NBER WORKING PAPER SERIES

THE EFFECT OF PERFORMANCE PAY INCENTIVES ON MARKET FRICTIONS:
EVIDENCE FROM MEDICARE

Atul Gupta
Guy David
Lucy (Kunhee) Kim

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

The Effect of Performance Pay Incentives on Market Frictions: Evidence from Medicare
Atul Gupta, Guy David, and Lucy (Kunhee) Kim
NBER Working Paper No. 30615
November 2022
JEL No. I11,I13,I18

ABSTRACT

Medicare has increased the use of performance pay incentives for hospitals, with the goal of increasing care coordination across providers, reducing market frictions, and ultimately to improve quality of care. This paper provides new empirical evidence by using novel operations and claims data from a large, independent home health care firm with the Hospital Readmissions Reduction Program (HRRP) penalty on hospitals providing identifying variation. We find that the penalty incentive to reduce re-hospitalizations passed through from hospitals to the firm at least for some types of patients, since it provided more care inputs for heart disease patients discharged from hospitals at greater penalty risk and that contributed more patients to the firm. This evidence suggests that HRRP helped increase coordination between hospitals and home health firms without formal integration. Greater home health effort does not appear to have led to lower patient readmissions.

Atul Gupta
Wharton Health Care Management
3641 Locust Walk, CPC 306
Philadelphia, PA 19104
atulgup@wharton.upenn.edu

Guy David
The Wharton School
University of Pennsylvania
305 Colonial Penn Center
3641 Locust Walk
Philadelphia, PA 19104-6218
and NBER
gdavid2@wharton.upenn.edu

Lucy (Kunhee) Kim
Google
kim.kunhee@gmail.com

I.   Introduction

US federal policy has emphasized the use of performance pay contracts to transmit quality and efficiency incentives to health care providers. However, there is considerable debate over the design and effectiveness of performance pay programs (Rosenthal et al., 2005; Christianson et al., 2008; Mullen et al., 2010; Jha et al., 2012). A key source of contention has been the use of patient outcomes as performance metrics and their association with unintended adverse effects due to gaming responses (Chen and Grabowski, 2017; Ibrahim et al., 2018; Wadhera et al., 2018). However, making pay contingent on outcomes can also be more efficient than using process (input) metrics if providers respond by pursuing more effective processes than would have been mandated otherwise. Policymakers need to factor in likely provider responses so they can optimally design performance pay contracts, however empirical evidence from health care remains limited. This paper attempts to fill this gap by documenting a novel response mechanism to outcomes-based penalties: overcoming market exchange frictions and shifting production to lower cost downstream providers. Our setting is the Hospital Readmissions Reduction Program (HRRP) — one of the largest performance pay programs in US healthcare.[1]

HRRP authorized the federal government to penalize hospitals if their re-hospitalization rate for Medicare patients admitted with select targeted conditions exceeds a specified benchmark. The Centers for Medicaid and Medicare Services (CMS) has imposed penalties of about $500 million on hospitals each year under HRRP. Prior studies vary in their assessments of the magnitude of its effect on the mean readmission rate, ranging from nearly zero to a one percentage point decrease, but the consensus is that it has decreased readmission rates for targeted patients (Carey and Lin, 2016; Desai et al., 2016; Mellor et al., 2016; Chen and Grabowski, 2017; Gupta, 2021; MedPAC 2018; Ziedan, 2019). While evidence on gaming responses such as up-coding (Ibrahim et al., 2018; Ody et al., 2019), patient selection and greater use of observation status (Gupta, 2021) has been forthcoming, evidence on quality improving mechanisms has remained speculative.

Hospitals could respond to the readmission penalty by sending more of their patients into post-acute care (PAC) as well as leverage their relationships with PAC partners to obtain more intense care for their patients – thus shifting some of the incremental care to a lower cost setting. Recent studies have shown the potential for PAC to improve hospital readmission rates (Rahman et al. 2017; David and Kim, 2018).

---

[1] The average penalty rate has consistently trended around 0.5% of total inpatient revenue from Traditional Medicare, or about 5% of revenue from the targeted conditions. Unlike other performance pay programs, there is no provision for bonus payments by doing better than the benchmark.

However, using Medicare claims data we find no economically or statistically significant changes in the proportion of hospital patients sent to any type of PAC—consistent with findings of recent studies (Gupta, 2021; Kim et al., 2019; Ziedan, 2019). Hence, we focus our attention on the intensive margin.

Among different types of PAC, we focus on home health care since in the HRRP context it is the only commonly used low-cost PAC. Forty percent of the patients targeted by HRRP and sent to PAC receive home health care. It costs only about 20% the price of a stay at a skilled nursing facility or at a hospital. Second, it is an important service in its own right—Medicare spending on home health care doubled between 2001 and 2014 (MedPAC, 2016). With the growth of payment bundles and the emphasis on cost containment, it is expected to continue to grow rapidly (Sood et al., 2011).

Freestanding Home Health Agencies (HHAs) compete for hospital referrals and adjust their protocols and intensity of care to preserve or enhance their business with referring hospitals. With the introduction of the readmission penalty, hospitals may have identified greater coordination with and effort by their HHA partners as a mechanism to decrease readmissions for their patients. Hence, market competition for hospital referrals offers a pathway for the hospital penalty incentive to pass-through to HHAs. To empirically test this hypothesis, we construct a measure of the salience of a hospital's penalty for an HHA and use this as the primary source of cross-hospital variation in incentive for an HHA to choose resource intensity for patients. We assume penalty salience increases in proportion with two factors—the share of the HHA's patient volume contributed by the hospital and the penalty rate determined under HRRP. To mitigate the potential for bias due to unobserved hospital quality, we use patients discharged with a condition not targeted by HRRP as a within-hospital comparison group. We combine these cross and within-hospital sources of variation in a difference-in-difference (D-D) research design where the first difference is between patients from targeted and non-targeted conditions and the second difference is between patients from hospitals at different levels of penalty salience for the HHA.

We implement the above approach using novel operational and claims data on over 12,000 patients from a large Home Health firm with 70 offices in 16 states. For each patient treated by the firm between 2012 and 2015, we observe rich episode level and visit level information. This dataset provides the usual details available in administrative claims data, but also includes rarely available operational details (such as time spent with the patient, and caregiver specialty) and an accounting-based cost estimate for each patient episode used by the firm in its internal business analytics. This allows us to not only quantify patient specific cost of care ('effort'), but also distinguish between alternative dimensions of effort, for example, increasing hours of care versus deploying more specialized caregivers who bill higher hourly rates.

Despite the advantages of this novel data, a key limitation is that we do not observe patients prior to the implementation of the HRRP penalty. This implies we cannot exploit the pre/post difference in outcomes in a traditional D-D framework as a source of identifying variation. Accordingly, the equivalent of the common trends assumption in our D-D design is that in absence of the penalty, differences in outcomes between patients in targeted and non-targeted conditions would be identical across patients from hospitals at different levels of penalty salience. We assess the credibility of this assumption and provide supporting evidence using multiple falsification tests.

Our results confirm that the firm incurred greater episode costs for heart attack and heart failure patients relative to non-targeted patients of similar risk at hospitals with higher levels of penalty salience. The implied increase in cost is economically meaningful—a heart failure patient from a hospital with 1 standard deviation greater penalty salience received care that cost about $60 more per episode (4.7% of the mean). While modest in absolute terms, we note that the firm spent only about $150 more at baseline for the sickest quartile of patients relative to the least sick quartile. Relative to this benchmark, it incurred 40% greater cost for targeted patients from a hospital with 1 s. d. greater penalty salience. However, we find no changes for pneumonia patients, implying heterogeneity in the pass-through of the penalty incentive within the targeted conditions.

We identify three channels through which the firm increased resource use. First, there is consistent evidence of greater effort in the first week of the episode. We interpret this as frontloading of effort, as recommended in the clinical literature (Rogers and Madigan, 2007; O'Connor et al., 2014; Murtaugh et al., 2017). Second, we show that the increase in cost is driven primarily by using more experienced caregivers and spending more time with targeted patients. This dimension of effort is typically unobserved in conventional claims data and may be an important reason why previous studies have not detected the role of PACs in responding to HRRP. Third, in the case of heart failure, the firm focused greater resource use on sicker patients. This is consistent with the HHA optimally directing greater resources to patients posing greater penalty risk to hospitals.

We present evidence from three different falsification tests to support the validity of our identification assumption. First, we show that the effects on episode cost disappear if we eliminate variation in penalty salience across hospitals, and instead assign the mean penalty rate across local area hospitals. Hence, the results are not driven by spurious responses to market-level penalty. Second, we use Medicare Advantage (MA) patients to implement a placebo test. MA patients are excluded from HRRP penalty calculations. Absent spillovers, hospitals and PACs should not invest additional resources for these patients due to HRRP. Reassuringly, we do not find differentially greater episode cost or lower readmissions for MA

patients. In fact, the coefficients typically suggest a reduction in resource intensity for MA patients, consistent with a multi-tasking effect. Third, we perform a randomization test, estimating the episode cost model 1,000 times. In each iteration we randomly assign a subset of patients (from targeted or non-targeted conditions) as the treated group and use the remaining patients as the control group. Thus, we obtain an empirical CDF of estimated effects on episode cost to compare against the coefficient obtained in the main results. We find that the true coefficient is at the 94th percentile of this distribution and can therefore reject the null hypothesis at conventional levels of significance.

This paper makes two contributions. First, we hypothesize and empirically test a novel mechanism through which providers responded to HRRP—one of the largest performance-pay programs in US health care. This program has received considerable attention, but prior studies have typically examined outcomes (readmissions and mortality) without offering insights into causal mechanisms (other than selection and coding). Overall, our results suggest that HRRP may have succeeded in its goal to improve coordination between hospitals and community-based providers by aligning incentives along the care continuum. We demonstrate that performance penalties imposed on hospitals passed through to downstream firms for heart disease patients, reallocating tasks across firm boundaries. This adds to the recent evidence on vertical integration in health care (Lehrman and Shore, 1998; David et al., 2013; Konetzka et al., 2018) and indicates that performance pay incentives may ease market frictions without changes in firm ownership. Second, these results inform policymakers on performance pay design in health care. It confirms that using outcomes as performance metrics motivates providers to deploy low-cost solutions, as suggested by economic theory (Shleifer, 1985; Lazear, 1986). It is implausible that policymakers could have identified and stipulated these specific aspects of coordination as input metrics.

The remainder of the paper is organized as follows. Section II provides the necessary institutional background. Section III describes our data sources and key variables. Section IV discusses our conceptual framework and presents the empirical strategy. We present our main results, including robustness and falsification checks in Section V, and conclude in Section VI.

II.   Background

A.   Performance pay in health care

Through the Affordable Care Act (ACA), the federal government committed to extensive use of performance pay in compensating hospitals for services to Medicare patients. With the passage of the Medicare Access and CHIP Reauthorization Act (MACRA) in 2015, performance pay incentives were extended to physician payments as well. Hence, performance pay is now a cornerstone of federal

payment policies. However, there is still considerable debate over whether provider performance should be measured through process metrics or patient outcomes.

This debate is not limited to health care and has a long tradition in the large theoretical literature on the use of performance pay incentives by firms (Prendergast, 1999). Economic theory predicts that linking pay to outcomes is more efficient when they are more easily measured than effort, and agents are better informed about the production function and their ability than the employer (Shliefer, 1985; Lazear, 1986). However, economic theory also predicts that the use of outcomes may lead to gaming (Baker, 1992) or distract agents from tasks that society values but does not reward explicitly, known as the multi-tasking effect (Holmstrom and Milgrom, 1991). The possibility of unintended adverse effects on patient health has made the use of outcomes controversial in health care (Ibrahim et al., 2018; Wadhera et al., 2018).

Nevertheless, both approaches are widely prevalent and often overlap. For example, CMS penalizes hospitals under HRRP if their heart attack patients have high readmission rates (an outcome measure). But it simultaneously rewards hospitals under the Hospital Value Based Purchasing (HVBP) if they provide heart attack patients with aspirin upon discharge (a process measure), regardless of readmission rates. Private insurers have taken the cue from Medicare and introduced performance pay programs of their own, using a mix of process and outcome measures (Chernew et al., 2011).

