effect for Ultra-High Rehab during RUG-IV, estimated from Equation 11. Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Robust standard errors are clustered at the SNF system level.

	(1) Code Intens.	(2) Code Intens.	(3) Code Intens.	(4) Code Intens.
Excess Rehab	0.0315***	0.0297***	0.0296***	0.0209***
	(36.03)	(34.44)	(34.55)	(29.79)
Quarter-Year FE	Yes	Yes	No	No
Patient Gender	No	Yes	Yes	Yes
Age Bucket	No	Yes	Yes	Yes
Patient Race	No	Yes	Yes	Yes
Diagnosis FE	No	Yes	No	No
County x Quarter FE	No	No	No	Yes
Diagnosis x Hospitalization Stay	No	No	Yes	Yes
Observations	6,500,296	6,500,290	6,488,504	6,484,773
Adjusted R^2	0.049	0.126	0.126	0.176

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table 3. Difference-in-Differences around PDPM Adoption

In this table, we explore the relationship between SNF system-level excess rehab during the Resource Utilization Group IV (RUG-IV) era (January 1, 2016-September 30, 2019) with billing intensity under PDPM. We estimate a difference-in-differences regression of the form:

$$y_{ijt} = \alpha + \beta HighRehab_j \times Post + \theta X_i t + \delta_t + \gamma_j + \epsilon_{ijt}$$

where y_{ijt} is a patient condition that qualifies for additional compensation under PDPM, but not under RUG-IV. $HighRehab_j$ is an indicator equal to one if a facility was in the highest tercile of excess rehab during RUG-IV according to Equation 11. $Acute$ is an indicator variable equal to one if patient i's primary diagnosis belongs to the Acute Neurologic clinical category, SLP is an indicator variable equal to one if at least one SLP-related comorbidity is included among patient i's diagnoses codes, and NTA is a weighted sum of NTA-related comorbidities. $Post$ is an indicator variable for admissions after PDPM becomes effective in October 2019. To control for time-invariant differences in propensity across facilities, we include an individual facility fixed effect in all specifications. Robust standard errors are clustered at the SNF system level and t statistics are reported in parentheses.

	(1) Acute	(2) SLP	(3) NTA	(4) Acute	(5) SLP	(6) NTA
High Rehab × Post	0.0533***	0.0240***	0.0892***	0.0383***	0.0173***	0.0769***
	(22.19)	(13.60)	(13.89)	(16.90)	(12.06)	(12.23)
Patient Gender	No	No	No	Yes	Yes	Yes
Age Bucket	No	No	No	Yes	Yes	Yes
Patient Race	No	No	No	Yes	Yes	Yes
County x Quarter FE	No	No	No	Yes	Yes	Yes
Diagnosis x Hospitalization Stay	No	No	No	Yes	Yes	Yes
Facility FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-Year FE	Yes	Yes	Yes	No	No	No
Observations	13,674,132	13,674,132	13,674,132	13,657,871	13,657,871	13,657,871
Adjusted R^2	0.046	0.031	0.063	0.274	0.140	0.188

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table 4. Instrumental Variables-Second Stage

In this table we examine the effect of excess rehab billing on PDPM billing intensity. We estimate a 2SLS regression of the form:

$$CodingIntensity_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt}$$

where y_{ijt} is a measure of billing intensity for patient i at facility j during quarter t and is an indicator for a patient being classified as either Low Function, Special Care High, Depression, SLP High or Dietary Restriction. We instrument for $ExcessRehab_j$ using $ExpectedFacilityRehab$ from Equation 5. Low Function, Special Care High, and Depression are classifications that impact the Nursing component of PDPM case-mix and increase reimbursement. SLP High and Dietary Restriction are classifications that increase PDPM reimbursements for the SLP component (as shown in Exhibits 1 and 2). Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Hospitalization length refers to the number of days for which a patient was hospitalized at an inpatient facility prior to beginning their SNF stay. Robust standard errors are double clustered at the SNF system and HSA level.

	(1) Low Func.	(2) S.C.H.	(3) Depression	(4) SLP High	(5) Diet	(6) Code Intens.
Excess Rehab	0.00360	0.0149**	0.0129***	0.00556***	0.00530*	0.0423***
	(1.14)	(2.51)	(2.63)	(2.61)	(1.85)	(3.18)
Patient Gender	Yes	Yes	Yes	Yes	Yes	Yes
Age Bucket	Yes	Yes	Yes	Yes	Yes	Yes
Patient Race	Yes	Yes	Yes	Yes	Yes	Yes
Diagnosis x Hospitalization Length	Yes	Yes	Yes	Yes	Yes	Yes
HSA FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,536,843	5,536,843	5,536,843	5,536,843	5,536,843	5,536,843
Kleibergen Paap F-Stat	52.34	52.34	52.34	52.34	52.34	52.34

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table 5. County Demographics, Facility Characteristics and PDPM Upcoding

This table explores the relationship between facility-level Coding Intensity and various measures of patient demographics, corruption, social capital and facility characteristics. All demographic data comes from the American Community Survey 5-year estimates and is measured in the 2015 year. Public Corruption is the number of public corruption convictions per million residents within each judicial district as in Griffin et al. (2025). Religious Affiliation is the percentage of a county's population with a religious affiliation. Religious affiliation data comes from the Association of Religious Data Archives. Ahsley Madison Useage is the percent of a county's population with a paid Ashley Madison account. Data for Ashley Madison usage comes from Griffin et al. (2019). Coefficients are estimated using ordinary least squares and counties are weighted by the number of skilled nursing visits.Facility size is the log number of patients treated by each facility. System Size is the log number of patients treated by all facilities controlled by an SNF system. Enforcement is the number of DOJ Medicare enforcement activities per 100,000. Competition is the county-level HHI. County-average is the average coding intensity by other facilities within a county, excluding facilities operated by the same SNF system. Other facilities in the system is the average coding intensity of other facilities within an SNF system.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Enforcement	-0.0921					-0.0619	-0.0281	-0.0531	-0.0623	-0.0476
	(-1.45)					(-1.34)	(-0.62)	(-1.30)	(-1.60)	(-1.38)
Competition		-0.0925				-0.0467	-0.0719	-0.0399	-0.0616	-0.0448
		(-1.11)				(-0.86)	(-1.15)	(-0.96)	(-1.43)	(-1.33)
System Size			0.215**			0.0861	0.0873	0.0865	0.0878	0.0890
			(2.52)			(1.57)	(1.59)	(1.57)	(1.58)	(1.61)
Facility Size			0.178*			0.101**	0.101*	0.103**	0.100*	0.104*
			(2.00)			(2.04)	(1.96)	(2.05)	(2.00)	(1.93)
County-Average				0.155		0.137**	0.115***	0.130***	0.132***	0.0868***
				(1.59)		(2.60)	(2.78)	(2.94)	(3.05)	(3.14)
Other Fac. in System					0.755***	0.715***	0.715***	0.715***	0.716***	0.713***
					(20.66)	(20.13)	(20.47)	(20.07)	(19.99)	(20.29)
Log Pop. Density							0.0929***			0.0986**
							(3.18)			(2.12)
Log Median Income							0.0406			0.0399
							(0.70)			(0.57)
Poverty Rate							0.0337			0.0319
							(0.91)			(0.68)
Unemployment Rate							-0.108			-0.0833
							(-1.56)			(-1.18)
Pct. College Educated							-0.164***			-0.158**
							(-3.26)			(-2.63)
Non-White							-0.00260			-0.00472***
							(-1.59)			(-3.08)
Public Corruption								0.0229		0.0424*
								(0.69)		(1.74)
Religious Affiliation								0.0804		0.0795*
								(1.24)		(1.97)
Ashley Madison Use								-0.0393**		0.0178
								(-2.07)		(0.26)
Clustering									0.125***	0.102***
									(3.51)	(3.34)
Support Ratio									-0.0213	-0.0892*
									(-0.88)	(-1.73)
Volunteer Rate									-0.0600	0.00637
									(-0.97)	(0.19)
Civic Organization									-0.0430	-0.0145
									(-1.00)	(-0.57)
Economic Connect.										-0.0694*
										(-1.79)
State Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,397	7,328	7,328	7,328	7,328	7,328	7,328	7,328	7,318	7,314
Adjusted R^2	0.179	0.215	0.274	0.219	0.541	0.558	0.560	0.558	0.558	0.561

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Supplemental Internet Appendix

8. PDPM Reimbursement Components

8.1. Nursing Component

The Nursing component of PDPM is determined by a patient's underlying diagnosis, intensity of services needed, physical function score, and depression status. Within broad hierarchical categories reflecting the extensiveness of required services, patients are assessed for depression. A final case-mix is assigned using a patient's level of physical function as shown in Exhibit 2. Patients with more intensive nursing needs, depression, or lower levels of physical function are eligible for higher daily reimbursements for the Nursing component. Daily reimbursements for the Nursing component (in 2022) range widely from $68.28 to $420.05.

When considering coding intensity, we first measure whether a patient is classified in the treatment category of Special Care High. To qualify for this level of care, a patient must have a Function Score of 14 or less and suffer from a specified serious medical condition (such as septicemia, daily respiratory therapy, comatose, or fever with additional symptoms). Patients classified into one of the Special Care High case-mixes are shown in Exhibit 2 by the pink boxes and arrows. The Special Care High case-mix groups have among the highest daily reimbursement amounts.[58]

Patients are then screened by skilled nursing staff for signs of depression using the Resident Mood Interview or Staff Assessment of Resident Mood. Nursing case-mixes with a depression diagnosis are denoted by the brown boxes and arrows on Exhibit 2. The final step in determining reimbursement for the Nursing component is classifying patients into distinct categories based on the level of physical function. Patients are scored on a scale of zero to 16 based on their ability to complete tasks involving mobility and self-care such as ability to get in and out of bed, sit/stand, and oral/toilet hygiene with a higher score indicating that a patient can complete more tasks independently. Patients who can complete fewer tasks on their own are eligible for higher daily reimbursements. For expositional purposes, we refer to case-mixes with a Nursing Function Score of five or below the lowest category, as "Low Function" (denoted by the turquoise boxes and arrows in Exhibit 2). Patients classified as Low Function need substantial support in every category of assessment and are unable to complete basic self-care and mobility tasks on their own.

[58] Extensive Services case-mix groups have higher daily reimbursements, but have extremely strict requirements to classify such as having an active Tracheostomy or Ventilator and are easier to verify.

Differences in reimbursement can be substantial—a patient qualifying for Special Care High with depression and low mobility would qualify for daily reimbursement of \$248.30 whereas a patient without such condition would qualify for a maximum daily reimbursement of \$147.95.

8.2. SLP Component

The Speech Language Pathology (SLP) component of reimbursement features 12 possible case-mix indices with daily reimbursement rates ranging from \$15.06 to \$93.25 per day as demonstrated in Exhibit 1. Patients are first screened for having an Acute Neurologic primary diagnosis, additional SLP-related comorbidities, and cognitive impairment. SLP reimbursement is increasing in the count of such conditions. SLP-related comorbidities are conditions that increase the cost of SLP-related care and include traumatic brain injuries, oral cancers, or speech and language deficits. Cognitive impairment is determined by staff and any level of cognitive impairment that is mild or above qualifies for higher reimbursement rates. For expositional purposes, we consider "SLP High" case-mixes as those in which a patient is billed for at least two of the three qualifying conditions and highlight such case-mix indices using yellow boxes and arrows in Exhibit 1.

Once patients have been sorted into an initial bucket, a SLP case-mix is assigned based on the presence or absence of dietary disorders. Patients qualify for a higher SLP reimbursement if they require a mechanically altered diet, have a swallowing disorder, or both. Swallowing disorders include symptoms such as coughing, choking, or experiencing pain while swallowing. We classify a patient as having a dietary restriction if they are billed as having both a mechanically altered diet and a swallowing disorder. Dietary Restriction case-mixes are illustrated in Exhibit 1 by the light blue boxes and arrows. Patients coded as having at least two of an Acute Neurologic primary diagnosis, SLP-related comorbidities or cognitive impairment along with a dietary restriction would qualify for a daily SLP reimbursement of at least \$78.19 compared to a maximum reimbursement of \$32.34 without such conditions.