B.   The Hospital Readmissions Reduction Program (HRRP)

HRRP authorized CMS to penalize hospitals whose 30-day re-hospitalization rates for select Traditional Medicare (TM) patients are greater than the national average. Hence the program explicitly discourages readmissions, treating them as adverse outcomes. CMS enjoys tremendous discretion in administering the exact rules of the program—the targeted conditions, the penalty formula, and timing. Hospitals were officially notified of these details in August 2011. Starting in fall of 2011, hospitals could predict the consequences of HRRP on their finances, and accordingly take actions to decrease their readmission rates. In the first three years program years, hospitals were penalized for re-hospitalization following an index admission for heart attack, heart failure or pneumonia.[2] All unplanned readmissions, including those occurring at a different hospital, count against the hospital where the index treatment was performed.

---

[2] An index hospital stay is defined as an inpatient stay that begins at least 30 days after a previous hospital stay for the same condition ended. As an example, if a patient is hospitalized for pneumonia and then subsequently for an AMI (heart attack) within 15 days of discharging, the AMI will be counted as a readmission under pneumonia, but also as an index case under AMI. If she is subsequently readmitted for another AMI 15 days later, the second AMI will be considered only as a readmission.

The key source of variation in penalty applied in any year is the extent to which a hospital's risk-adjusted readmission rate deviates from the national average, all computed over a three-year baseline period. While a hospital does not receive a bonus payment if it does better (i.e., lower) than the threshold, it is penalized if its baseline rate exceeds the threshold. This rate is normalized against the national average and adjusted by a multiplier based on the hospital's total Medicare revenue, which is typically close to one.[3] Penalty values are computed separately for each targeted condition and a weighted sum is calculated at the hospital level. This ensures that hospitals have an incentive to improve on each condition, since excelling in one does not compensate for poor performance on other conditions. This aspect of the program motivates our empirical approach of using the variation in condition-specific penalty.

The condition-specific penalty dollar amount is directly proportional to the revenue billed by the hospital for the targeted condition ("base"), i.e., all else equal, hospitals serving more heart attack patients will incur a greater penalty for readmissions following heart attacks. To isolate variation due to baseline performance in readmission, we normalize the penalty amount by this base revenue, and refer to the ratio as the penalty per dollar of base revenue, or penalty rate in short. Figure 1 illustrates variation in this penalty rate across hospitals for pneumonia in the first year of the program. The dots represent penalty rates for individual hospitals, while the solid line plots a hypothetical linear increase. The plot confirms that baseline variation in readmissions performance is the major determinant of HRRP penalty variation.[4] The patterns are similar across the three targeted conditions. We use this as our preferred measure of penalty variation across hospitals in the empirical analysis.

In any given year, hospitals are penalized based on baseline performance, but the penalty is applied two years later. For example, the penalty was first applied to 2012-13 revenue based on baseline performance over July 2008–June 2011. In every successive year, hospital penalty status is updated based on a correspondingly updated three-year performance window (e.g., July 2009–June 2012 is the performance window for the next program year). This program design has an important implication for

---

[3] The multiplier term is the ratio of total Medicare hospital revenue in the penalty year to the mean annual Medicare revenue in the baseline period. We assume that the hospital can accurately predict its Medicare revenue at the start of the penalty year and therefore compute this penalty rate. Since Medicare revenue does not change substantially from year to year, this is a plausible assumption.

[4] The schematic also provides intuition that each readmission increases the penalty by a constant amount. The formula is designed such that each readmission beyond the benchmark invites a penalty equal to 5-6 times the mean reimbursement paid to the hospital for the index case (MedPAC, 2018). Table 1 Panel A shows that the weighted mean total reimbursement across targeted conditions is about $11,000. Since CMS pays hospitals about $8,000 per index case on average (70% of 11,000), this implies that each readmission incurs about $40−50,000 in penalty, substantially exceeding the revenue the hospital would receive from the index and readmission stays together.

our empirical analysis. Penalties in the first year were based on hospital performance during a period that ended prior to when the formula and other details were revealed. Hence, hospitals could not influence their first-year penalty status by changing patient mix, clinical condition mix, care protocols or coordination with other providers after August 2011. Penalty status in future years, however, could itself be influenced by hospital responses to HRRP, i.e., a reverse causality concern. Hence, in our empirical work we hold hospital penalty status constant as of the first year of the program. In fact, we go a step further and compute the applicable penalty rates based on hospitals readmissions over July 2007–June 2010 instead. This is much removed from the outcomes we study, which is post-acute care use during 2012–15.

C. Home Health Agencies (HHAs)

PAC providers care for patients and assist them with daily activities after receiving treatment in an acute care setting, such as inpatient care. PAC occurs in a variety of settings depending on the intensity of care required – in skilled nursing facilities, long-term acute care, and at home through home health or hospice care. The importance and role of PAC has grown over time as the share of hospital patients discharged to PAC has increased.[5] In fact, for Medicare patients discharged with heart attack or heart failure, PAC has been the fastest growing category of care (Chandra et al., 2013).

Home health care involves several general care and recovery tasks which could plausibly occur in the hospital. However, these tasks are much costlier to produce in the hospital setting, primarily due to the capital cost of hospital infrastructure and the alternative costs associated with lower throughput (Candrilli and Mauskopf, 2009; Metlife, 2011). HHAs also enable regular monitoring of patient health, offering the potential to prevent medical emergencies, at least in theory. Home health care agencies and industry associations emphasize that a key component of their value proposition for Medicare is their ability to avoid costly hospital readmissions (AHHQI, 2012). Since HRRP emphasized readmission reductions — an outcome that HHAs also target — it is a very salient program for home health agencies.

Freestanding home health care firms compete for hospital patient referrals (Study, 2014). While patients make the final choice, their decision is influenced by hospital physicians and discharge planners, who inform them about the different options, their strengths, and weaknesses. These decisions are subject to informal negotiations over the quality, intensity of care, and responsiveness that a home health firm can commit to the hospital and patient. Competitive pressure may induce HHAs to offer

---

[5] MedPAC. June, 2017. Data Book: Health Care Spending and the Medicare Program, Section 8: Post-acute care. http://www.medpac.gov/docs/default-source/data-book/jun17_databooksec8_sec.pdf.

additional services or intensity of treatment to win patient referrals. In the aftermath of the readmission penalty, this could include additional effort to help the hospital reduce readmissions (Worth, 2014).

Home health care is relatively inexpensive. Table 1 Panel B shows that Medicare paid about $2,600 for an episode following a hospital stay for one of the HRRP conditions. In contrast, the average reimbursement for a nursing home stay was about $10,000—comparable to the cost of the hospital stay. HHAs determine the care resources necessary for patients when the episode begins. Home health care involves home visits and care by both skilled (nurses, physical therapists) and semi-skilled employees (assistants, aides). In case of traditional Medicare patients, the HHA receives a prospective payment for an episode that can be up to 60 days long. Like hospital prospective payments, the amount varies based on the provider's assessment of intensity of care required for the patient and local market costs.[6] Hence, the HHA is not reimbursed on the margin for additional care within the same episode for the typical patient.

III.   Data

A.   Data sources

a)   Traditional Medicare claims data

We use a 100% sample of Medicare inpatient claims data to compute hospital penalty under HRRP per dollar of base revenue, as described in Section II.B above. We compute hypothetical condition-specific penalties applicable based on the performance period of July 2007–June 2010, following the formula used by CMS.[7] The mean unconditional penalty per dollar of revenue for hospitals across target conditions is about 7%. This implies that hospitals would expect to be penalized about 7% of the annual revenue they billed Medicare for the targeted conditions over the baseline period of 2007–10.

We also use a 20% random sample of Medicare claims to examine national patterns in use of PAC for hospital patients targeted by HRRP. This data spans July 2008 through June 2015 and therefore allows observing patterns before and after introduction of the penalty. Table 1 describes

---

[6] HHAs assess the patient's condition and needs at the start of the episode using the Outcome and Information Assessment Set (OASIS) instrument. Based on the patient's clinical severity, expected medical supplies, and number of therapy visits, patients are assigned to one of 153 Home Health Resource Groups (HHRG). CMS uses this code assignment and geographic adjustment factors to adjust the home health base payment rate, which was $2,969 in 2016. More details available at https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HomeHealthPPS/index.html.

[7] More details on the penalty rules and formula are available at https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html. We deployed the SAS code released by the CMS contractor (based at Yale school of public health).

this sample. There were about 1.3 million index hospital stays for the targeted conditions, with heart failure contributing the largest share (~45%). About 40% of these stays were followed by a PAC episode within 7 days of discharge from the hospital, with skilled nursing (18%) and home health (16%) the two largest types of care. Figure 2 presents time series patterns on the probability of being sent to some type of PAC within 7 days of discharge from a hospital stay for one of the targeted conditions (Panel a), and the mean number of PAC providers used by hospitals for each of the categories. The data suggests that the patterns remained relatively stable over this period, with no noticeable shifts immediately after 2011. We formally test for changes in PAC use in Section V.E. (a).

b)   Proprietary data from large HHA

The primary data source is a novel patient level operational and claims data from a large for-profit, independent home health care firm that operates 70 offices in 16 states.[8] The firm allows each office to operate rather autonomously, with control over human resources management, care protocols and scheduling. Each office is also held accountable as an independent profit center. Offices logically serve as the unit of our analyses. This feature of autonomous operations within the firm also helps partially alleviate concerns about the generalizability of our results to other HHAs.

We have data on all Traditional Medicare and Medicare Advantage patients treated by the firm over calendar years 2012–15. Patients in the sample were discharged following an inpatient hospitalization at one of 297 hospitals. The data contains detailed information for each care visit—length of the visit, date and time, specific provider who performed the care (and her service type), and estimated cost of the visit, including allocated personnel time. This information uniquely allows us to measure treatment intensity and patterns, such as time spent with the patient and personnel cost. These details remain unobserved in typical claims data. The files also contain rich patient health status assessment at the start of the episode, which allow us to control granularly for differences in patient status when they were discharged from the hospital.

---

[8] We exclude data from 30 offices since they were not operational during the pre-HRRP period, which we used to compute baseline patient flows. The 16 states include Arizona, Colorado, Connecticut, Delaware, Florida, Massachusetts, Maryland, North Carolina, New Jersey, New Mexico, Oklahoma, Pennsylvania, Rhode Island, Texas, Virginia, and Vermont. These states account for about 40% of index hospital stays in the targeted conditions, and a slightly greater proportion of related home health care episodes. Patient demographics and associated reimbursements are similar as in the national sample.

We define years based on July–June periods, to be consistent with the annual measurement cycle used by CMS for HRRP. The analysis sample is limited to patients who received home health care during July 2012–June 2015 and were discharged from inpatient care.[9]

Table 2 presents descriptive statistics on the analysis sample separately for patients with targeted and non-targeted conditions. The total sample contains approximately 9,000 Traditional Medicare episodes, equivalent to 130 episodes per office and 30 per hospital on average. Targeted conditions contribute about 2,200 patients of which heart failure is the largest (1,250). The control group is highly diffused across many conditions, each making a small contribution to the sample. The most frequent diagnosis in the control group is Osteoarthrosis with about 1,000 episodes or 14% of the control group.[10]

We use data from 2012 to establish baseline values for key variables as well as to produce summary statistics on distributions of outcomes like resource use and readmissions. We assign severity scores to patients as the sum of observable risk indicators coded by the firm at the start of the episode.[11] We use the severity scores to assign patients to top or bottom quartile of sickness. Table A. 1 Panel A presents descriptive statistics on demographics and risk factors for patients in the two groups. The sickest patients are more likely to be female and non-white and diverge dramatically on risk factors. For example, eighty percent of the sickest patients have a history of 2+ hospital stays, while less than 10% of the least sick patients do.[12]

Table A. 2 helps assess the representativeness of the proprietary sample. It presents data comparing inpatient stays and PAC use over 2009–11 on the universe of hospitals (Col. 1), the 16 states where the firm operates (Col. 2), and for the approximately 300 hospitals appearing in the proprietary data (Col. 3). Hospitals dealing with the HHA were more likely to use PAC and home

---

[9] In addition, we make some minor sample restrictions. We restrict to patients who had a single 60-day episode, which is a unit of TM payment also widely used under MA plans. The Home health episode must begin within 30 days of discharging from the hospital. In practice, the vast majority commence within a week of discharge. Limiting to June 2015 helps avoid overlap with a performance pay program for HHAs imposed by CMS in 9 states starting in January 2016. The selection of the 9 states was announced in the CY 2015 home health prospective payment system proposed rule on July 10, 2015 available at https://www.gpo.gov/fdsys/pkg/FR-2015-07-10/pdf/2015-16790.pdf.