Exhibit 1. Nursing Case-Mix

This diagram demonstrates how patients are sorted into nursing case-mixes for the Nursing component of PDPM. Patients are first sorted into six buckets in a hierarchical order based on underlying patient conditions with earlier classifications qualifying for higher daily reimbursement. Patients are then sorted based on the presence of a tracheostomy or ventilator, depression, or intensity of nursing rehab. Finally, patients are sorted into a final case-mix using their functional score on section GG. Special Care High denotes patients that are classified into the Special Care High classification (pink bucket), Depression denotes that a patient was billed for depression (brown boxes), and finally Low Function means that a patient was billed for a functional score of five or below (turquoise boxes). Daily reimbursement for each case-mix is displayed in the further right column.

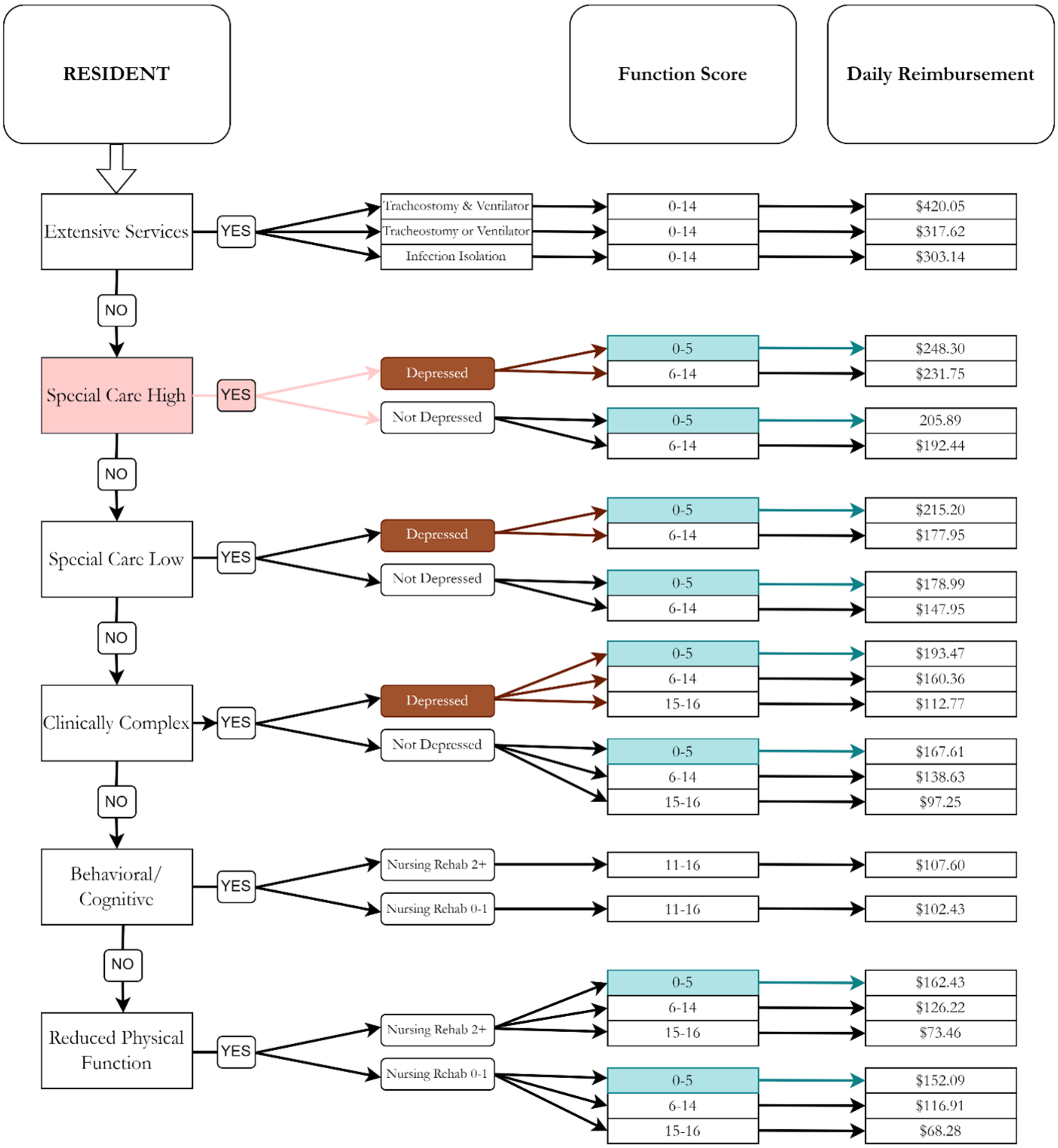


Exhibit 2. Speech Language Pathology (SLP) Case-Mix

This exhibit displays how patients are sorted into possible case-mixes for the Speech Language and Pathology component of PDPM care. Patients are first sorted by whether they have an Acute Neurologic primary diagnosis, an SLP-related comorbidity, or cognitive impairment. We classify a patient at SLP High if they have at least two of the conditions as denoted by the yellow boxes. Patients are then sorted based on dietary restrictions including whether a patient has a swallowing disorder or a mechanically altered diet. We classify a dietary restriction if a patient has both a swallowing disorder and a mechanically altered diet as denoted by the blue boxes. The right-most column displays how much reimbursement each case-mix is entitled to.

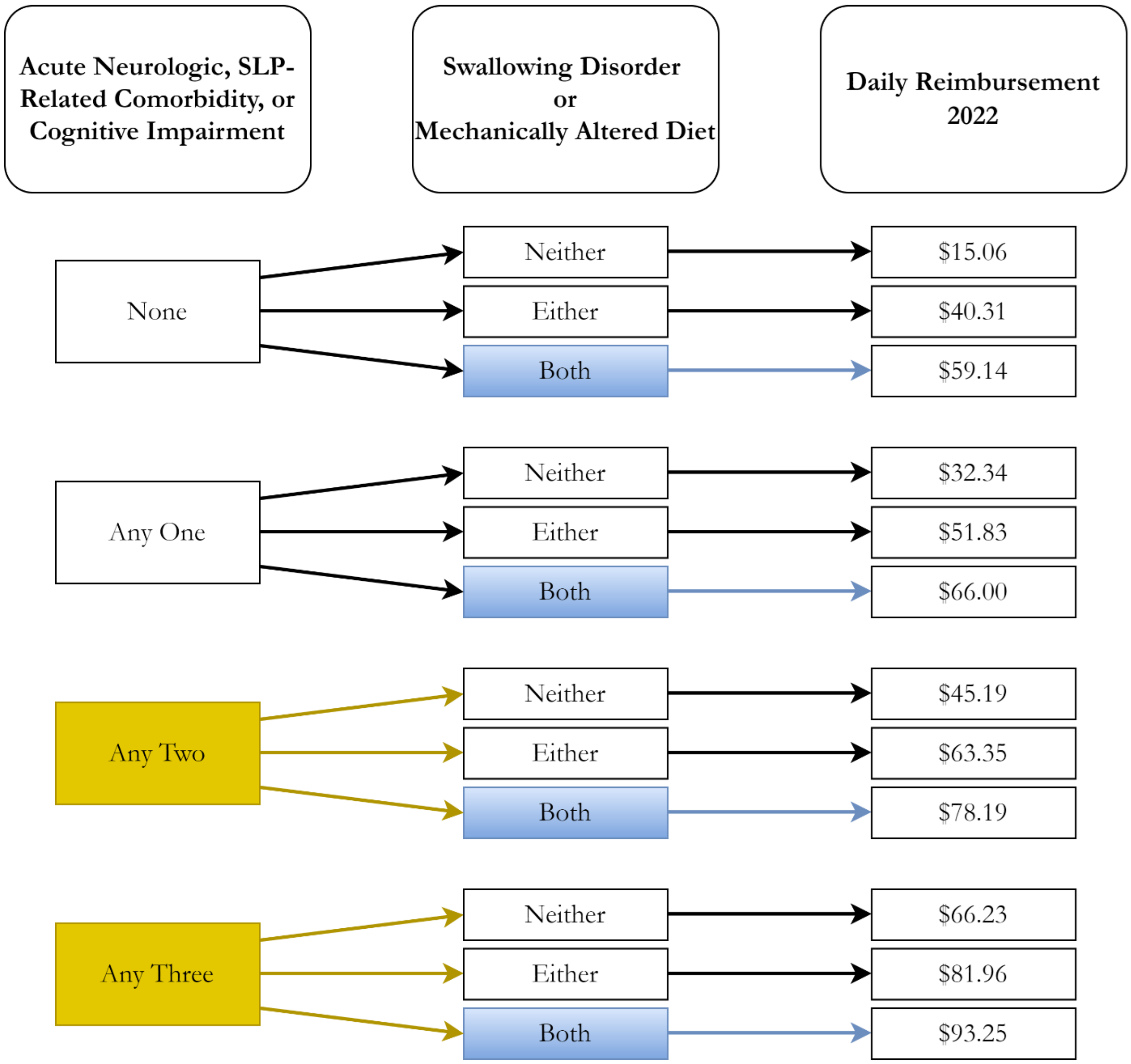
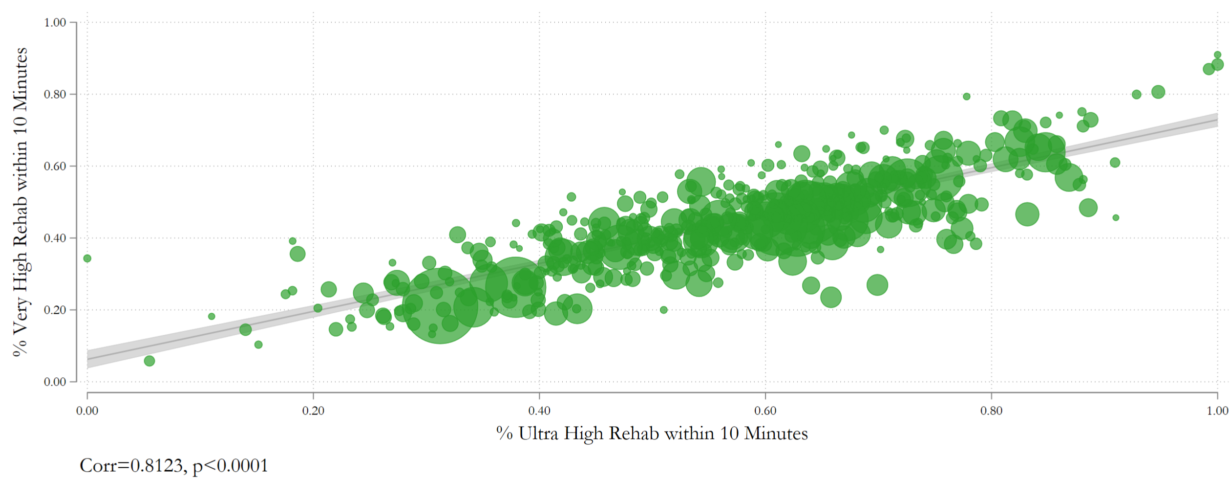


Figure IA.1. Excess Rehab and Threshold Treatment

This figure explores the relationship between system excess rehab and the propensity to treat within a tight window around the cutoff. Panel A displays a scatter plot between a system's excess rehab and either the proportion of all days billed between 720-729 minutes (left subgraph) or the proportion of all days billed more than 730 minutes (right subgraph). Each observation is an SNF system, and the size corresponds to the number of patients treated by each system. Panel B presents a scatter plot of the proportion of Ultra-High rehab patients billed within 10 minutes of the threshold against the proportion of Very-High rehab patients billed within 10 minutes of the threshold (500-509 minutes per week) by system. Finally, Panel C displays the distribution of Ultra-High rehab care by system type and whether care falls within 10 minutes of the cutoff.