[10] The top five conditions in the control group, ICD9 code, and # of patients respectively are 1) Osteoarthrosis and allied conditions 715 990, 2) Other disorders of urethra and urinary tract 599 717, 3) General symptoms 780 558, 4) Disorders of muscle ligament and fascia 728 456, and 5) Symptoms involving nervous and musculoskeletal systems 781 430. Collectively these account for about 46% of the control group sample.

[11] These include indicators for alcohol dependency, drug dependence, heavy smoking, obesity, disruptive behavior, memory loss, impaired decision making, intractable pain, urinary incontinence, and indicators for pre-home health care history of 2+ falls, 2+ hospital stays, recent decline in mental capacity, and take 5+ medications.

[12] We use data from calendar year 2012 and limit to the three targeted conditions to arrive at the sample in this table, about 860 patients. We then assign patients to quartiles using the severity score. Those with scores 0 or 1 are in the bottom quartile, while those with scores of 4 or more are in the top quartile.

health care at baseline but had similar 30-day unadjusted readmission rates. Table A. 2 Panel B presents descriptive statistics on demographics for home health care patients. The patient demographics at hospitals represented in the sample are very similar to the national averages, which is reassuring. Comparing these means to the corresponding demographics of patients in the proprietary sample in Table 2 panel C column 1, we find that patients at the firm have similar age and gender mix as in the population but are more likely to be non-white.

B. Outcome variables

a) Episode cost and other measures

Table 2 panel B presents descriptive statistics on resources deployed by the firm. Our preferred measure is an index of aggregate spending for each home health episode using detailed cost accounting data from the firm. The average episode costs the firm about $1,270 and costs are similar across targeted and non-targeted patients. It is composed of two segments. The first segment, called visit cost, reflects the specialty mix of caregivers and materials used and accounts for about 60% of total cost. For example, for two otherwise identical episodes if one received care entirely by nurses while the other received care by a mix of nurses and aides, then the former will have greater visit cost. The second segment is personnel cost, i.e., salary and overheads of the specific individual making the visit, and accounts for about 40% of total cost. This segment enables differentiation in cost between two caregivers of the same specialty. For example, a more experienced and qualified nurse will have greater personnel cost than a less experienced one. Finally, total cost also includes incidental expenses such as travel reimbursements, but these account for only 3% of total cost.

Total episode cost is our preferred measure since it incorporates multiple margins of effort (time spent with the patient, expertise and experience of the caregiver, and materials used) and translates them into dollars. It captures more information than using number of visits or total time spent with the patient, which we also observe.[13] Note that in the Medicare claims data we only observe total charges and number of visits.

b) Probability of readmission

---

[13] Regression analysis confirms that the cost components discussed above (visit and personnel costs) capture additional information about episode cost beyond what is conveyed by the number of visits and total time spent. After conditioning on the number of visits and total time spent, a 10% increase in visit cost is associated with a 6% (p<0.01) increase in total cost. The corresponding estimate for personnel cost is 7% (p<0.01).

We use readmission to a hospital within 30 days of discharge from the hospital as our key health outcome. In addition to being an important marker of quality, it is common among Medicare patients, and is the metric targeted by HRRP. The mean probability of readmission within 30 days for patients with targeted conditions in our sample is 16%.[14] Table A. 1 presents baseline readmission rates for the healthiest and sickest patient quartiles. The sickest patients are more than twice as likely to be re-hospitalized within 30 days, at 20 percentage points (pp) versus 9 pp for those in the bottom quartile.

IV. Empirical Strategy

A. Conceptual framework

This section discusses the home health agency's objective and why it would invest additional effort to reduce hospital readmissions in response to the penalty. Nearly 90% of agencies in the US are for-profit, hence it is natural to assume the average firm maximizes total profit (MedPAC, 2016). This is also true of the specific firm studied in this paper. According to MedPAC, home health firms make relatively high operating margins on Medicare patients – about 11% in 2014. This estimate accounts for fixed costs, implying the marginal profit on a Medicare patient is higher still. Hence, agencies are incentivized to serve as many Medicare patients as their capacity allows them to. Hospitals are the dominant source of patient referrals for agencies (HHCN, 2019). However, freestanding firms must compete for hospital referrals. Figure 2 shows that hospital patients are discharged to 3.5 different home health firms on average.

Hospitals have no contractual arrangement with home health firms, nor can the latter pay hospitals to obtain patient referrals. However, hospitals still have substantial leverage over home health firms since they direct Medicare patient volume to the firms. Hospital discharge planners and physicians are key influencers in helping patients choose home health agencies. This leads to a familiar situation in health care where hospital personnel serve as de facto agents for patients. Hence, agencies invest in building relationships with hospital administrators, physicians, and particularly discharge planners. They differentiate themselves on dimensions like responsiveness, specializing in certain disease types, and offering additional services, sometimes even outside of what Medicare reimburses (Study, 2014). Home health firms have to pitch their services to hospitals and tailor their services to suit the preferences of both hospitals and patients, rather than of patients alone. Since Medicare patients

---

[14] Following previous studies (David and Kim, 2018), we exclude patients who receive a home-health visit on the same day as they are re-hospitalized. This avoids situations where a provider visit uncovered an emergent condition requiring immediate hospital care.

are financially lucrative, firms have a cushion to tradeoff higher cost per patient to increase patient volume and total profits.

In theory, home health care can help patients avoid hospital readmissions, though the empirical evidence is mixed (Linertova et al., 2011; Murtaugh et al., 2017; David and Kim, 2018). A review of the trade press reveals that home health firms and their industry groups emphasize this potential benefit as a key point in their marketing communications (Lotven, 2012; Mullaney, 2015). Both this background and discussions with executives at our data partner suggest that HRRP was a salient policy change for home health firms. Driven by demand from hospitals worried about the readmissions penalty and firms that strategically offered additional services to differentiate themselves and strengthen their market share, we hypothesize that the penalty incentive passed through from hospitals to home health firms that served their patients.

In addition to the per-episode cost of care, the home health firms also incur some fixed costs of business development to source referrals from hospitals and fixed costs of operations to maintain service capacity. Due to the presence of these fixed costs, hospitals contributing greater patient revenue become more profitable for the firm. We hypothesize these relationships accordingly become more valuable for the office manager and their demands more salient. All else equal, the agency should be willing to invest more resources in patients from hospitals that contribute a greater share of patients.[15]

B.   Penalty salience

This section develops a measure to empirically operationalize pass through of the penalty incentive from the hospital to the home health firm. Imagine that heart failure patients from three hospitals, A, B and C, contributing 40%, 40% and 20% of the patient volume at a HHA office are referred simultaneously to that office. Further, assume the three hospitals are penalized at 0.5%, 1% and 0.5%, respectively, on heart failure under HRRP. We hypothesize that the readmission penalty is more salient for the HHA office for patients arriving from hospital A relative to those from C even though both are penalized at 0.5%, since A contributes a greater share of patients to the office. We also posit that the penalty is more salient for patients discharged from B relative to those from A since B has a greater penalty rate and hence is more likely to assert this issue in discussions with the HHA.

---

[15] There are alternative ways to motivate this relationship. For example, hospital leverage may increase not in the hospital's existing contribution to patient volume, instead in its unrealized potential for contributing to the HHA. For example, imagine a leading academic medical center (AMC) that refers many patients to HHA care, but only a very small fraction to a new HHA entrant in the market. The new entrant would plausibly be willing to offer some additional service features to patients from the AMC to steal some of this potential business. Our data limits us from accurately observing unrealized potential, and hence we cannot operationalize this approach.

The above discussion pertains to patients discharged from hospitals with one of the targeted conditions. Patients not subject to the penalty, i.e., those discharged with non-targeted conditions or those covered by other insurers, should not receive additional care from the HHA since effort is costly and not reimbursed on the margin.

We develop an empirical measure of condition (c) specific penalty salience, P_{hjc} for each HHA office j that varies across referring hospitals h. Following the discussion above, we construct this variable as the product of the HRRP penalty per dollar of revenue, $penalty_{hc}$ and the share of patients at HHA office j contributed by hospital h, $share_{jh0}$.

$$P_{hjc} = penalty_{hc} * share_{jh0} \qquad (1a)$$

$$share_{jh0} = \frac{n_{hj0}}{n_{j0}}, \qquad (1b)$$

where n_{hj0} is the total number of patients across all conditions referred by hospital h to office j during the period January–June 2012, and n_{j0} is the corresponding total number of patients across all conditions served by office j during that period. We use the total patient volume (including MA patients[16]) to be consistent with our conceptual framework that the importance of the hospital depends on its total contribution to the office.

We hold penalty salience constant over time even though both penalty rate and hospital shares change annually. However, all such updates are potentially contingent on the initial penalty and hence are endogenous to the penalty. The mean (unconditional) penalty salience across targeted patients is 0.02 with a standard deviation of 0.04. Consistent with the discussion above, we set $penalty_{hc} = 0$ for all patients with non-targeted conditions.

C. Research design

We develop a difference-in-difference research design that exploits the natural experiment created by HRRP's readmission penalty and is feasible given our data. The first difference is taken across hospitals at different levels of penalty salience. This is a pre-determined, policy driven variation in the penalty perceived by the same office for patients discharged from different hospitals. A key limitation is that we do not have data from the period prior to the HRRP penalty. Hence, we *cannot* implement a traditional research design with pre/post variation providing the second difference. We

---

[16] MA patients contribute about a quarter of total patients for the firm, but they are distributed unevenly across offices. This is a result of variation in the penetration of MA across different markets. Over the period January–June 2012, the inter-quartile range of MA share across offices of the firm was about 18 percentage points. Ignoring MA patients would therefore introduce non-random measurement error in computing hospitals' patient share of each office.

overcome this limitation by using patients in non-targeted conditions as a within-hospital comparison group. Hence, the second difference is between patients discharged from the same hospital with targeted and non-targeted conditions. This approach therefore allows home health effort and readmissions to differ between patients from the same hospital belonging to the treated and comparison cohorts, but tests if these differences are correlated with penalty salience.

The key identifying assumption that enables interpreting the DD coefficient as causal is that in absence of the penalty, differences in outcomes between targeted and non-targeted patients would be similar across hospitals at different levels of penalty salience. This is equivalent to the common trends assumption in a conventional difference-in-difference design which would have assumed that treated patients from hospitals at different levels of penalty salience would have progressed on parallel trends in absence of the penalty.

We assess the validity of our identifying assumption through three different falsification exercises under which we would expect null effects. These are, respectively, shutting down variation in penalty salience across hospitals for an office and simply using a mean market-level penalty rate, replicating the analysis with MA patients who are not subject to HRRP, and a randomization test where we estimate the model 1,000 times assigning a random subset of patients as a placebo treatment group in each iteration. This exercise generates an empirical distribution of placebo effects on effort against which we can compare the estimate from the main model. Reassuringly, in all three cases, we find small and statistically insignificant effects relative to the coefficients from the main analysis.

A caveat with the proposed research design is that spillover effects to non-targeted patients will bias the estimated effects away from the true effect of the penalty. Spillover effects could be positive if the home health firm adopts new protocols uniformly across patients or hires better quality nurses for all patients. On the other hand, if the firm has capacity constraints, greater resources on targeted patients may come at the expense of non-targeted patients. Hence, it is difficult to predict the direction of the bias.

Prior studies on HRRP have found evidence of net positive spillovers for non-targeted patients (Chen and Grabowski, 2019; Batt et al., 2020; Gupta, 2021), i.e., patients in non-targeted conditions or those covered by other insurers also experienced a net reduction in readmissions. If this is also the case in our setting, then our estimates will be biased downward and understate the true effect. Spillover effects have been found to be larger for patients in conditions related to those targeted by the program. To mitigate the potential for spillover effects in our analysis, we exclude patients discharged from the hospital with a cardiovascular or respiratory condition— those with a principal ICD 9 diagnosis code within the 390xx-519xx range.

D.   Estimating equation

We conduct our analyses at the patient episode level. For each patient episode i receiving treatment at HHA office j and discharged from hospital h with condition c, we model outcome Y_i as below:

$$Y_i = \alpha_j + \alpha_h + \alpha_t + \alpha_{d_c} + \sum_{\substack{d_c=AMI, \\ HF,Pneum}} \beta_c \, P_{hjc} \, 1(d_c = 1) + \gamma_1 X_i + \epsilon_{ijht} \tag{2}$$

We include office (α_j) and hospital (α_h) fixed effects.[17] The former eliminate stable unobserved differences in personnel and managerial quality across offices. The latter eliminate time invariant differences in quality across hospitals and ensure we compare treated and control patients discharged from the same hospital. We include indicators for each targeted condition, α_{d_c} to allow for stable differences in outcomes relative to non-targeted conditions. P_{hjc} is the penalty salience for patients discharged with condition c from hospital h to HHA office j. It is set to zero for patients with non-targeted conditions and calculated separately for each of the three penalized conditions.[18]

The coefficients of interest in this model are β_c, estimates of the differential effect of being discharged with targeted condition c relative to non-targeted conditions from a hospital with penalty salience equal to one, relative to the same difference at a hospital with zero penalty salience, all within the same office. Table 2 Panel A shows that the variation in penalty salience is split nearly equally into the within-office and between-office components. Hence, there is substantial variation in penalty salience across hospitals served by the same office.

We control for observable differences across patients by including a vector of co-morbidity indicators X_i. Table 2 Panel C presents descriptive statistics on these variables. ϵ_{ijht} is an idiosyncratic error term capturing unobserved disturbances that may affect the outcome. We cluster standard errors by hospital and office to account for potential correlation in outcomes for patients discharged from the same hospital or served by the same office.

V.   Results

---

[17] We would like to thank an anonymous referee for the suggestion to include hospital fixed effects in our estimating equation.

[18] Note that if P_{hjc} is included by itself in the equation it will drop out since it is set to zero for all non-targeted patients.

A. Descriptive evidence

We first visually explore if unadjusted patterns of episode cost and patient readmissions suggest a differential response for patients from penalized hospitals. Figure 3 presents time series trends of mean episode cost (Panel a) and readmission rates (Panel b) using data from the proprietary data sample. We plot the mean values separately for patients discharged from hospitals with non-zero penalty salience ('treated') versus those from hospitals with zero penalty salience ('comparison'). The first group includes approximately 1,100 patients discharged with targeted conditions, while the latter group includes all patients with non-targeted conditions and about half the patients with a targeted condition discharged from hospitals for which we calculated a zero-penalty rate (approx. 7,900 patients). In each plot, the mean values are presented over the calendar years 2012–15, with the first and last data points representing six-month averages. For expositional clarity, we normalize the raw values for both groups by their respective means in 2012, but do not adjust the data otherwise. Note that we do not have data from what we would consider as the pre-HRRP period (prior to August 2011).

Figure 3a shows that episode costs increased in 2013 for patients from hospitals facing a non-zero penalty salience and declined steadily later. In contrast, episode costs for patients from zero penalty salience hospitals do not meaningfully increase in 2013 but do exhibit the same declining trend. The declining trend over 2014 and 2015 is common to both groups and may also reflect a structural shift in how the firm computes these cost estimates. At the end of the period, the cost for the treated group is similar to what it was in 2012, while for the comparison group, it is about 10% lower. Panel b presents mean values of thirty-day readmission rates for patients in the two groups. There is a clear decline for the treated group in 2014, which persists in 2015 (about 10% lower relative to 2012), while the trend is flat for the comparison group.

This evidence suggests differential movements in episode cost and patient outcomes for patients from hospitals likely to be penalized and contributing greater volume to the HHA. These patterns are unadjusted and may therefore partly reflect differences in patient or office composition over time. Hence these are only suggestive, and in the next section we present results obtained by estimating equation 2 discussed above.

B. Resource intensity

Table 3 columns 1–6 presents results on Home Health resource intensity, obtained by estimating equation 2. For each measure of resource intensity, we present two sets of results – columns 1, 3, and 5 pertain to effects for the entire episode, while columns 2, 4, and 6 present effects only after the first week of the episode. We use outcomes at one-week duration to test if the firm responded by increasing

effort disproportionately at the beginning, i.e., a frontloaded response. Each column presents two sets of condition specific coefficients. The average effect of being discharged with a targeted condition relative to a non-targeted condition regardless of penalty salience is presented in rows 1, 3, and 5. The coefficient of interest is that on the interaction of the condition dummy and penalty salience (rows 2, 4, and 6).

We begin by exploring effects on episode cost (columns 1 and 2). The coefficients suggest differentially greater episode cost for heart failure patients from hospitals with greater penalty salience. The implied effect for heart failure patients discharged from a hospital with 1 s. d. greater penalty salience is 4.7% greater episode spending ($e^{0.78} - 1 = 1.18 \times 0.04 = 0.047$). Since the mean episode cost for targeted patients was about \$1,280, the implied amount is \$60. Since this is not a conventional DD design, a brief interpretation of this coefficient is warranted. The estimate implies that, within an office, the difference in episode cost between heart failure and non-targeted condition patients from the same hospital is \$60 greater for hospitals with 1 s. d. greater penalty salience. The remaining DD estimates from our main model can be interpreted in a similar fashion.

When we limit our attention to the first week, we find statistically significant greater cost for heart attack patients, while the coefficient for heart failure patients is also positive but marginally significant. Accordingly, we focus more on the effect for heart attack patients. An increase of 1 s. d. in penalty salience is associated with a statistically significant 6.9% greater cost of care in the first week of the episode for heart attack patients ($e^1 - 1 = 1.71 \times 0.04 = 0.069$). This implies \$28 greater cost in the context of our sample (6.9% of approximately \$400).

When interpreting the magnitude of these implied effects, it is useful to consider the range of costs observed in the baseline period between the sickest and least sick patient groups (about \$150, see Table A. 1). If we regard this as an empirical benchmark of the extent by which the firm can increase episode resource intensity, then a 1 s. d. greater penalty salience is associated with a 40% (60/150) greater episode cost for heart failure patients, and the first week costs for heart attack patients are differentially greater by more than the range of cost observed in the baseline period (\$12). Taken together, the firm appears to use more resources in an economically meaningfully way for heart disease patients from hospitals with greater penalty salience.

Table A. 3 presents the results of estimating equation 2 on individual cost segments – visit and personnel costs (in logs). The results indicate that the differentially greater cost for heart failure patients is driven by both greater visit and personnel cost categories. A 1 s. d. greater penalty salience is associated with a 4% greater visit cost ($e^{0.7} - 1 = 1.03 \times 0.04 = 0.04$), or about \$30 (4% of 760) and a similar amount in additional personnel cost. Hence, there is a greater relative increase in personnel

cost, since this category accounts for only 40% of total episode cost, but about half of the increase. The estimated effects on first week costs for heart attack patients indicate a similar pattern – personnel costs account for 55% of the total effect. The increase in personnel costs implies that the firm is deploying more experienced caregivers for the targeted patients from hospitals with greater penalty salience.

Greater visit cost could be driven by two channels – spending more time with the patient or deploying more specialized workers (e.g., registered nurses or physical therapists rather than semi-skilled aides). We directly observe the time spent with patients and study the effect on this dimension. Table 3 columns 3-4 and 5-6 present results on the total time spent and the number of visits, respectively. The point estimates imply an increase in the amount of time spent with both heart attack and heart failure patients. For example, the coefficients imply the firm spends about 20 minutes more (3% of the mean) for heart failure patient episodes from a hospital with 1 s. d. greater penalty salience. The coefficients imply that the increase in time spent with the patient is comparable to the estimated increase in total episode cost (e.g., we find a 7% increase in both the time spent and the episode cost in the first week for heart attack patients). Hence, we interpret these results as suggesting that greater visit cost is largely driven by spending more time with the patient.

Note that standard claims data would allow us to observe only the number of visits. Even for this outcome, the 20% Medicare sample would likely produce downward biased estimates due to measurement error. First, sampling error would lead to noise in the hospital's share of HHA volume. Second, without the corresponding MA claims for each hospital, we cannot compute the hospital's true share of the HHA's Medicare patient volume accurately.

We note here that we do not find an increase in costs across all targeted conditions – the estimated effects for pneumonia patients are typically statistically insignificant. This may be due to sampling error, but the coefficients are consistently negative across the different components of cost, suggesting a decline in resource intensity.

C.   Readmissions

Table 3 columns 7 and 8 present the estimated effects on the probability of 30-day and 7-day readmissions, respectively. The estimates indicate differentially lower readmission probability in the first week for heart failure patients and – surprisingly – no decline over the 30-day period. The estimated effects imply that heart failure patients from a hospital with a 1 s. d. greater penalty salience experience a 1 percentage point lower probability of readmission ($0.26*0.04 = 0.01$) in the first week. This effect is about 20% of the mean readmission rate in the first week and is about half the difference in readmissions between the sickest and least sick patient groups. Alternatively, a heart failure patient

from a hospital with mean penalty salience experienced a 0.5 percentage point lower probability of readmission (0.26*0.02 = 0.0052). This is well within the range of estimates from previous studies of the decrease in readmissions due to HRRP (Desai et al., 2016; Gupta, 2021). However, these gains don't persist to the 30-day point.

The estimated effects on patient readmissions are puzzling given that the firm invests greater resources for targeted patients without delivering consistent improvement in readmissions. It may signal lack of effectiveness of home health care in addressing hospital readmissions, or heterogeneity in patient outcomes that is masked by small average effects. Accordingly, in the next section we test for heterogeneous effects.

D. Heterogeneity

We hypothesize that since incremental costs borne for these patients come at the cost of the firm's episode level surplus, it will selectively target patients for additional care. Since the objective behind the incremental effort is to help reduce penalty risk for hospitals, the firm should prioritize more resources for patients at greater readmission risk. We test this hypothesis by exploiting variation in observed risk across patients as an additional dimension that may determine the firm's response.

The patient severity score (described in Section III.A) is an excellent predictor of readmission risk – the sickest patient quartile have more than twice the 30-day probability of readmission as do patients in the bottom quartile (see Table A. 1). Accordingly, we construct an indicator that takes value one if the patient has a severity score greater than the sample median (i.e. $d_s = 1[s_i \geq s_{50}]$). We construct the indicator by comparing targeted and non-targeted patients to their group-specific median values. We then estimate the model in equation 3 below.

$$Y_i = \alpha_j + \alpha_h + \alpha_t + \alpha_{d_c} + \sum_{\substack{d_c=AMI, \\ HF,Pneum}} \beta_{1c} P_{hjc} 1(d_c = 1) + \sum_{\substack{d_c=AMI, \\ HF,Pneum}} \beta_{2c} 1(d_c = 1) 1(d_s = 1)$$

$$+ \sum_{\substack{d_c=AMI, \\ HF,Pneum}} \delta_c P_{hjc} 1(d_c = 1) 1(d_s = 1) + \gamma_1 X_i + \epsilon_{ijh} \qquad (3)$$

This model has three sets of coefficients of interest. β_{1c} is equivalent to the main interaction coefficient from equation 2 and estimates the differential effect for patients discharged with targeted condition c from hospitals with penalty salience $P_{hjc} = 1$. However, this coefficient now estimates the main effect for patients in the bottom half of sickness severity (henceforth, healthier patients). β_{2c}

estimates the average effect for sicker patients discharged with targeted condition c from hospitals with zero penalty salience. This coefficient soaks up any differential responses for sicker patients even from non-penalized hospitals. Finally, δ_c is the coefficient of interest in this model and it isolates the differential response for sicker patients from high penalty salience hospitals, in addition to the effect β_{1c} discussed above.[19] As in the base specification, we continue to control for observable differences across patients (X_i, which now includes the indicator d_s). When interpreting the coefficients, note that the reference group is now *healthier* non-targeted patients.

We estimate equation 3 with our main outcomes – total cost and probability of readmission and present the results in Table 4. Since the main results suggest greater effort and lower readmissions in the first week, we also present evidence on these outcomes after one week of care. For brevity of presentation, we omit the condition specific coefficients α_{d_c}. The coefficients estimating the effect for sicker targeted patients from non-penalized hospitals (β_{2c}) were generally small and statistically insignificant across both outcomes – suggesting little differential effort for sicker *targeted* patients, if discharged from a hospital with zero penalty salience. These are also not reported. The table presents the main interaction coefficients (β_{1c}) and differential response for sicker patients (δ_c).