Panel A. Relationship Between Excess Rehab and Ultra-High Billing by Threshold

Panel B. Facility-Level Billing Across Threshold



Panel C. Threshold Billing by Facility Excess Rehab

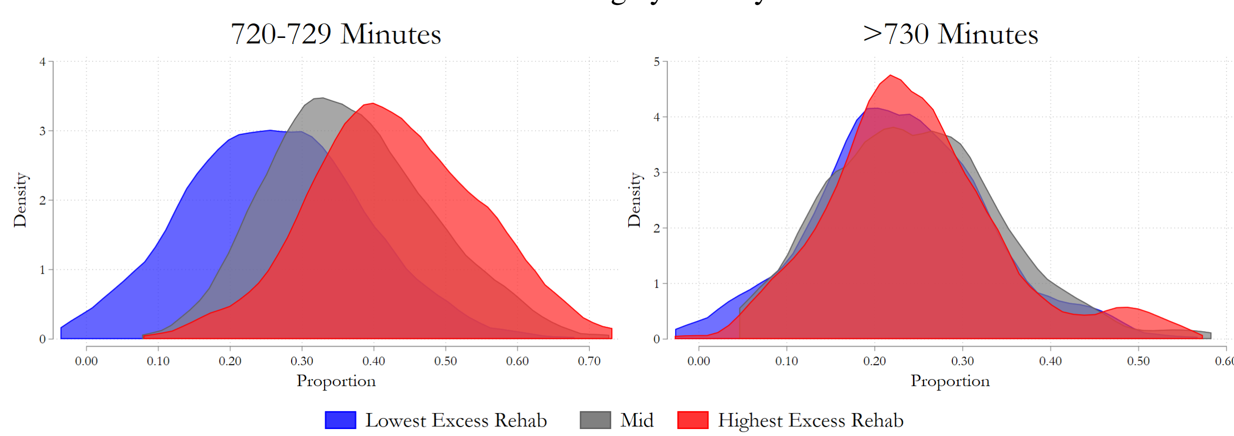


Figure IA.2. Excess Rehab and Threshold Treatment

This figure explores the useage of differing therapy categories by facility type. Facilities are first sorted into deciles based on the abnormal excess rehab estiamted using Equation 11. This figure plots the proportion of patients receiving therapy within a range of 511-719 minutes (e.g. at the Very High level) as well as the proportion of patient days receiving just enough therapy to qualify for Ultra-High case mixes, defined as between 720-729 minutes of therapy per week.

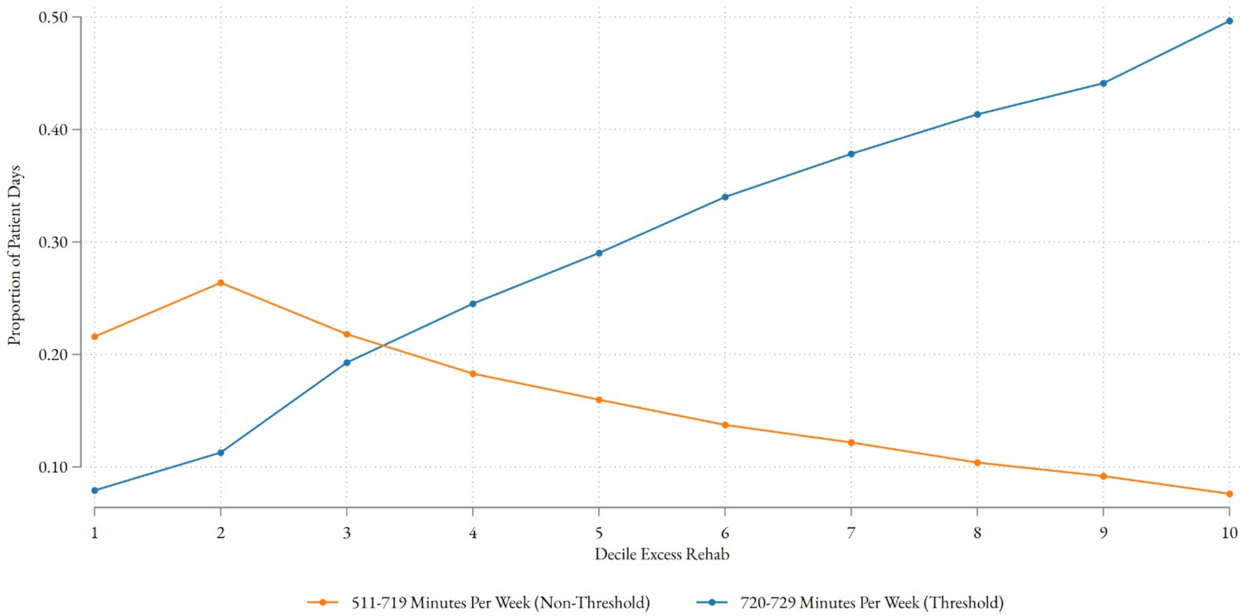
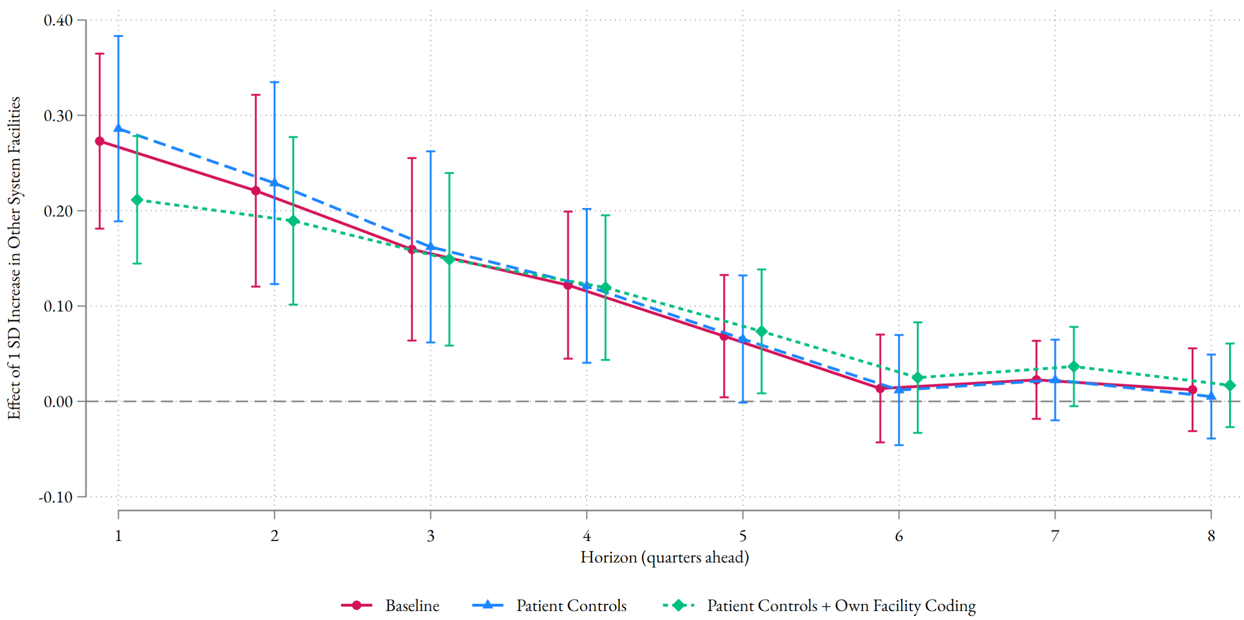


Figure IA.3. Diffusion of System Practices by Facility Location

This figure explores within-system diffusion of coding practices. Panel A explores the predicted effect on future upcoding for a one standard deviation increase in contemporaneous upcoding at other facilities within an SNF system located in a different state up to eight quarters out. Panel B explores the predicted effect on future upcoding for a one standard deviation increase in contemporaneous upcoding at other facilities within an SNF system which are located within the same state up to eight quarters out. Standard errors denote 95% confidence intervals.

Panel A: Predictive Regressions of Out-of-State System Practices on Future Upcoding



Panel B: Predictive Regressions of In-State System Practices on Future Upcoding

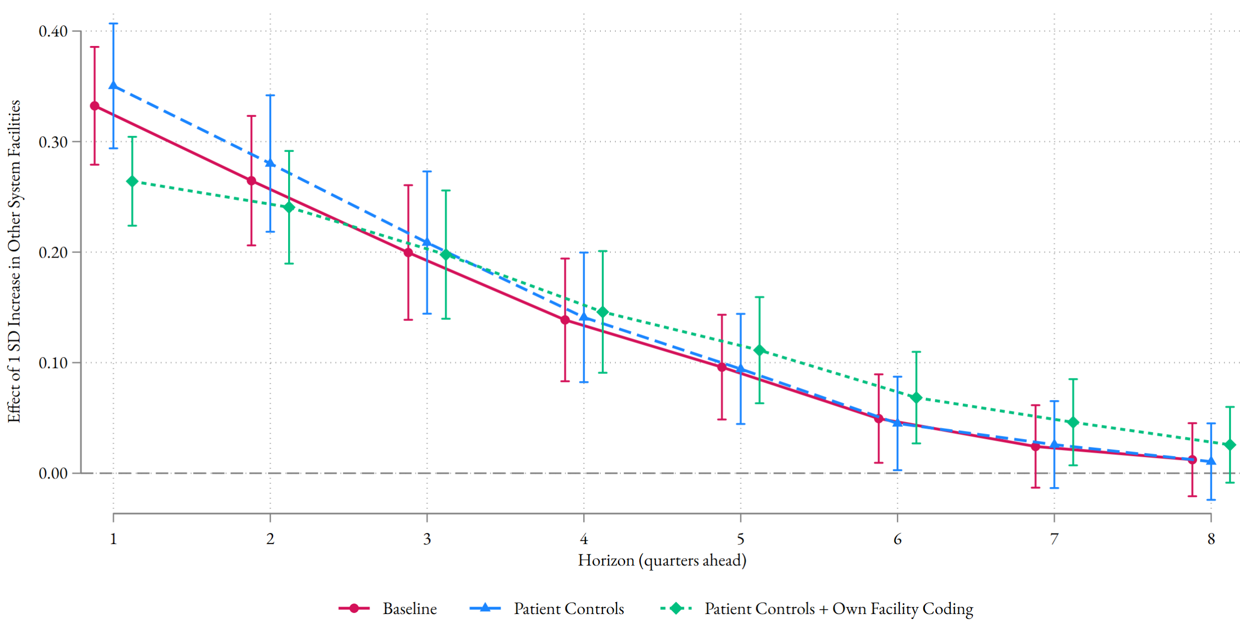


Figure IA.4. Diffusion of System Practices by Facility Location-Distance Thresholds

This figure explores within-system diffusion of coding practices. This figure plots the predicted effect on future upcoding for a one standard deviation increase in contemporaneous upcoding at other facilities within an SNF system located different distances away for up to eight quartrs in the future. Standard errors denote 95% confidence intervals.

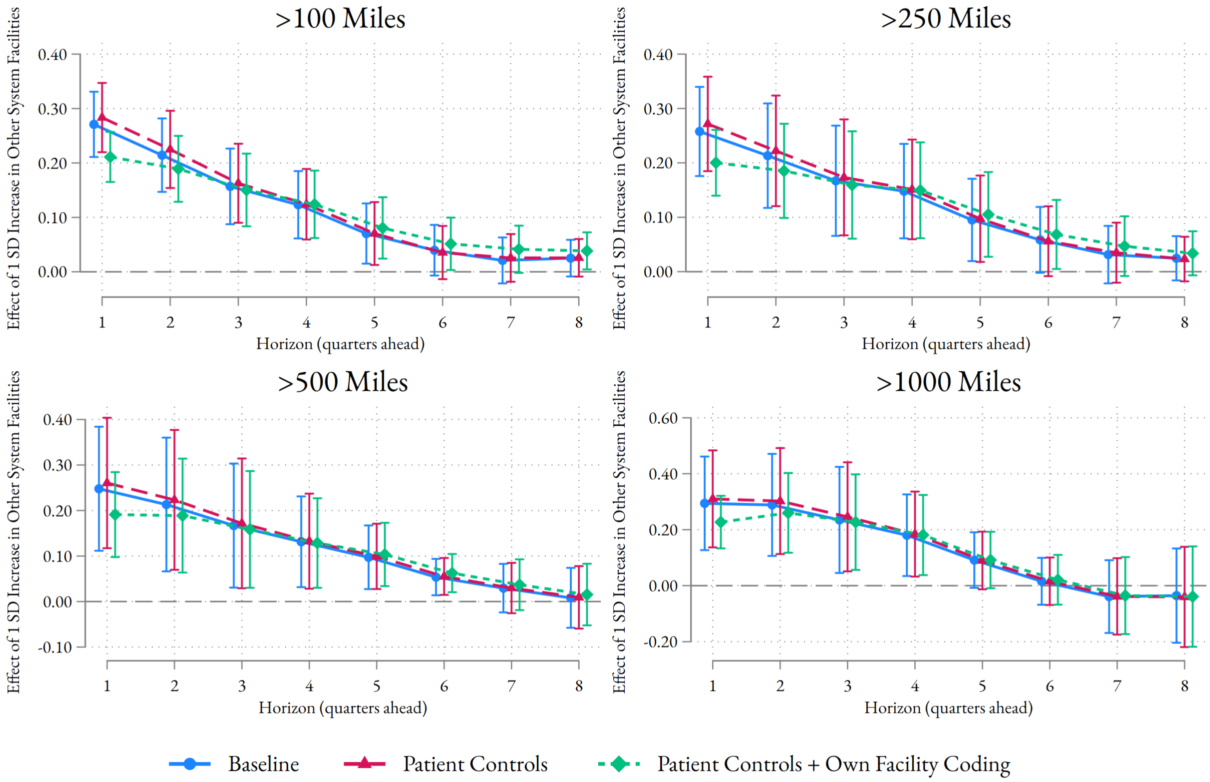


Figure IA.5. Facility Rehab and Compensating Comorbidities Times-Series

This figure displays the undajusted levels of compensating comorbidities by a facility's level of RUG-IV rehab.

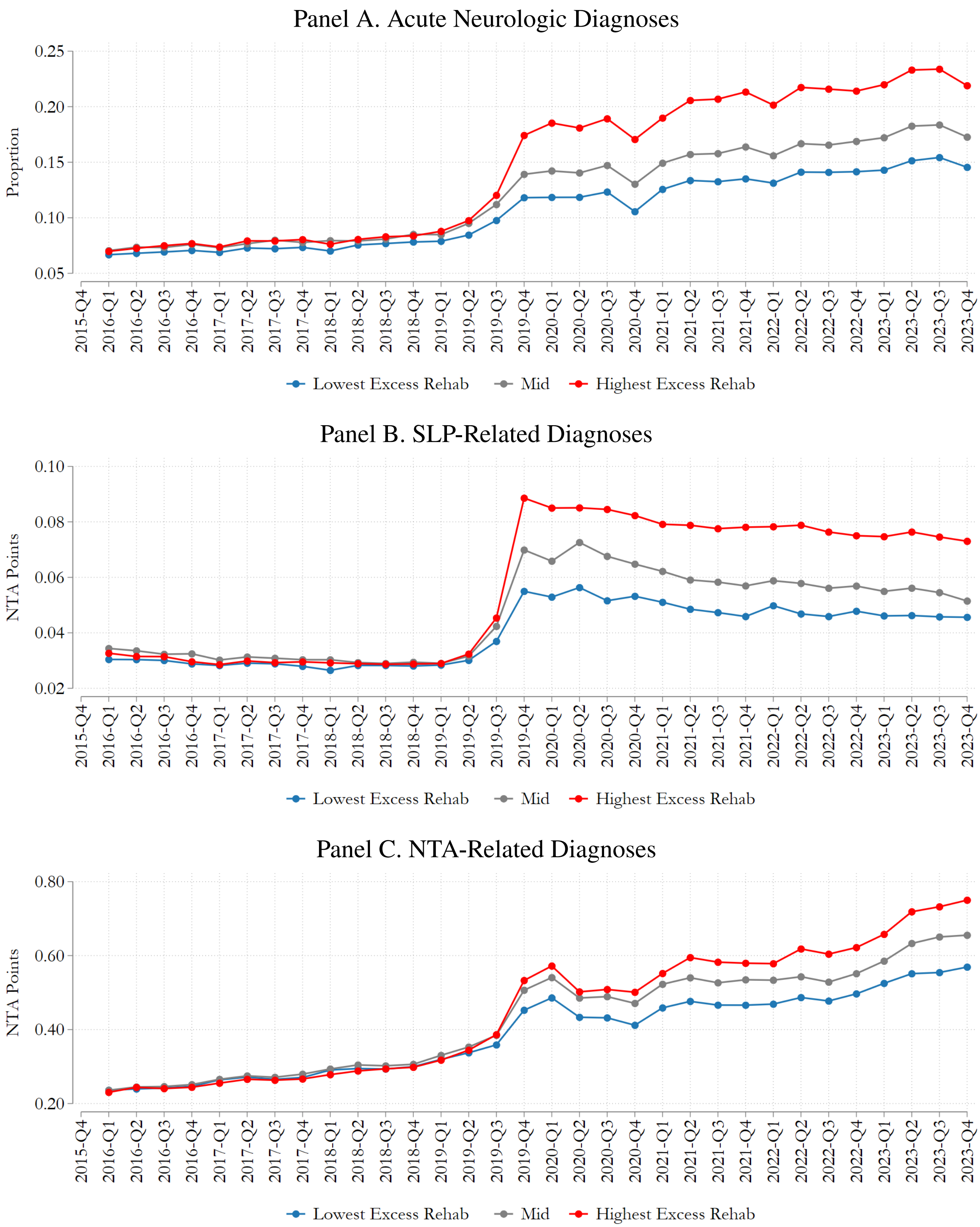
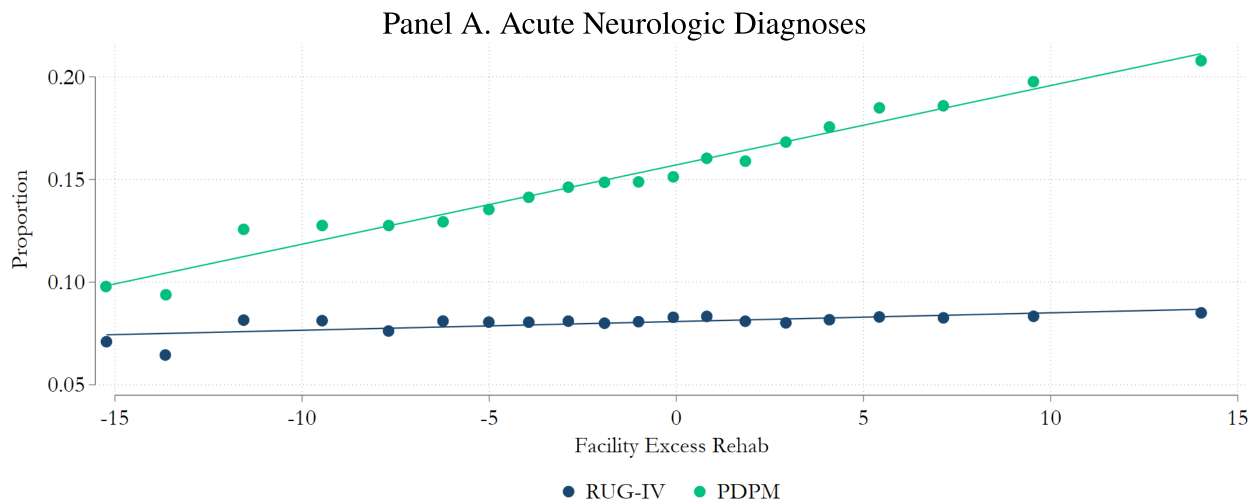
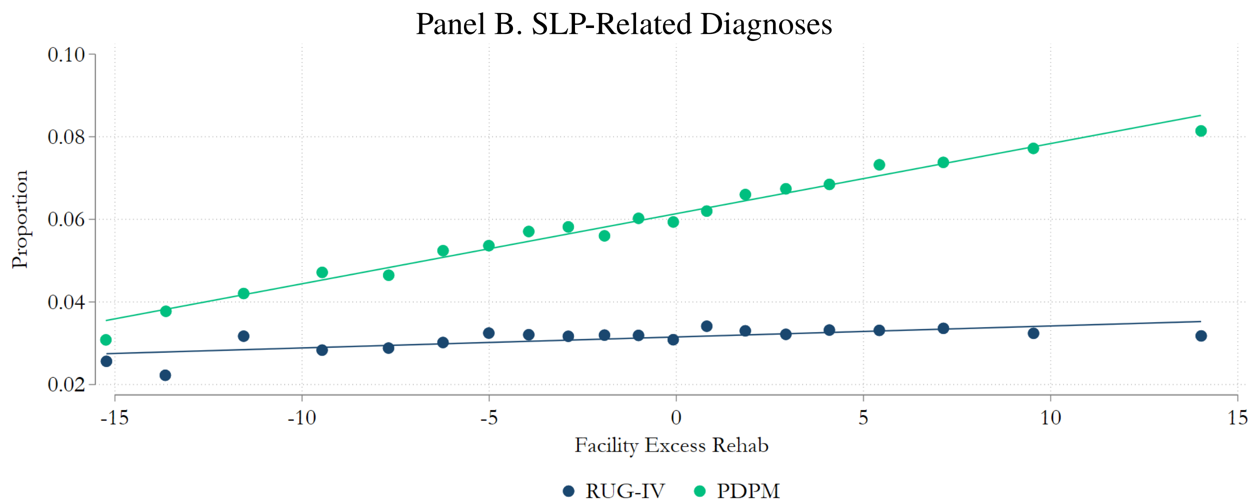


Figure IA.6. Facility Rehab and Compensating Comorbidities Times-Series Binscatter

This figure displays the undajusted levels of compensating comorbidities by a facility's level of RUG-IV rehab.





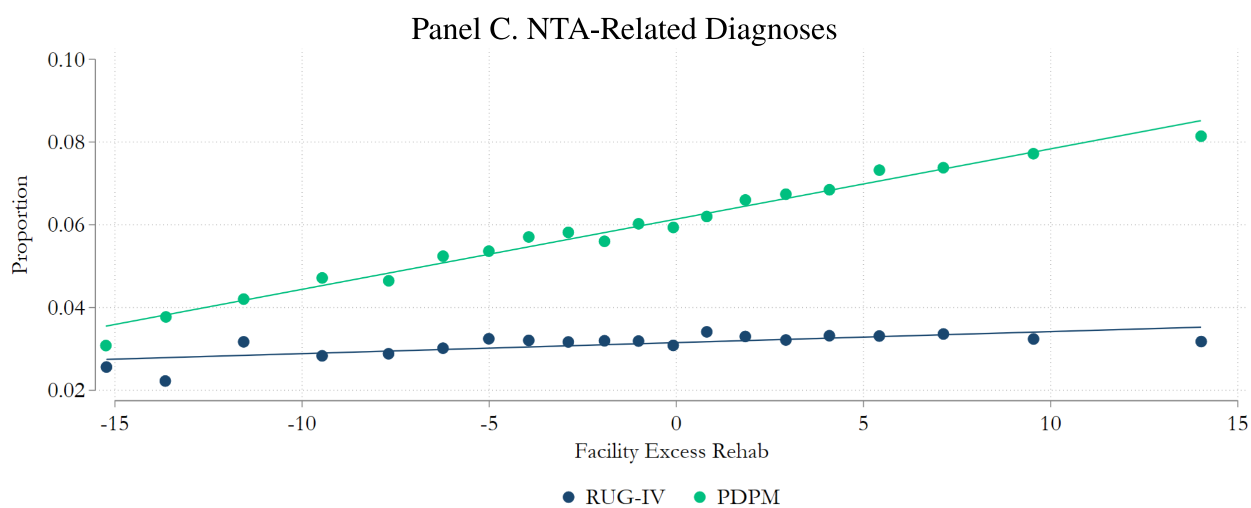
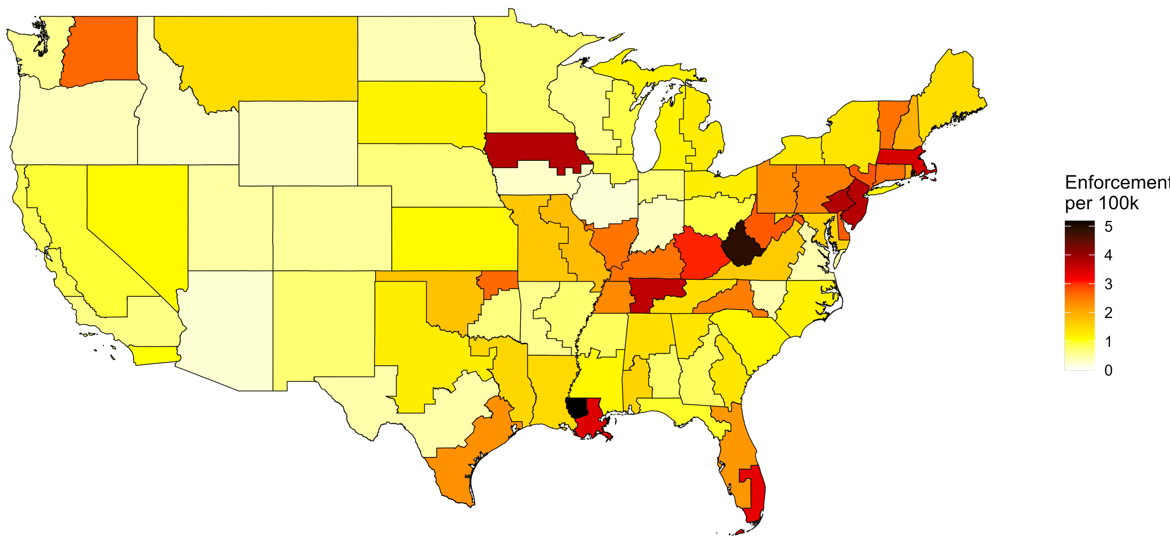


Figure IA.7. DOJ Medicare Enforcement Activity

Notes: This figure plots the level of DOJ enforcement activity by judicial district. Enforcements come from the DOJ Press Release Archives. To focus on relevant enforcement actions, we require that press releases mention Medicare. The number of enforcement actions are normalized by the population of the judicial district. Panel B displays facility-level Coding Intensity within each judicial district. The x-axis denotes the normalized enforcement activity.