The results on total cost clarify that the firm provided greater resources to sicker heart failure patients, particularly in the first week of care. Column 2 indicates that while first week cost may have been lower for healthier heart failure patients, cost for sicker heart failure patients from a hospital with 1 s. d. greater penalty salience is about 5% greater ($e^{0.97-0.16} - 1 = 1.25 \times 0.04 = 0.05$). Table 4 columns 3 and 4 present corresponding results on patient readmissions in 30-days and 7-days, respectively. Column 4 shows that the decrease in first week readmissions for heart failure patients is largely due to patients sicker than the median.

The heterogeneity analysis results on episode cost and readmissions helps resolve some, but not all the puzzles raised by the average effects discussed in the previous section. For example, among heart failure patients sicker patients appear to have received differentially greater treatment intensity. They also experience fewer readmissions at the 7-day mark. However, these gains do not persist to 30 days. Taken together, the evidence suggests that greater home health care intensity is successful in delaying readmissions but did not meaningfully avoid hospital care.

E.   Threats to identification and robustness checks

---

[19] The term $\sum_{\substack{d_c=AMI, \\ HF, Pneum}} \beta_{3c} P_{hjc} \, 1(d_{s_i} = 1)$ i.e. the interaction of hospital penalty salience and patient being in the sicker half of the sample soaks up the differential response for sicker patients from hospitals of greater penalty salience, regardless of whether they were targeted or not. However, since we define $P_{hjc} = 0$ for patients from non-targeted conditions, this term is collinear with the main coefficient, δ_c and drops out.

a) Changes in hospital referral behavior due to HRRP

Since patients must receive home health care to enter the proprietary data sample, a natural concern is whether hospitals' response to HRRP includes changes in the patient mix referred to home health care. If penalized hospitals responded by differentially sending more patients to home health, and if the marginal patients were more costly to serve, then we would spuriously find higher spending for these patients. The opposite response is also possible, that is, the marginal patient discharged to home health may be healthier than before, in which case we would be biased against finding an increase in spend. It is therefore difficult to sign the bias and assess its importance without additional analysis.

To formally study changes in extensive margin use of post-acute care, we used the Medicare claims data to estimate linear probability differences-in-differences models presented in equation 4 below. The models were estimated independently for each targeted condition, and hence the subscript c is omitted. The dependent variable is an indicator for use of any PAC or specific sub-type within 7 days of discharge from the index hospital stay. The coefficient of interest is β which quantifies the mean change in use of PAC for a hospital with penalty rate, $penalty_h = 1$ over the period 2012–15 relative to the pre-HRRP period. Note that we hold hospital penalty rate constant as of 2012 to circumvent reverse causality concerns.

$$1(PAC)_i = \alpha_h + \alpha_t + \beta \, penalty_h \, 1(T_t = 1) + \gamma_1 X_i + \epsilon_{iht} \tag{4}$$

Table A. 4 presents corresponding estimates for each of the targeted conditions. Barring a couple of coefficients, none are statistically significant. More importantly, the implied magnitudes are very small. For example, a 1 s. d. increase in penalty rate is associated with a decrease of 0.001 (0.06 x -0.02) in the probability of a heart failure patient being sent to home health care. This is negligible relative to the mean usage rate of home health care in the sample (0.17). The other coefficients are of comparable magnitude. We also estimated a flexible, non-parametric specification where we used an indicator d_h for hospitals in the top tertile by $penalty_h$. This is presented in equation 5 below. This model also tests for differential pre-trends in the use of PAC across hospitals. For brevity, this exercise focuses on the two largest categories of PAC – home health and skilled nursing.

$$1(PAC)_i = \alpha_h + \alpha_t + \sum_{s \neq 2008} \beta_s \, d_h \, 1(t_t = s) + \epsilon_{iht} \tag{5}$$

Figure A.1 plots the coefficients for any PAC (Panel a), home health (Panel b), and skilled nursing (Panel c), respectively. There is no distinct pattern before or after the implementation of the penalty for any of the conditions or any type of PAC. Taken together, these results imply no change in the use of PAC for the average patient with one of the targeted conditions.

Another concern is that while there is no change in the fraction of patients sent to PAC, the risk profile of patients in PAC may have changed. Table A. 5 presents results from triple difference models testing for heterogeneous effects based on patient sickness severity.[20] Columns 1, 3, and 5 reproduce the results discussed above for comparison. Columns 2, 4, and 6 present the triple difference coefficient for each of the targeted conditions. These are also small and statistically insignificant. This helps alleviate the concern that hospitals may be sending sicker or healthier patients to home health care post-HRRP. These findings are consistent with those of previous studies on HRRP which have not found evidence of changes in PAC use by hospitals (Gupta, 2021; Kim et al., 2019; Ziedan, 2019).

b)   Falsification tests

(i)   Eliminating variation in penalty across hospitals

We interpret differentially greater episode cost for targeted patients as evidence of greater effort by the home health firm in response to HRRP. The identification assumption is that if the penalty were not implemented, there would be no systematic variation in the difference in outcomes between targeted and non-targeted patients from hospitals at different levels of penalty salience. We assess the plausibility of this assumption through a falsification test where we shut down variation in penalty salience across hospitals and assign to targeted patients the average penalty rate across all hospitals sending patients to the same office. Non-targeted patients continue to be assigned zero penalty salience. We then re-estimate our main models. Since penalty salience now only varies across markets, the key coefficient now captures whether offices in higher penalty markets spend more on targeted condition patients relative to those in lower penalty markets. If our results largely reflect the firm's response to local market sentiment regarding the penalty, then we should continue to find meaningful effects even with this aggregate penalty measure. Table 5 Panel A presents the corresponding results. Reassuringly, the estimates are small relative to the main coefficients and statistically insignificant across all three conditions. For example, the

---

[20] We used history of hospital use to identify sicker patients. Specifically, we generated an indicator for patients with a hospital stay in the 6 months prior to the index hospital stay. This strongly predicts readmission ($p<0.01$) and is not susceptible to up-coding at the hospital.

DD coefficient on episode cost for heart failure patients is 0.32 (about two-fifth the size of the main estimate) and is statistically insignificant. This exercise confirms that our results are not driven by a generic response to the introduction of the penalty.[21]

(ii)   Medicare Advantage patients only

CMS does not calculate or consider readmissions for MA patients to compute a hospital's penalty under HRRP. MA patients provide a suitable control group since they are of similar age as fee-for-service patients. Under our conceptual framework, the hospital has no incentive to insist on differentially greater resource use for these patients, and the HHA obviously has no reason to do so on its own. Since there are only 23 heart attack patients in the Medicare Advantage sample, we exclude heart attack from this analysis.

Table 5 Panel B presents the corresponding results when we replicate our main analysis on the sample of MA patients. Columns 1 and 2 present estimated effects on episode and first week costs, respectively. The coefficients imply that heart failure patients from high penalty salience hospitals receive either the same or differentially lower resource intensity. A slight decrease in resources for MA patients with targeted conditions could be due to the multitasking effect alluded to in Section II.A. This pattern is also found in the case of pneumonia patients, where we estimate a marginally significant reduction in episode cost in the first week.

Table 5 Panel B columns 3 and 4 present the corresponding results on the probability of readmission in 30 days and the first week, respectively. The coefficients do not display a consistent pattern across conditions or even within a condition. In the case of pneumonia, we find a positive and marginally significant coefficient on readmissions in the first week, which is plausible given the estimated reduction in resource intensity discussed above. But this pattern doesn't hold at 30 days for pneumonia or for heart failure patients.

(iii)   Placebo treatment effects

Our research design exploits the fact that HRRP incentivized improvement in readmissions only for patients in select conditions. However, one may worry that our results are capturing spurious effects. We assess this concern by generating an empirical distribution of treatment effects by

---

[21] All coefficients on episode cost and readmissions in Table 5 Panel A are statistically insignificant and small relative to the main estimates, except one. We find a large, positive coefficient on the probability of 30-day readmissions for heart failure patients. This could reflect sampling error or simply capture a mechanical pattern of greater readmission rates in markets with higher penalty rates.

randomly assigning a subset of the sample as heart failure patients and re-estimating our main specification 1,000 times. We then compare the observed effect on heart failure patients against this distribution.[22] Figure 4 presents the CDF of the estimated effects and indicates the true coefficient by a red line. The true coefficient is larger in magnitude than about 94% of the placebo coefficients. Hence, we interpret the main results as capturing a systematic effect on episode cost for heart failure patients.

c) Alternative characterization of penalty salience

We took a reduced form approach when we constructed our penalty salience measure and assumed that it increases in proportion to a hospital's share of the firm's patient volume. While this is intuitive, there are alternative ways to motivate this mechanism. For example, one may argue that hospitals will not bring up increased service standards with all their HHA partners. If enforcing these standards requires greater monitoring effort by the hospital, then it is costly for the hospital to do so. It follows that hospitals will prioritize monitoring for HHAs that account for a larger share of their patient discharges.

We replicate our main results using a modified measure of penalty salience to operationalize this interpretation. This measure, denoted $\tilde{P}_{hjc}$, also incorporates the firm's share of the hospital's patients sent to home health care in the baseline period, $\tilde{s}_{hj0}$, as shown below.[23]

$$\tilde{P}_{hjc} = penalty_{hc} * share_{jh0} * \tilde{s}_{hj0} \tag{3a}$$

$$\tilde{s}_{hj0} = \frac{n_{hj0}}{n_{h0}} \tag{3b}$$

In the baseline period, the patient weighted mean value of $\tilde{s}_{hj0}$ across office-hospital pairs was 0.41, with a standard deviation of 0.39.[24] However the firm's share of hospital patients is positively correlated with the hospital's share of the firm's volume ($share_{jh0}$), hence this alternative measure of penalty salience is only slightly smaller in magnitude than the main measure

---

[22] We exclude heart attack and pneumonia patients from the sample for this exercise to focus attention on heart failure where find a substantial increase in episode cost, as shown in Table 3. We retain heart failure patients in the sample in the spirit of a randomization test. So, in each iteration, some heart failure patients may also be assigned as placebo treated units.

[23] We obtained the total number of home health care episodes initiated by hospitals following the targeted conditions in 2012 using a 100% sample of home health claims data.

[24] This is calculated as the average of $\tilde{s}_{hj0}$ across all hospital-office pairs, weighted by the number of patients of this dyad served by the firm in the baseline period.

on average (mean of 0.017 vs. 0.02). Using this alternate measure as the key source of cross-hospital variation, we replicate our main results on episode cost and probability of readmission. Table 5 panel C presents the corresponding interaction coefficients. Reassuringly, these results have similar patterns – across conditions, outcomes, and time horizon, i.e., whether we consider the entire episode or first week only. The coefficients are also similar in size and imply similar magnitude effects as the main results. For example, the estimates imply a differentially greater cost of 4.9% ($e^{0.8} - 1 = 1.22 * 0.04 = 0.049$) for heart failure patients from a hospital with 1 s. d. greater penalty salience.

VI. Discussion and Conclusions

Performance pay has become a pervasive feature of provider contracts in US healthcare, although there remains considerable debate over design. Using patient health outcomes to anchor performance pay contracts exposes providers to greater risk but may also be more efficient if providers respond by deploying effective and low-cost solutions. To design policies optimally, policymakers need to factor in potential provider response to incentives. But there is little empirical evidence from the health care sector to guide policymakers. This paper begins to fill these gaps by examining whether a performance pay penalty imposed on hospitals passed through to post-acute care providers and resulted in greater use of resources outside the hospital.

We do so in the context of one of the largest performance-pay programs in US health care, the Hospital Readmissions Reduction Program (HRRP). Previous studies of this program have documented improvements in 30-day readmissions, however there is debate over the causal channels. Leveraging novel operational and claims data from a large home health care firm operating in 16 states, we find that the firm provided more inputs for patients targeted by HRRP, discharged from hospitals at greater penalty risk and which account for a greater share of the firm's patient volume. The data also permits us to identify three mechanisms used by the firm – frontloading care in the first week of the episode, using more experienced caregivers and spending more time with the patient, and targeting sicker patients.

The data is novel and allows us to observe changes in effort that would remain unobserved in traditional claims data. The results are also encouraging from a policy perspective. However, the use of data from one provider does limit the generalizability of the conclusions.

We find effects on episode cost for heart disease patients, but not for pneumonia patients. Within heart disease, our results at the 30-day mark are stronger for heart failure patients. However, prior studies on HRRP have found greater effects on readmissions for heart attack and pneumonia patients

relative to heart failure. (Desai et al., 2016; Mellor et al., 2016; Ziedan, 2019). Our results may differ because our sample is limited to home health care patients (about 16% of hospital stays for the targeted conditions). These episodes represent more complex patients and effects for this group may differ from population level effects documented in prior studies. Further, it is plausible that the low-intensity, frequent care provided by home health care disproportionately benefits patients with chronic conditions such as heart failure (Naylor et al., 1999; 2004).

These results indicate that even in the absence of vertical integration, downstream post-acute care firms can internalize penalty incentives imposed on hospitals, to the extent quality of care affects demand for their services. As post-acute care gains prominence in policy proposals, quantifying the welfare effects for society assumes more importance. This was outside the scope of the current study but remains an important avenue for future research.

References

AHHQI. 2012. "Comprehensive Analysis of Medicare Data Provides Roadmap to Savings in Post-Acute Care." Alliance for Home Health Quality and Innovation. Available at https://ahhqi.org/images/pdf/cacep-wp1-press-release.pdf.

Baker, George P. 1992. "Incentive Contracts and Performance Measurement." Journal of Political Economy 100 (3): 468–505. https://doi.org/10.1086/261831.

Candrilli, S. and J. Mauskopf, 2006, "How Much does a Day in the Hospital Cost?" Value in Health, 9, A56.

Carey, Kathleen, and Meng-Yun Lin. "Hospital Readmissions Reduction Program: safety-net hospitals show improvement, modifications to penalty formula still needed." Health Affairs 35, no. 10 (2016): 1918-1923.

Chandra, Amitabh, Maurice A. Dalton, and Jonathan Holmes. "Large increases in spending on postacute care in Medicare point to the potential for cost savings in these settings." Health Affairs 32, no. 5 (2013): 864-872.

Chen, Min, and David C Grabowski. 2017. "Hospital Readmissions Reduction Program: Intended and Unintended Effects." Medical Care Research and Review. SAGE Publications Sage CA: Los Angeles, CA, 1077558717744611.

Chernew, Michael E., Robert E. Mechanic, Bruce E. Landon, and Dana Gelb Safran. "Private-payer innovation in Massachusetts: the 'alternative quality contract'." Health Affairs 30, no. 1 (2011): 51-61.

Christianson, Jon B., Sheila Leatherman, and Kim Sutherland. "Lessons from evaluations of purchaser pay-for-performance programs." Medical Care Research and Review 65, no. 6_suppl (2008): 5S-35S.

David, Guy, Evan Rawley, and Daniel Polsky. "Integration and task allocation: Evidence from patient care." Journal of Economics & Management Strategy 22, no. 3 (2013): 617-639.

David, Guy, and Kunhee Lucy Kim. 2018. "The Effect of Workforce Assignment on Performance: Evidence from Home Health Care." *Journal of Health Economics* 59: 26–45. https://doi.org/10.1016/j.jhealeco.2018.03.003.

Desai, Nihar R, Joseph S Ross, Ji Young Kwon, Jeph Herrin, Kumar Dharmarajan, Susannah M Bernheim, Harlan M Krumholz, and Leora I Horwitz. 2016. "Association between Hospital Penalty Status under the Hospital Readmission Reduction Program and Readmission Rates for Target and Nontarget Conditions." JAMA 316 (24). American Medical Association: 2647–56.

Gupta, Atul. "Impacts of performance pay for hospitals: The readmissions reduction program." American Economic Review 111, no. 4 (2021): 1241-83.

HHCN. "2019 Home Health Care Outlook Report." Homecare Homebase. 2019. Downloaded from https://homehealthcarenews.com/survey/hhcn-hchb-2019-home-health-care-outlook-report/.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." Journal of Law, Economics, & Organization 7. JSTOR: 24–52.

Ibrahim, Andrew M., Justin B. Dimick, Shashank S. Sinha, John M. Hollingsworth, Ushapoorna Nuliyalu, and Andrew M. Ryan. "Association of coded severity with readmission reduction after the Hospital Readmissions Reduction Program." JAMA internal medicine 178, no. 2 (2018): 290-292.

Jha, Ashish K., Karen E. Joynt, E. John Orav, and Arnold M. Epstein. "The long-term effect of premier

pay for performance on patient outcomes." New England Journal of Medicine 366, no. 17 (2012): 1606-1615.

Kim, K. Lucy, Li Li, Meng Kuang, Leora I. Horwitz, and Sunita M. Desai. "Changes in Hospital Referral Patterns to Skilled Nursing Facilities Under the Hospital Readmissions Reduction Program." Medical Care 57, no. 9 (2019): 695-701.

Konetzka, R. Tamara, Elizabeth A. Stuart, and Rachel M. Werner. "The effect of integration of hospitals and post-acute care providers on Medicare payment and patient outcomes." Journal of health economics (2018).

Lazear, Edward P. "Salaries and piece rates." Journal of business (1986): 405-431.

Lehrman, S. and K. Shore, 1998, "Hospitals' Vertical Integration into Skilled Nursing: A Rational Approach to Controlling Transaction Costs," Inquiry-Blue Cross & Blue Shield Association, 35, 303–314.

Linertová, Renata, Lidia García-Pérez, José Ramón Vázquez-Díaz, Antonio Lorenzo-Riera, and Antonio Sarría-Santamera. "Interventions to reduce hospital readmissions in the elderly: in-hospital or home care. A systematic review." Journal of evaluation in clinical practice 17, no. 6 (2011): 1167-1175.

Lotven, Amy, 2012. "Home health alliance says analysis underscores cost-savings potential." Inside Health Policy. Available at https://ahhqi.org/media-center/14.

Medicare Payment Advisory Commission. "Report to the Congress: promoting greater efficiency in Medicare. June 2007." (2007).

Medicare Payment Advisory Commission. "Report to the Congress: Medicare Payment Policy." Chapter 8: Home Health Care Services, pp. 209–234. (2016)

Medicare Payment Advisory Commission. " Mandated report: Effects of the Hospital Readmissions Reduction Program. June 2018." (2018).

Mellor, Jennifer, Michael Daly, and Molly Smith. "Does it pay to penalize hospitals for excess readmissions? Intended and unintended consequences of Medicare's hospital readmissions reductions program." Health economics 26, no. 8 (2017): 1037-1051.

MetLife, 2011, "Market Survey of Nursing Home, Assisted Living, Adult Day Services, and Home Care Costs," Available at: https://www.aarp.org/content/dam/aarp/livable-communities/old-learn/health/the-metlife-market-survey-of-nursing-home-assisted-living-adult-day-services-and-home-care-costs-2011-aarp.pdf. Accessed November 28, 2018.

Mullaney, Tim, 2015. "Home Health Agencies Protect Referrals by Cutting Hospital Readmissions." Home health care news. Available at https://homehealthcarenews.com/2015/02/home-health-agencies-protect-referrals-by-cutting-hospital-readmissions/.

Mullen, Kathleen J., Richard G. Frank, and Meredith B. Rosenthal. "Can you get what you pay for? Pay-for-performance and the quality of healthcare providers." The Rand journal of economics 41, no. 1 (2010): 64-91.

Murtaugh, Christopher M., Partha Deb, Carolyn Zhu, Timothy R. Peng, Yolanda Barrón, Shivani Shah, Stanley M. Moore et al. "Reducing readmissions among heart failure patients discharged to home health care: effectiveness of early and intensive nursing services and early physician follow-up." Health services research 52, no. 4 (2017): 1445-1472.

Naylor, Mary D., Dorothy Brooten, Roberta Campbell, Barbara S. Jacobsen, Mathy D. Mezey, Mark V. Pauly, and J. Sanford Schwartz. "Comprehensive discharge planning and home follow-up of hospitalized elders: a randomized clinical trial." JAMA 281, no. 7 (1999): 613-620.

Naylor, Mary D., Dorothy A. Brooten, Roberta L. Campbell, Greg Maislin, Kathleen M. McCauley, and J. Sanford Schwartz. "Transitional care of older adults hospitalized with heart failure: a randomized, controlled trial." Journal of the American Geriatrics Society 52, no. 5 (2004): 675-684.

O'Connor, Melissa, Kathryn H. Bowles, Penny H. Feldman, Mary St. Pierre, Olga Jarrín, Shivani Shah, and Christopher M. Murtaugh. 2014. "Frontloading and Intensity of Skilled Home Health Visits: A State of the Science." *Home Health Care Services Quarterly* 33 (3): 159–75. https://doi.org/10.1080/01621424.2014.931768.

Ody, Christopher, Lucy Msall, Leemore S. Dafny, David C. Grabowski, and David M. Cutler. "Decreases In Readmissions Credited To Medicare's Program To Reduce Hospital Readmissions Have Been Overstated." Health Affairs 38, no. 1 (2019): 36-43.

Prendergast, Canice. "The provision of incentives in firms." Journal of economic literature 37, no. 1 (1999): 7-63.

Rahman, Momotazur, John McHugh, Pedro L. Gozalo, Dana Clay Ackerly, and Vincent Mor. 2017. "The Contribution of Skilled Nursing Facilities to Hospitals' Readmission Rate." Health Services Research 52 (2): 656–75. https://doi.org/10.1111/1475-6773.12507.

Rogers, Joanne, Mary Perlic, and Elizabeth A. Madigan. "The effect of frontloading visits on patient outcomes." Home Healthcare Now 25, no. 2 (2007): 103-109.

Rosenthal, Meredith B, Richard G Frank, Zhonghe Li, and Arnold M Epstein. 2005. "Early Experience with Pay-for-Performance: From Concept to Practice." JAMA 294 (14). American Medical Association: 1788–93.

Shleifer, Andrei. 1985. "A Theory of Yardstick Competition." The RAND Journal of Economics 16 (3): 319. https://doi.org/10.2307/2555560.

Study, Emily. 2014. "Home Health and Hospital Referrals: How to Get Your Agency on the Magic List." Home health care news. Downloaded from https://homehealthcarenews.com/2014/12/home-health-and-hospital-referrals-how-to-get-your-agency-on-the-magic-list/.

Sood, Neeraj, Peter J. Huckfeldt, José J. Escarce, David C. Grabowski, and Joseph P. Newhouse. "Medicare's bundled payment pilot for acute and postacute care: analysis and recommendations on where to begin." Health Affairs 30, no. 9 (2011): 1708-1717.

Wadhera, Rishi K., Karen E. Joynt Maddox, Jason H. Wasfy, Sebastien Haneuse, Changyu Shen, and Robert W. Yeh. "Association of the Hospital Readmissions Reduction Program with mortality among Medicare beneficiaries hospitalized for heart failure, acute myocardial infarction, and pneumonia." JAMA 320, no. 24 (2018): 2542-2552.

Worth, Tammy. "Home health on road to reducing readmits." Healthcare Finance, April 7 (2014).

Ziedan, Engy. 2019. "The Intended and Unintended Consequences of the Hospital Readmission Reduction Program."