Panel A: DOJ Enforcement Activity by Judicial District



Panel B: Enforcement Activity and Coding Intensity

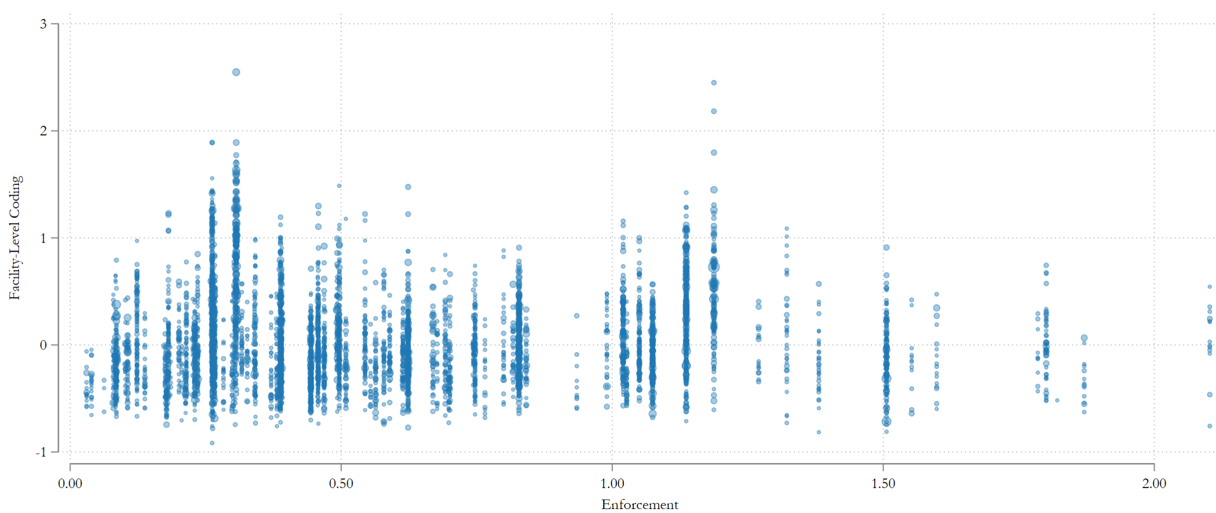


Figure IA.8. SNF Billing Practices and Cost of Patient Care

This figure explores the costs of skilled nursing care per patient stay. We compute this amount as the total of the Medicare allowed amount which includes the sum that Medicare pays, deductible or coinsurance amounts that the beneficiary is responsible for, and any amount owed by a third party divided by the total count of SNF stays provided. Dollar amounts are adjusted for inflation. SNF systems are sorted into deciles of excess rehab which is defined according to Equation (1). The blue line denotes revenue during RUG-IV while the orange line denotes revenue during PDPM. The shaded areas denote 95% confidence intervals.

Panel A. SNF Total Revenue by Decile Excess Rehab

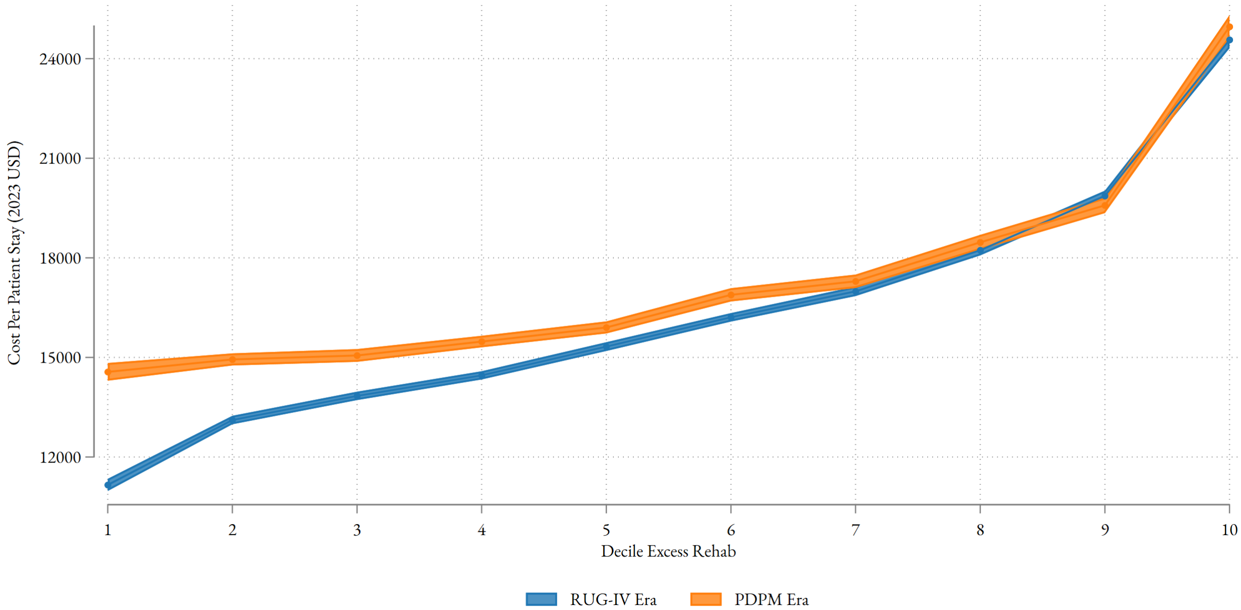


Figure IA.9. Use of Ultra-High Rehab for Expiring Patients

This figure explores the relationship between system excess rehab and the provision of Ultra-High rehab care for patients who die at a skilled nursing facility. The y-axis is the proportion of patients who received Ultra-High rehab care for every day that they were in the skilled nursing facility. The x-axis is a system's excess rehab as defined by Equation 11. The line denotes a binscatter that has been fit to the data. The *t* statistic is presented in the bottom left corner and standard errors are clustered at the SNF system level.

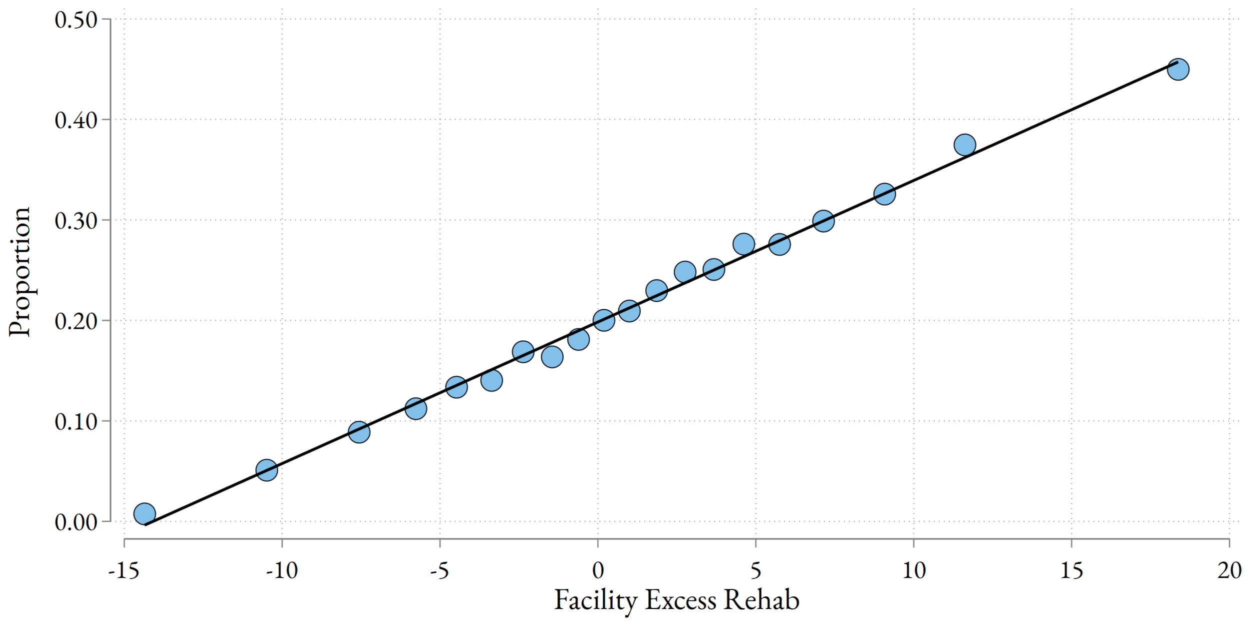


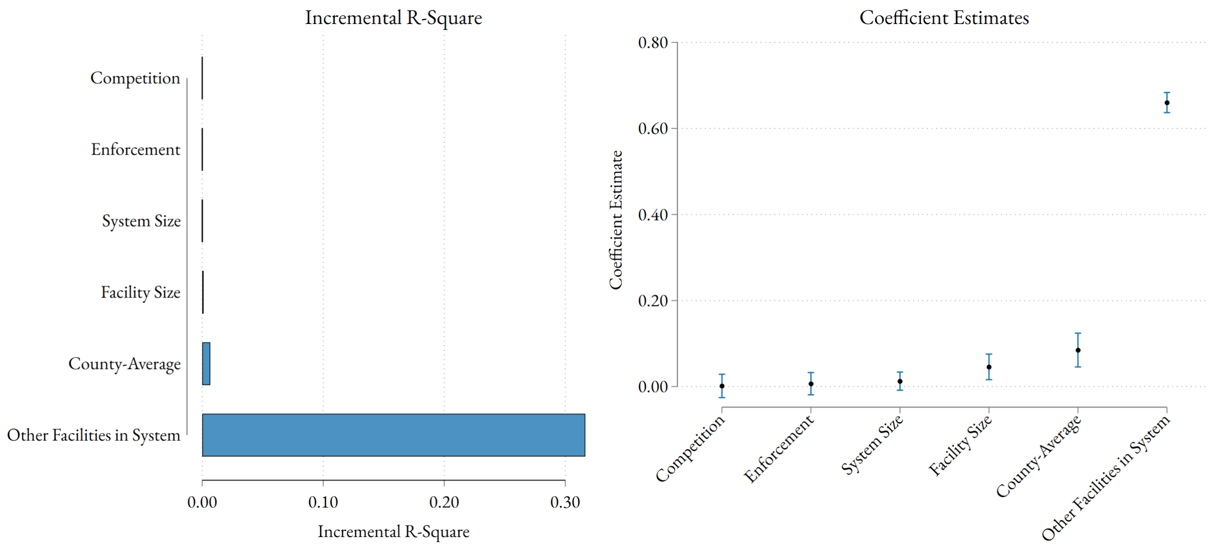
Figure IA.10. OLS Determinants of Facility-Level Fraud

This figure explore the role of potential determinants in explaining cross-sectional variation in coding intensity across skilled nursing facilities. Coding Intensity is residualized on patient characteristics and aggregated up to the individual facility. We then consider cross-sectional facility-level regressions of the form:

$$ResidCodeIntens_{jm} = \alpha + \theta X_{jm} + \epsilon_{jm} \tag{12}$$

where X_{jm} is a facility-level characteristic. The six variables considered are Enforcement, County Concentration (HHI), Facility Sizy, System Size, County Average Upcoding, and the Coding rates at other faciltiies within the system. Panel A plots the incremental R-Squared from adding each additional candidate determinant (left subgraph) as well as the estimated coefficient on each variable (right subgraph) along with 95% confidence intervals. For ease of comparison, Coding Intensity and each determinant have been standardized. Panel B plots a binscatter illustrating the slope between Coding Intensity and each determinant.