(a): Penalty incentive
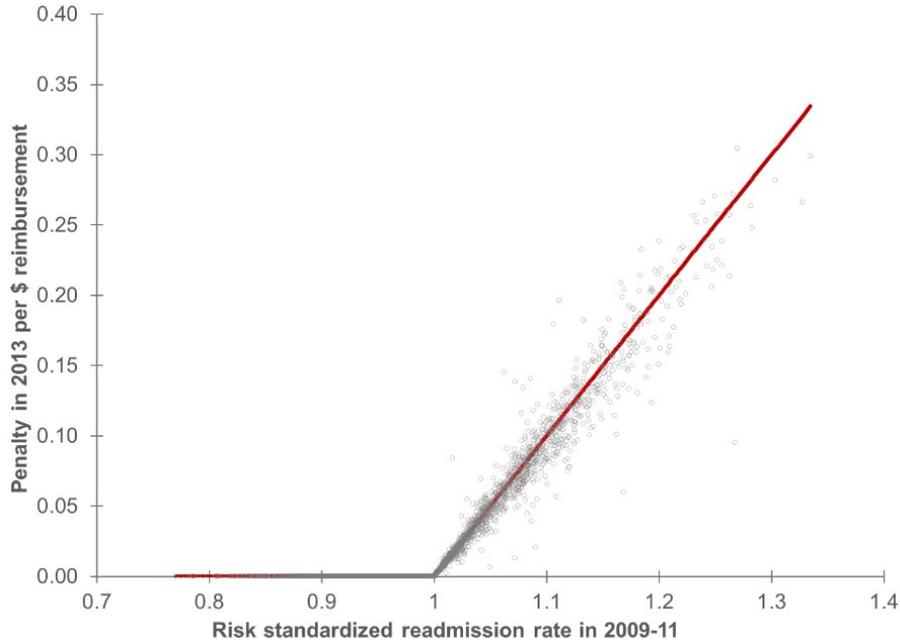
Figure 1: HRRP penalty incentive and mean readmission rate

Note: This figure provides necessary background about the HRRP program penalty incentive. It illustrates the penalty incentive introduced by HRRP using data from pneumonia patients. The figure plots our measure of the penalty rate in the first year of the program on the Y-axis against the risk-adjusted readmission rate over the period 2009–11 on the X axis. Each circle denotes an acute-care hospital. There are about 3,000 hospitals represented in the figure. We compute the penalty rate as the ratio of the penalty value assessed due to pneumonia in dollars divided by the mean annual fee-for-service Medicare revenue for the hospital from pneumonia during the baseline period (2009−11). Hence, it is the penalty per dollar of reimbursement, and we interpret it as the penalty rate. Following CMS, the risk-adjusted readmission rate is normalized by the mean, and this is known as the risk-standardized readmission rate (RSRR). The solid red line provides a schematic of the relationship between the RSRR and the penalty rate. The actual penalty formula depends on other factors and hence the circles deviate slightly from the red line. The incentive patterns are similar for other targeted conditions.

(a): Extensive margin use

(b): Number of PAC providers used
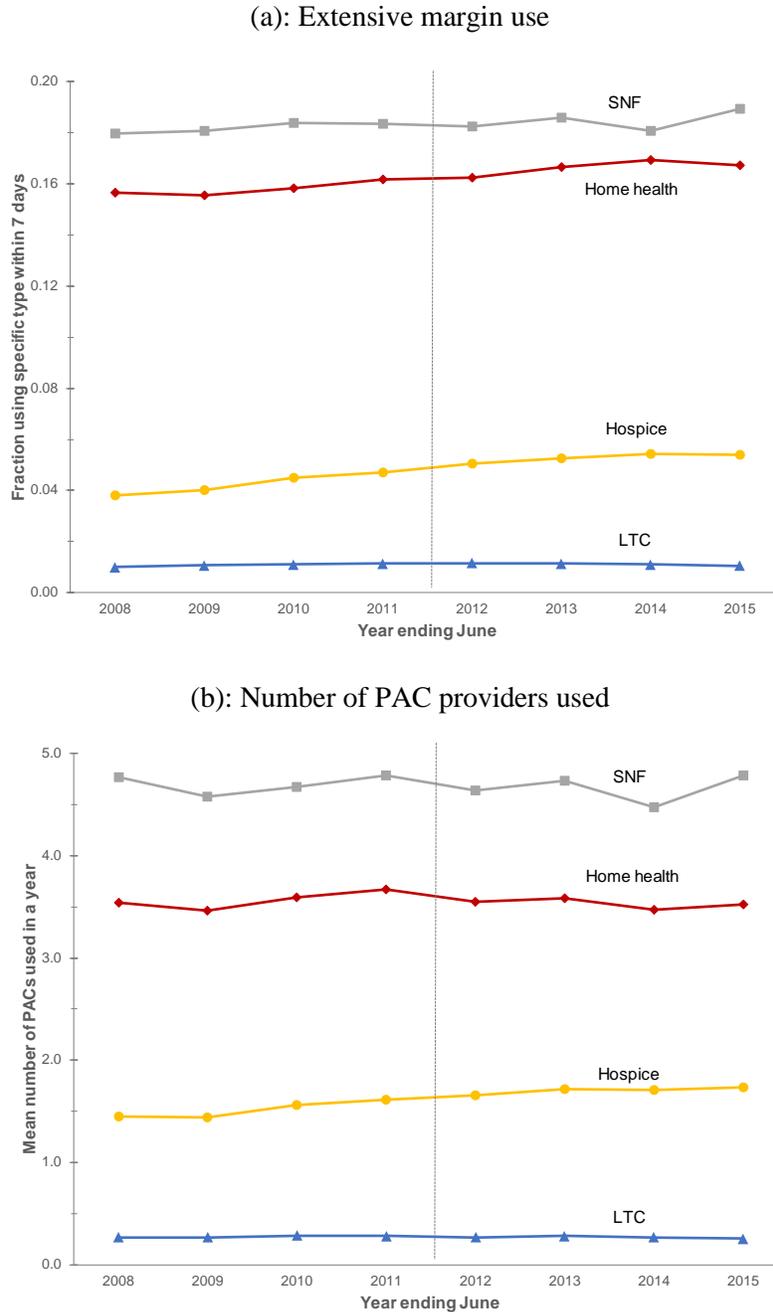
Figure 2: Use of PAC in Traditional Medicare

Note: This figure presents time series trends on the extensive margin use of post-acute care (Panel a) and the mean number of PAC providers used, by type (Panel b), using claims data from a 20% random sample of Traditional Medicare beneficiaries. Panel (a) presents the prevalence of four different types of PAC starting within seven days of discharge from hospital stays for the targeted conditions. The vertical dotted line indicates the formal announcement of HRRP details in August 2011. Panel (b) excludes hospitals that used only one PAC throughout the period. In both panels, years are defined as July–June periods rather than calendar years to be consistent with CMS computations under HRRP.

(a): Mean episode cost
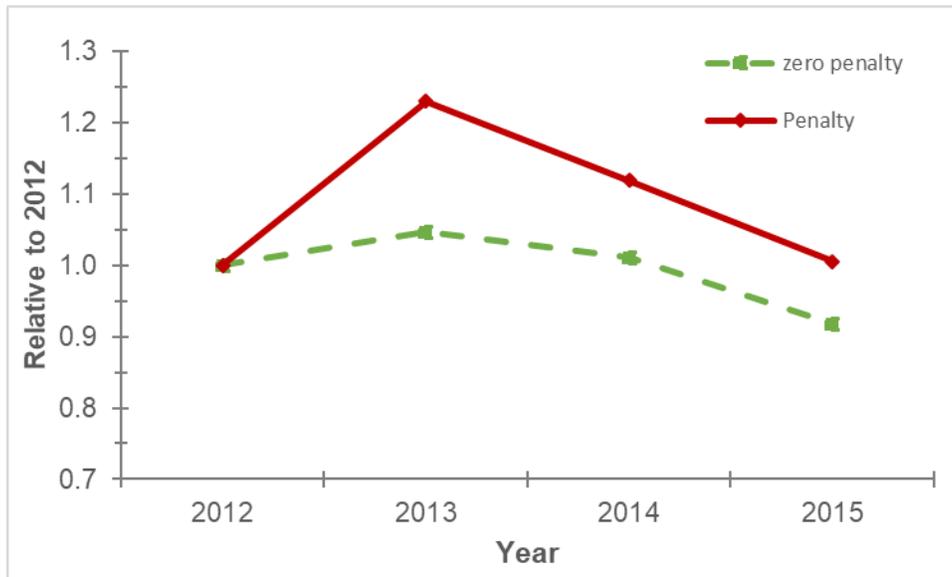
(b): Probability of readmission in 30 days
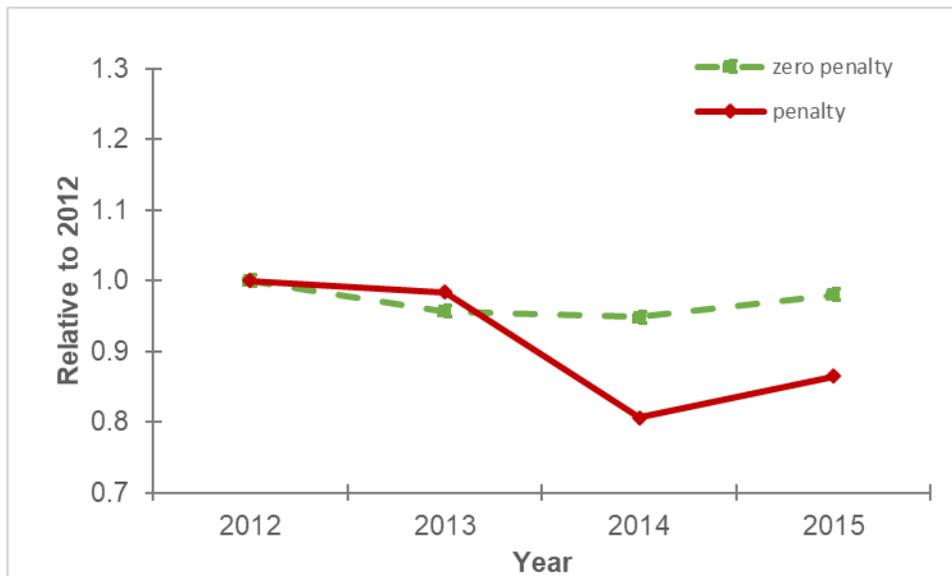
Figure 3: Time series of key outcomes by penalty salience

Note: This figure presents time series trends on the key outcome variables in the proprietary home health care data – mean episode cost (Panel a), and the probability of readmission in 30 days (Panel b), against calendar year on the X-axis. Both panels plot the time series separately for patients from hospitals with zero penalty salience (dashed), and non-zero penalty salience (solid). Since the levels of the two groups of hospitals are quite different, we normalize the values against their respective levels in 2012. Note that episode costs are deflated to be in 2015 dollars. The zero-penalty salience group includes patients from targeted conditions that were discharged from a hospital assessed zero penalty.

33

Note: p-value = .063
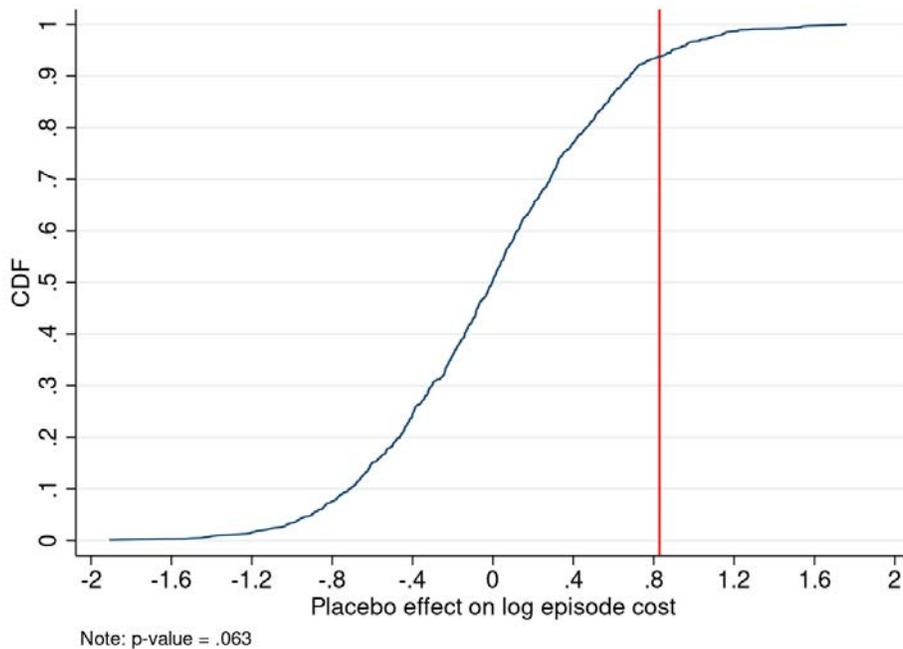
Figure 4: Empirical distribution of placebo effects on episode cost

Note: This figure presents the empirical distribution of placebo effects on log episode cost obtained by implementing a randomization test. Our main result finds an increase in episode cost for heart failure patients with a coefficient of about 0.8. We obtain a distribution of placebo effects and plot them relative to the observed effect on heart failure (indicated by the red line). Accordingly, we excluded heart attack and pneumonia patients from the sample but retained heart failure patients. We then estimated the model 1,000 times. In each iteration we randomly assigned a subset of patients as placebo heart failure patients, keeping the same proportion as the actual number of heart failure patients. The figure plots the CDF of the estimates obtained. It indicates that the true effect is larger than about 94% of the placebo estimates.