Panel A: Determinants of Facility-Level Billing



Panel B: Relationship Between Facility Billing and Candidate Explanatory Variables

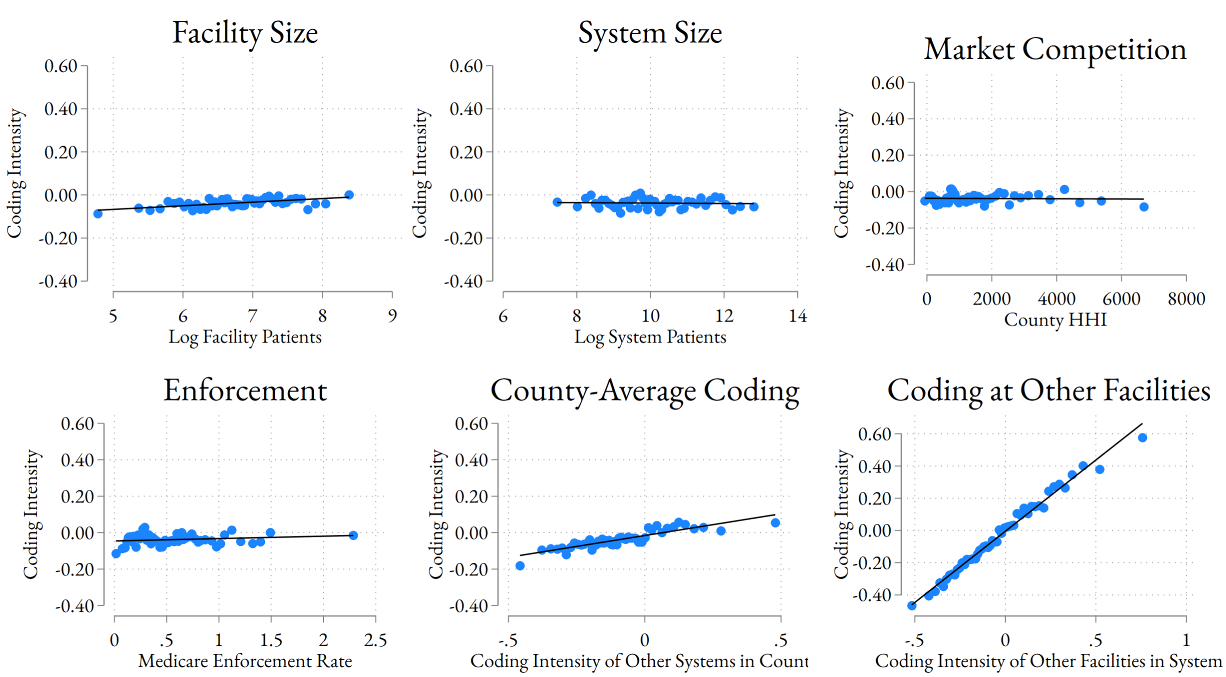


Figure IA.11. Excess Rehab and PDPM Coding Intensity by System

This figure explores billing practices under the RUG-IV and PDPM billing regimes. Each scatter plot denotes a patient diagnosis at the CCSR level. The x-axis displays the mean number of Ultra-High rehab days (left panel) or PDPM coding intensity (right panel) by opportunistic systems. Opportunistic SNF systems are those in the top tercile of excess rehab as measured using Equation 11. The y-axis displays the mean number of Ultra-High rehab days (left panel) or PDPM coding intensity (right panel) that patients with the same inpatient diagnosis receive at other facilities. Red denotes diagnosis groups for which patients receive higher Ultra-High rehab or higher coding intensity at opportunistic systems.

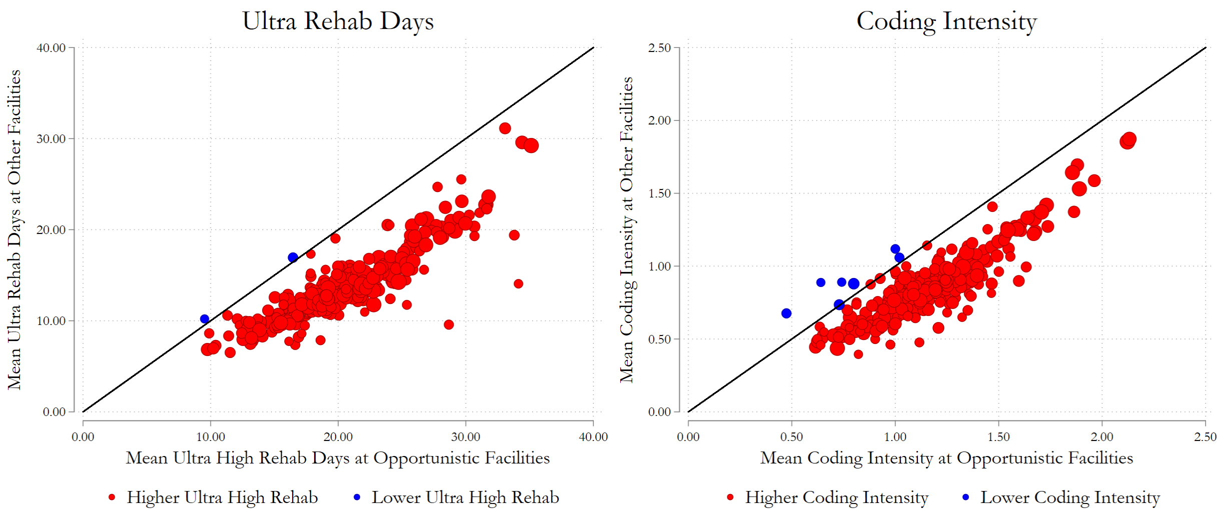
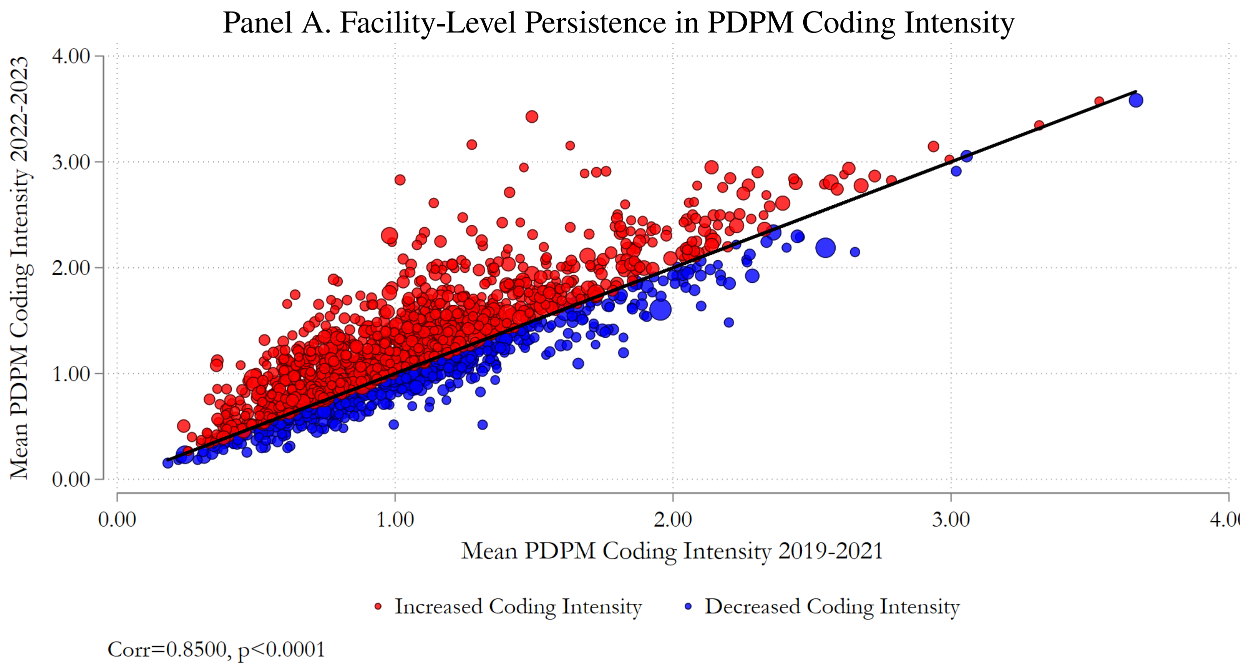


Figure IA.12. Persistence and Time-Series Variation in PDPM Coding Intensity

This figure explores the persistence and variation in PDPM coding intensity over time. $CodingIntensity$ is the sum of indicators for whether a patient is classified as Low Function, Special Care High, Depression, SLP High or Dietary Restriction. $CodingIntensity$ ranges from zero to five and a higher number denotes greater billing intensity. Panel A explores the persistence of PDPM coding intensity at the skilled nursing facility level by plotting the average PDPM coding intensity for the skilled nursing facility from 2019-2021 on the x-axis against the average coding intensity from 2022-2023 on the y-axis. The size of the markers denotes the facility size. Red markers denote facilities that experienced an increase in coding intensity during the second half of the sample. The correlation coefficient is presented at the bottom left. Panel B explores the time-series variation in coding intensity. Coding intensity is residualized on fixed effects for patient gender, age, race, and diagnosis at the CCSR-level as documented at the referring hospital. Skilled nursing facilities are sorted into terciles based on their excess rehab from January 1, 2016-September 30, 2019.



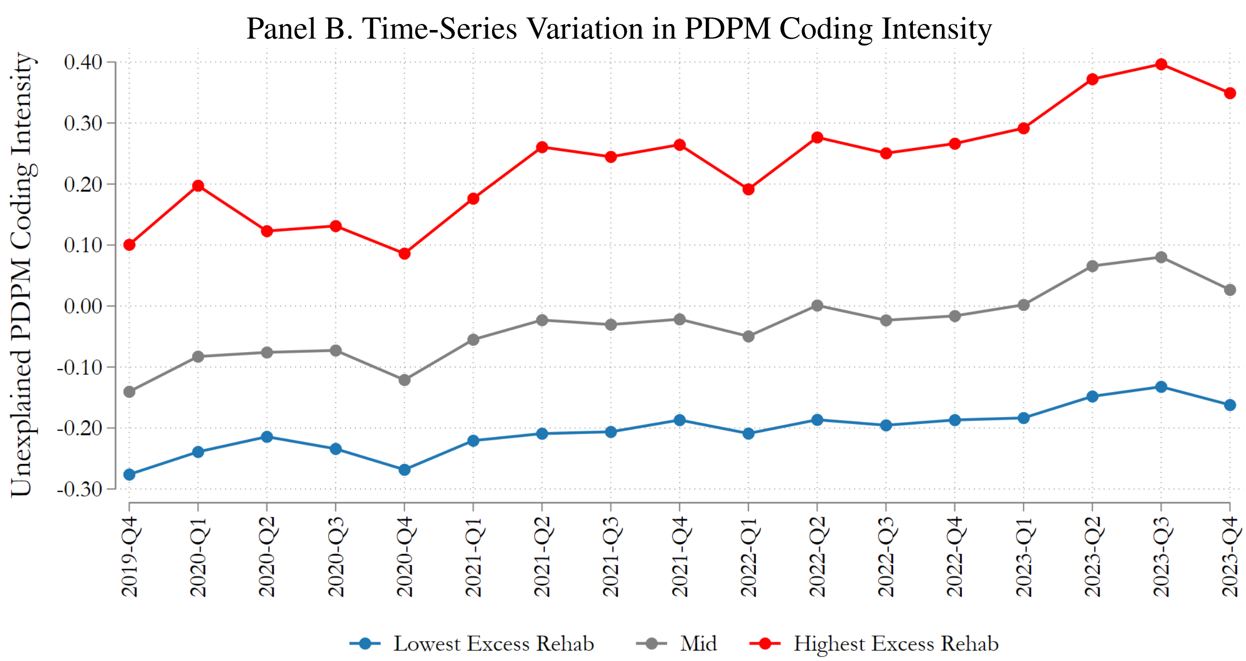


Figure IA.13. Change in Billing Intensity around Acquisition by Individual Category

This figure investigates the incidence of billing intensity around acquisition by Opportunistic Systems. Specifically, this figure displays coefficients from the following dynamic stacked cohort difference-in-difference regression:

$$y_{ijct} = \alpha + \sum_{t=1}^{T} \beta_t Period_t \times AcquiredHighRehab_j + \Gamma_{jc} + \delta_{tc} + \epsilon_{ijt}$$

where Y_{ijct} is an indicator for each individual billing category patient i receives at facility j. $AcquiredHighRehab_j$ is an indicator variable equal to one if a facility was acquired by an opportunistic SNF system, defined as systems in the highest tercile of excess rehab during RUG-IV. $Period_t$ is an indicator denoting time relative to acquisition. Confidence intervals denote 95% levels. Γ_{jc} is a facility by cohort fixed effect and δ_{tc} is a quarter by cohort fixed effect.

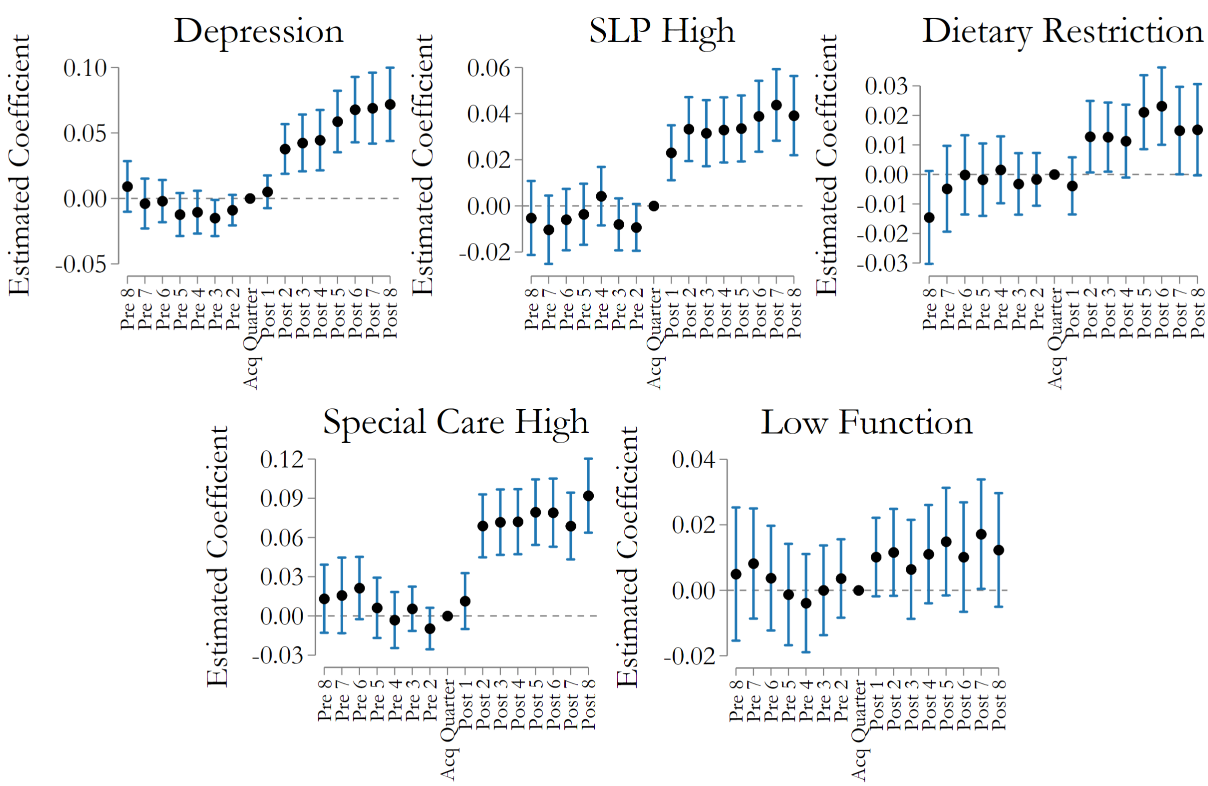


Figure IA.14. Competitor Occupancy and Probability of Admission

This figure visually demonstrates the effect of occupancy constraints on facility selection. Competitor occupancy is the occupancy rates of other facilities within the same Hospital Service Area (HSA) in the month prior to SNF admission. The y-axis shows the probability that a patient i chooses a given facility j within HSA h. All facilities located within the same HSA are considered as candidate facilities. The data is fit using a local polynomial specification.

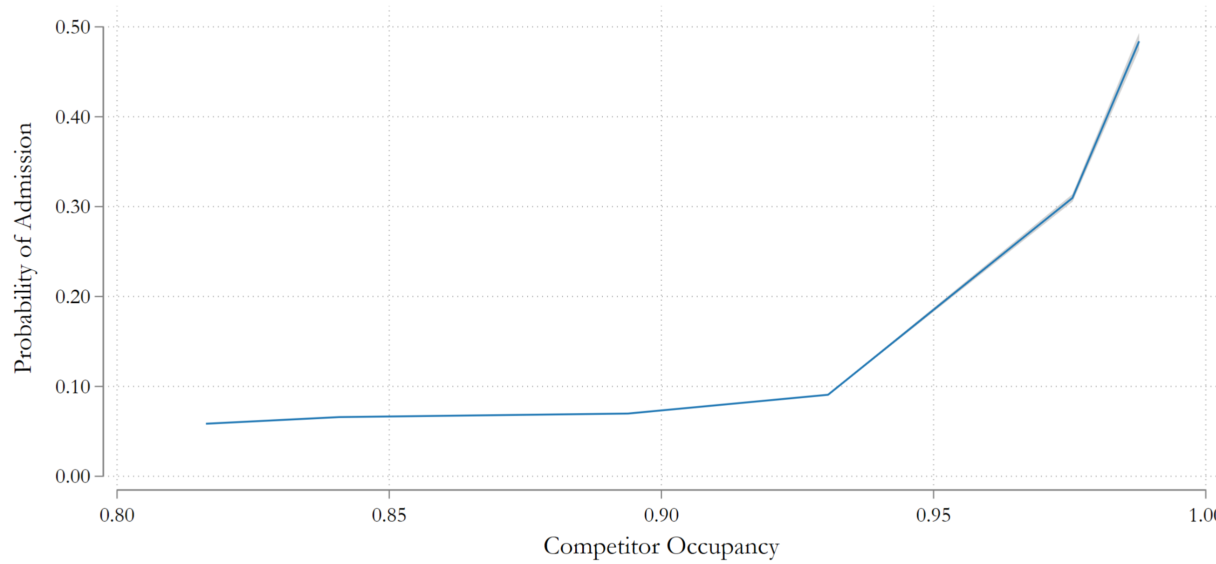


Figure IA.15. Within-Patient Billing Codes

This figure examines the potential fraudulent classification of patients by testing the likelihood that patients who were billed as receiving Ultra-High rehab care for every day of care under RUG-IV era that are subsequently coded as being low mobility patients with a categorization of Special Care High or Low Function once PDPM billing takes effect. Only those patients who were in a facility at the time of billing switch (October 1, 2019) are included. The graph shows a quadratic fit of the variable of interest and excess rehab. 95% confidence intervals are denoted by the grey shaded area. Special Care High is shown in the leftmost subgraph, Low Function, in the middle, and a category for either is presented in the far right subgraph.

Panel A. Low Mobility Classification among Ultra-High Rehab Patients

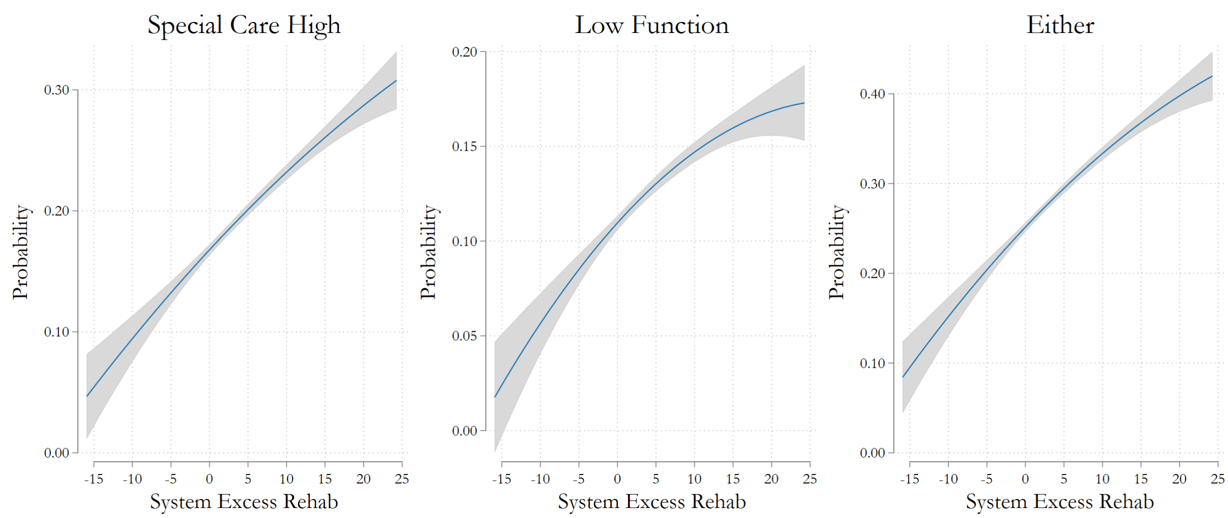


Table IA.1. County-Level Determinants of PDPM Billing

This table explores the relationship between county-level Coding Intensity and various measures of patient demographics, corruption, and measures of social capital. All demographic data comes from the American Community Survey 5-year estimates and is measured in the 2015 year. Public Corruption is the number of public corruption convictions per million residents within each judicial district as in Griffin et al. (2025). Religious Affiliation is the percentage of a county's population with a religious affiliation. Religious affiliation data comes from the Association of Religious Data Archives. Ahsley Madison Useage is the percent of a county's population with a paid Ashley Madison account. Data for Ashley Madison usage comes from Griffin et al. (2019). Coefficients are estimated using ordinary least squares and counties are weighted by the number of skilled nursing visits.

	(1) Code Intens.	(2) Code Intens.	(3) Code Intens.	(4) Code Intens.
Log Pop. Density	0.0686**	0.0644**	0.0746*	0.0695
	(2.47)	(2.09)	(1.78)	(1.60)
Log Median Income	0.108*	0.0997	0.0944	0.0869
	(1.87)	(1.25)	(1.45)	(1.06)
Poverty Rate	0.0816**	0.0814**	0.0683	0.0685
	(2.09)	(2.12)	(1.54)	(1.56)
Unemployment Rate	-0.0274	-0.0279	-0.0303	-0.0297
	(-1.27)	(-1.17)	(-1.22)	(-1.11)
Pct. College Educated	-0.0708***	-0.0745***	-0.0694**	-0.0724***
	(-2.81)	(-3.19)	(-2.39)	(-2.84)
Public Corruption		0.00453		0.00397
		(0.18)		(0.17)
Religious Affiliation		0.00705		0.0106
		(0.54)		(0.65)
Ashley Madison Useage		0.0119		0.00888
		(0.31)		(0.25)
Economic Connectedness			-0.0124	-0.0109
			(-0.41)	(-0.39)
Clustering			0.0510*	0.0495*
			(1.77)	(1.93)
Support Ratio			-0.0327*	-0.0336*
			(-1.90)	(-1.96)
Volunteer Rate			0.000954	0.00364
			(0.07)	(0.27)
Civic Organization			-0.0123	-0.0143
			(-1.18)	(-1.50)
State Fixed Effect	Yes	Yes	Yes	Yes
Observations	2,758	2,758	2,727	2,727
Adjusted R^2	0.636	0.637	0.646	0.646

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Table IA.2. Demographics, Corruption, Social Capital and PDPM Upcoding

This table explores the relationship between facility-level Coding Intensity and various measures of patient demographics, corruption, and measures of social capital. All demographic data comes from the American Community Survey 5-year estimates and is measured in the 2015 year. Public Corruption is the number of public corruption convictions per million residents within each judicial district as in Griffin et al. (2025). Religious Affiliation is the percentage of a county's population with a religious affiliation. Religious affiliation data comes from the Association of Religious Data Archives. Ahsley Madison Useage is the percent of a county's population with a paid Ashley Madison account. Data for Ashley Madison usage comes from Griffin et al. (2019). Coefficients are estimated using ordinary least squares and counties are weighted by the number of skilled nursing visits.

	(1) Code Intens.	(2) Code Intens.	(3) Code Intens.	(4) Code Intens.
Log Pop. Density	0.160***	0.149***	0.161***	0.143***
	(3.44)	(3.06)	(3.16)	(2.81)
Log Median Income	0.155	0.143	0.155	0.143
	(1.50)	(1.14)	(1.50)	(1.22)
Poverty Rate	0.121**	0.118**	0.115**	0.116**
	(2.22)	(2.15)	(2.06)	(2.06)
Unemployment Rate	-0.0506**	-0.0450	-0.0320	-0.0255
	(-2.10)	(-1.30)	(-1.29)	(-0.77)
Pct. College Educated	-0.123**	-0.124***	-0.111**	-0.112**
	(-2.54)	(-2.71)	(-2.32)	(-2.65)
Non-White	-0.00427***	-0.00426***	-0.00529***	-0.00539***
	(-4.59)	(-3.15)	(-3.36)	(-3.96)
Public Corruption		-0.0136		-0.00562
		(-0.49)		(-0.22)
Religious Affiliation		0.0414*		0.0516**
		(1.72)		(2.15)
Ashley Madison Useage		0.0132		0.00519
		(0.19)		(0.11)
Economic Connectedness			-0.0368	-0.0240
			(-0.99)	(-1.03)
Clustering			0.0657*	0.0554*
			(1.77)	(1.97)
Support Ratio			-0.0501	-0.0544*
			(-1.42)	(-1.77)
Volunteer Rate			0.00657	0.0139
			(0.22)	(0.65)
Civic Organization			-0.0128	-0.0213
			(-0.50)	(-1.38)
State Fixed Effect	Yes	Yes	Yes	Yes
Observations	14,412	14,412	14,366	14,366
Adjusted R^2	0.190	0.191	0.192	0.193

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Table IA.3. Facility Characteristics and PDPM Upcoding

This table explores the relationship between facility-level Coding Intensity and facility characteristics. Facility size is the log number of patients treated by each facility. System Size is the log number of patients treated by all facilities controlled by an SNF system. Enforcement is the number of DOJ Medicare enforcement activities per 100,000. Competition is the county-level HHI. County-average is the average coding intensity by other facilities within a county, excluding facilities operated by the same SNF system. Other facilities in the system is the average coding intensity of other facilities within an SNF system.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Facility Size	0.177***						0.0457***
	(4.95)						(3.02)
System Size		0.212***					0.0125
		(2.89)					(1.16)
Enforcement			0.00300				0.00663
			(0.11)				(0.50)
Competition				-0.0449**			0.00143
				(-2.06)			(0.10)
County-Average					0.126***		0.0848***
					(3.68)		(4.24)
Other Facilities in System						0.673***	0.660***
						(57.86)	(55.55)
State Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,430	7,430	7,430	7,430	7,430	7,430	7,430
Adjusted R^2	0.270	0.265	0.249	0.250	0.258	0.589	0.595

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Table IA.4. Excess Rehab and PDPM Billing–Individual Categories

In this table, we examine the relationship between SNF system-level excess rehab during the Resource Utilization Group IV (RUG-IV) era (January 1, 2016-September 30, 2019) with billing intensity during the Patient Driven Payment Model (PDPM) era (October 1, 2019-December 31, 2023). We estimate an OLS regression of the form:

$$y_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt}$$

where y_{ijt} is a measure of billing intensity for patient i at facility j during quarter t and is an indicator for a patient being classified as either Low Function, Special Care High, Depression, SLP High or Dietary Restriction. Low Function, Special Care High, and Depression are classifications that impact the Nursing component of PDPM case-mix and increase reimbursement. SLP High and Dietary Restriction classification increase PDPM reimbursements for the SLP component (as shown in Exhibits 2 and 1). Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Hospitalization length refers to the number of days for which a patient was hospitalized at an inpatient facility prior to beginning their SNF stay. Robust standard errors are clustered at the SNF system level.

	(1) Low Func.	(2) S.C.H.	(3) Depress	(4) SLP High	(5) Diet
Facility Rehab	0.00228***	0.00564***	0.00557***	0.00376***	0.00369***
	(12.80)	(20.30)	(18.39)	(23.14)	(22.19)
Patient Gender	Yes	Yes	Yes	Yes	Yes
Age Bucket	Yes	Yes	Yes	Yes	Yes
Patient Race	Yes	Yes	Yes	Yes	Yes
County x Quarter FE	Yes	Yes	Yes	Yes	Yes
Diagnosis x Hospitalization Stay	Yes	Yes	Yes	Yes	Yes
Observations	6,484,773	6,484,773	6,484,773	6,484,773	6,484,773
Adjusted R^2	0.074	0.114	0.110	0.198	0.065

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table IA.5. Instrumental Variables-Second Stage

In this table we examine the effect of excess rehab billing on PDPM billing intensity. We estimate a 2SLS regression of the form:

$$CodingIntensity_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt}$$

where y_{ijt} is a measure of billing intensity for patient i at facility j during quarter t and is an indicator for a patient being classified as either Low Function, Special Care High, Depression, SLP High or Dietary Restriction. We instrument for $ExcessRehab_j$ using $ExpectedFacilityRehab$ from Equation 5. Low Function, Special Care High, and Depression are classifications that impact the Nursing component of PDPM case-mix and increase reimbursement. SLP High and Dietary Restriction are classifications that increase PDPM reimbursements for the SLP component (as shown in Exhibits 1 and 2). Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Hospitalization length refers to the number of days for which a patient was hospitalized at an inpatient facility prior to beginning their SNF stay. Robust standard errors are clustered at the SNF system level.

	(1) Low Func.	(2) S.C.H.	(3) Depression	(4) SLP High	(5) Diet	(6) Code Intens.
Avg Rehab	0.00682***	0.0125***	0.0119***	0.00631***	0.00689***	0.0444***
	(8.32)	(11.74)	(6.93)	(15.77)	(13.55)	(12.71)
Patient Gender	Yes	Yes	Yes	Yes	Yes	Yes
Age Bucket	Yes	Yes	Yes	Yes	Yes	Yes
Patient Race	Yes	Yes	Yes	Yes	Yes	Yes
Diagnosis x Hospitalization Stay	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,53,6846	5,536,846	5,536,846	5,536,846	5,536,846	5,536,846

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table IA.6. Instrumental Variables-Second Stage (90% Cutoff)

In this table we examine the effect of excess rehab billing on PDPM billing intensity. We estimate a 2SLS regression of the form:

$$CodingIntensity_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt}$$

where y_{ijt} is a measure of billing intensity for patient i at facility j during quarter t and is an indicator for a patient being classified as either Low Function, Special Care High, Depression, SLP High or Dietary Restriction. We instrument for $ExcessRehab_j$ using $ExpectedFacilityRehab$ from Equation 5, except that the cutoff for $Constrained$ is 90%. Low Function, Special Care High, and Depression are classifications that impact the Nursing component of PDPM case-mix and increase reimbursement. SLP High and Dietary Restriction classification increase PDPM reimbursements for the SLP component (as shown in Exhibits 1 and 2). Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Hospitalization length refers to the number of days for which a patient was hospitalized at an inpatient facility prior to beginning their SNF stay. Robust standard errors are clustered at the SNF system level.

	(1) Low Func.	(2) S.C.H.	(3) Depression	(4) SLP High	(5) Diet	(6) Code Intens.
AvgRehab	0.00681***	0.0125***	0.0119***	0.00630***	0.00688***	0.0444***
	(8.32)	(11.84)	(6.96)	(15.80)	(13.56)	(12.75)
Patient Gender	Yes	Yes	Yes	Yes	Yes	Yes
Age Bucket	Yes	Yes	Yes	Yes	Yes	Yes
Patient Race	Yes	Yes	Yes	Yes	Yes	Yes
Diagnosis x Hospitalization Stay	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,536,846	5,536,846	5,536,846	5,536,846	5,536,846	5,536,846

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table IA.7. Instrumental Variables-Second Stage (90% Cutoff)

In this table we examine the effect of excess rehab billing on PDPM billing intensity. We estimate a 2SLS regression of the form:

$$CodingIntensity_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt}$$

where y_{ijt} is a measure of billing intensity for patient i at facility j during quarter t and is an indicator for a patient being classified as either Low Function, Special Care High, Depression, SLP High or Dietary Restriction. We instrument for $ExcessRehab_j$ using $ExpectedFacilityRehab$ from Equation 5, except that the cutoff for $Constrained$ is 90%. Low Function, Special Care High, and Depression are classifications that impact the Nursing component of PDPM case-mix and increase reimbursement. SLP High and Dietary Restriction classification increase PDPM reimbursements for the SLP component (as shown in Exhibits 1 and 2). Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Hospitalization length refers to the number of days for which a patient was hospitalized at an inpatient facility prior to beginning their SNF stay. Robust standard errors are double clustered at the SNF system and HSA level.

	(1) Low Func.	(2) S.C.H.	(3) Depression	(4) SLP High	(5) Diet	(6) Code Intens.
Avg Rehab	0.000542	0.0140***	0.0126***	0.00553***	0.00493**	0.0376***
	(0.27)	(3.16)	(3.36)	(2.91)	(2.33)	(3.99)
Patient Gender	Yes	Yes	Yes	Yes	Yes	Yes
Age Bucket	Yes	Yes	Yes	Yes	Yes	Yes
Patient Race	Yes	Yes	Yes	Yes	Yes	Yes
Diagnosis x Hospitalization Stay	Yes	Yes	Yes	Yes	Yes	Yes
HSA FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,536,843	5,536,843	5,536,843	5,536,843	5,536,843	5,536,843

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table IA.8. Instrumental Variables-Second Stage (95% Cutoff)

In this table we examine the effect of excess rehab billing on PDPM billing intensity. We estimate a 2SLS regression of the form:

$$CodingIntensity_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt}$$

where y_{ijt} is a measure of billing intensity for patient i at facility j during quarter t and is an indicator for a patient being classified as either Low Function, Special Care High, Depression, SLP High or Dietary Restriction. We instrument for $ExcessRehab_j$ using $ExpectedFacilityRehab$ from Equation 5, except that the cutoff for $Constrained$ is 95%. Low Function, Special Care High, and Depression are classifications that impact the Nursing component of PDPM case-mix and increase reimbursement. SLP High and Dietary Restriction are classifications that increase PDPM reimbursements for the SLP component (as shown in Exhibits 1 and 2). Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Hospitalization length refers to the number of days for which a patient was hospitalized at an inpatient facility prior to beginning their SNF stay. Robust standard errors are double clustered at the SNF system and HSA level.

	(1) Low Func.	(2) S.C.H.	(3) Depression	(4) SLP High	(5) Diet	(6) Code Intens.
Avg Rehab	0.00682***	0.0126***	0.0119***	0.00631***	0.00690***	0.0445***
	(8.31)	(11.75)	(6.92)	(15.72)	(13.58)	(12.71)
Patient Gender	Yes	Yes	Yes	Yes	Yes	Yes
Age Bucket	Yes	Yes	Yes	Yes	Yes	Yes
Patient Race	Yes	Yes	Yes	Yes	Yes	Yes
Diagnosis x Hospitalization Stay	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,536,846	5,536,846	5,536,846	5,536,846	5,536,846	5,536,846

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010

Table IA.9. Instrumental Variables-Second Stage (95% Cutoff)

In this table we examine the effect of excess rehab billing on PDPM billing intensity. We estimate a 2SLS regression of the form:

$$CodingIntensity_{ijt} = \alpha + \beta ExcessRehab_j + \theta X_{it} + \delta_t + \epsilon_{ijt}$$

where y_{ijt} is a measure of billing intensity for patient i at facility j during quarter t and is an indicator for a patient being classified as either Low Function, Special Care High, Depression, SLP High or Dietary Restriction. We instrument for $ExcessRehab_j$ using $ExpectedFacilityRehab$ from Equation 5, except that the cutoff for $Constrained$ is 95%. Low Function, Special Care High, and Depression are classifications that impact the Nursing component of PDPM case-mix and increase reimbursement. SLP High and Dietary Restriction are classifications that increase PDPM reimbursements for the SLP component (as shown in Exhibits 1 and 2). Patient controls include gender, age (in five-year increments), and race. Fixed effects are as indicated at the bottom of each column. Patient diagnosis comes from the referring hospital and is at the CCSR-level. Hospitalization length refers to the number of days for which a patient was hospitalized at an inpatient facility prior to beginning their SNF stay. Robust standard errors are double clustered at the SNF system and HSA level.

	(1) Low Func.	(2) S.C.H.	(3) Depression	(4) SLP High	(5) Diet	(6) Code Intens.
Avg Rehab	0.00332	0.0130**	0.0133**	0.00506**	0.00570*	0.0405***
	(1.14)	(2.30)	(2.50)	(2.24)	(1.85)	(2.92)
Patient Gender	Yes	Yes	Yes	Yes	Yes	Yes
Age Bucket	Yes	Yes	Yes	Yes	Yes	Yes
Patient Race	Yes	Yes	Yes	Yes	Yes	Yes
Diagnosis x Hospitalization Stay	Yes	Yes	Yes	Yes	Yes	Yes
HSA FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,536,843	5,536,843	5,536,843	5,536,843	5,536,843	5,536,843

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.010