Table 1: Medicare sample descriptive statistics

	(1)	(2)	(3)	(4)
	Heart Attack	**Heart Failure**	**Pneumonia**	**Aggregate**
A. Hospital stays				
Total # of index cases	238,323	606,843	482,475	1,327,641
30-day readmission rate	0.18	0.22	0.17	0.20
Mean reimbursement ($)	17,284	9,725	8,640	10,688
B. Use of Post-Acute care				
Any post-acute care	0.36	0.40	0.39	0.39
Home health care	0.15	0.17	0.14	0.16
Skilled nursing care	0.16	0.18	0.20	0.18
Hospice	0.04	0.05	0.05	0.04
Long-term care hospital	0.01	0.01	0.01	0.01
C. Home health care episodes				
Total number of cases	32,673	98,916	63,288	194,877
Mean reimbursement ($)	2,444	2,556	2,622	2,560
Mean charge ($)	2,384	2,582	2,471	2,506
Mean # visits	15.1	16.3	15.6	15.8

Note: This table presents summary statistics on key outcomes from a 20% sample of Traditional Medicare claims pertaining to the conditions targeted by HRRP. Panel A describes hospital stays at general acute care hospitals with minimum 15 patients in the relevant condition over July 2008–June 2011, the benchmark for year 1 of the penalty. Panels B and C describe post-acute care use starting within 7 days of discharge from the index hospital stay. In Panel B, the same patient may use two types of post-acute care, and hence the sum of the individual categories may exceed the share of those using any post-acute care. The sample in Panel C is limited to home health facilities with 15 or more patients over the benchmark period, like the restriction used for hospitals. Total number of observations (Panels A, C) pertain to July 2008–June 2015. All mean values pertain to the benchmark period (2009–11). All spending/charges are deflated to be in 2017 dollars using the CPI-U.

Table 2: Proprietary data descriptive statistics

	(1) **Target conditions**	(2) **Non-target conditions**	(3) **All patients**
A. Descriptive statistics			
Observations (episodes)	2,183	6,844	9,027
Mean (S. D.) penalty per dollar of revenue	0.07 (0.09)	0	
Mean (S. D.) penalty salience	0.02 (0.04)	0	
Within-office S.D. in penalty salience:			
Across conditions	0.022		
Heart Attack	0.023		
Heart Failure	0.019		
Pneumonia	0.011		
30-day readmission rate	0.16	0.08	0.10
B. Care inputs			
Total cost index - mean	1280	1265	1269
Total cost index - S. D.	846	807	817
- Visit cost	762	756	758
- Personnel cost	485	473	475
Episode length (days)	30	27	28
Mean number of visits	12.9	12.6	12.7
C. Patient characteristics (all patients)			
Risk for hospitalization: History of 2+ falls	0.23	0.25	0.25
Risk for hospitalization: 2+ hospitalizations	0.46	0.31	0.34
Risk for hospitalization: Recent decline in Mental	0.07	0.07	0.07
Risk for hospitalization: Take 5+ medications	0.93	0.83	0.85
Age	80.7	77.6	78.4
Female	0.59	0.62	0.61
White	0.77	0.83	0.82
D. Referring hospital characteristics			
Share of hospitals > 300 beds	0.44	0.53	0.51
Not for-profit	0.88	0.86	0.87
For-profit	0.10	0.11	0.11
Teaching	0.70	0.69	0.69
Have a hospital-based HHA	0.14	0.14	0.14

Note: This table presents summary statistics on key independent variables and outcomes from the proprietary sample from the home health firm. Each observation is a patient episode. Columns 1 and 2 present values for the targeted and non-targeted conditions respectively, while column 3 presents results for all patients together. Panel A presents the number of observations and key metrics related to the readmissions reduction program—penalty per dollar of revenue (i.e., penalty rate normalized by inpatient base revenue from the condition), penalty salience (the product of penalty per dollar of revenue and the hospital's contribution to the office's patient flow over Jan–June 2012), and 30-day readmission rate. It also indicates the within-office variation in penalty salience across patients from the same targeted condition and across conditions. Panel B presents mean values of our key measures of home health care inputs. Panel C presents selected patient characteristics. Panel D presents characteristics of the 297 hospitals in the sample. All cost variables are deflated to be in 2015 dollars using the CPI-U.

Table 3: Provider effort and patient readmissions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln Total cost ($)		Ln Total time (min)		Ln Num of visits		Prob. Readmission	
	Overall	First week	Overall	First week	Overall	First week	30 Days	First week
1(Heart Attack)$_i$	-0.126	-0.144	-0.089	-0.119	-0.138	-0.133	-0.072	-0.046
	(0.135)	(0.119)	(0.121)	(0.127)	(0.118)	(0.098)	(0.100)	(0.071)
x P$_{hjc}$	-0.164	1.001***	0.040	1.018***	-0.102	0.813***	0.095	-0.022
	(0.286)	(0.200)	(0.247)	(0.148)	(0.247)	(0.157)	(0.163)	(0.079)
1(Heart Failure)$_i$	-0.025	-0.019	-0.031	-0.057	-0.028	-0.024	0.032	0.027
	(0.108)	(0.061)	(0.109)	(0.065)	(0.094)	(0.048)	(0.052)	(0.027)
x P$_{hjc}$	0.776***	0.403*	0.609***	0.145	0.654***	0.319	0.145	-0.258***
	(0.198)	(0.236)	(0.200)	(0.217)	(0.178)	(0.196)	(0.152)	(0.071)
1(Pneumonia)$_i$	-0.003	-0.077***	-0.008	-0.068***	0.007	-0.046**	0.072***	0.015
	(0.038)	(0.021)	(0.039)	(0.023)	(0.034)	(0.018)	(0.018)	(0.010)
x P$_{hjc}$	-1.151	-0.278	-0.921	-0.400	-0.940	-0.341	-0.244	0.372
	(1.384)	(1.292)	(1.295)	(1.146)	(1.173)	(1.045)	(0.566)	(0.365)
Mean dep. var.	1,280	407	584	199	12.9	3.8	0.16	0.05

Note: This table presents effects on resource use and patient readmissions obtained by estimating equation 2. Two sets of coefficients are presented—on the indicator for each condition, and the interaction of an indicator for the condition and penalty salience, P_{hjc}. For each outcome variable we present estimated effects over the entire episode (Cols. 1, 3, 5, and 7), and in the first week only (Cols. 2, 4, 6, and 8). About 8% of the episodes end within 7 days. We construct a total cost index (2015 dollars) based on cost accounting estimates provided by the home health care firm. The two main cost categories are visit cost (~60%) and personnel cost (~40%). Appendix Table A. 3 presents detailed results on these two elements of total cost. To interpret magnitudes, note that mean (s. d.) of penalty salience was 0.02 (0.04) across targeted conditions. Standard errors are presented in parentheses and are clustered by hospital and HHA office. The sample size for all regressions is 8,968.

Table 4: Heterogeneity by patient severity

	(1)	(2)	(3)	(4)
	Ln Total cost ($)		Probability of readmission	
	Episode	First week	30-days	First week
$1(sick)_i$	0.083***	0.001	0.046***	0.017***
	(0.022)	(0.016)	(0.008)	(0.005)
$1(\text{Heart Att.})_i$ x P_{hjc}	-0.068	1.077***	0.310***	0.025
	(0.284)	(0.165)	(0.112)	(0.052)
$1(\text{Heart Att.})_i$ x P_{hjc} x $1(sick)_i$	-0.033	-0.220	-0.585	-0.114
	(0.585)	(0.401)	(0.420)	(0.181)
$1(\text{Heart Fail.})_i$ x P_{hjc}	0.664*	-0.156	0.138	-0.100
	(0.390)	(0.373)	(0.250)	(0.123)
$1(\text{Heart Fail.})_i$ x P_{hjc} x $1(sick)_i$	0.303	0.973***	0.067	-0.253*
	(0.574)	(0.343)	(0.212)	(0.130)
$1(\text{Pneumonia})_i$ x P_{hjc}	-0.431	-0.052	-0.152	0.599
	(1.512)	(1.408)	(0.495)	(0.407)
$1(\text{Pneumonia})_i$ x P_{hjc} x $1(sick)_i$	-1.015	-0.350	-0.278	-0.560
	(3.052)	(2.627)	(1.184)	(0.715)
Mean dep. var.	1,280	407	0.16	0.05

Notes: This table presents regression results of the effects on episode cost (Columns 1 and 2) and probability of readmission (Columns 3 and 4) using equation 3 which exploits variation in observed patient sickness in addition to penalty salience, P_{hjc}. We construct an indicator $d_s = 1(sick)_i$ for the patient being above the median severity score, based on history of co-morbidities, risky behaviors, and utilization history. For brevity, not all coefficients are presented, but are available on request. To interpret the coefficients, note that mean (s. d.) of penalty salience was 0.02 (0.04) across targeted conditions. For each outcome, we also present results at the end of the first week of episode. About 8% of episodes end within the first week. Standard errors are presented in parentheses and are clustered by hospital and HHA office. The sample size for all regressions is 8,968.

	(1)	(2)	(3)	(4)
	Ln Total cost ($)		Probability of readmission	
	Overall	First week	30 days	First week
Panel A: Market-level penalty				
1(Heart Attack)$_i$ x P$_{hjc}$	0.552	0.527	-0.013	0.154
	(0.630)	(0.560)	(0.279)	(0.196)
1(Heart Failure)$_i$ x P$_{hjc}$	0.322	0.131	0.323***	0.057
	(0.208)	(0.182)	(0.119)	(0.082)
1(Pneumonia)$_i$ x P$_{hjc}$	0.234	-0.178	0.041	0.081
	(0.376)	(0.306)	(0.197)	(0.164)
Mean dep. var.	1,280	407	0.16	0.05
Panel B: MA sample				
1(Heart Failure)$_i$ x P$_{hjc}$	-0.971	-0.310	-0.257	0.156
	(1.058)	(0.589)	(0.321)	(0.270)
1(Pneumonia)$_i$ x P$_{hjc}$	1.410	-3.310***	1.561	1.507*
	(2.456)	(1.196)	(1.384)	(0.779)
Mean dep. var.	1,232	390	0.16	0.07
Panel C: Alternate measure of salience				
1(Heart Attack)$_i$ x P$_{hjc}$	-0.215	0.940***	0.050	-0.020
	(0.277)	(0.179)	(0.136)	(0.083)
1(Heart Failure)$_i$ x P$_{hjc}$	0.798***	0.399	0.089	-0.274***
	(0.199)	(0.253)	(0.130)	(0.065)
1(Pneumonia)$_i$ x P$_{hjc}$	-1.314	-0.749	-0.659	0.216
	(1.368)	(1.151)	(0.402)	(0.415)
Mean dep. var.	1,280	407	0.16	0.05

Notes: This table presents results using falsification tests (Panels A and B) and a robustness check (Panel C), respectively, as described in Section V.E. The two outcomes presented are episode cost (Columns 1 and 2) and probability of readmission (Columns 3 and 4). For each outcome variable we present effects on effort over the entire episode and in the first week. About 8% of episodes end within the first week. For brevity, throughout we present only the coefficients (β_c) on the interaction of penalty salience and indicators for targeted conditions in equation 2. Panel A shuts down variation in penalty rate and patient shares across hospitals and uses the mean penalty rate across local hospitals for all target patients at an office. Panel B uses Medicare Advantage patients and hence has a different sample and means. There are very few heart attack patients in the MA sample hence we exclude them. Panel C tests robustness to using an alternate measure of penalty salience. Standard errors are presented in parentheses and are clustered two-way by hospital and HHA office.

(a): Any post-acute care

(b): Home health

(c): Skilled nursing
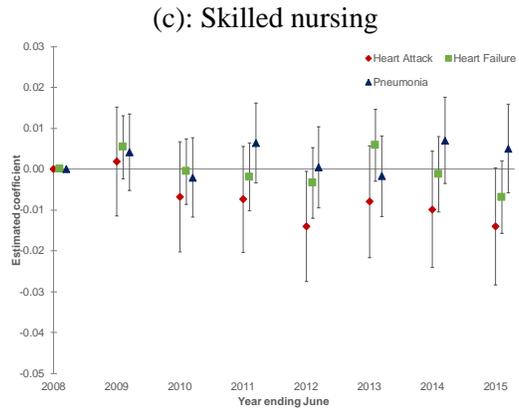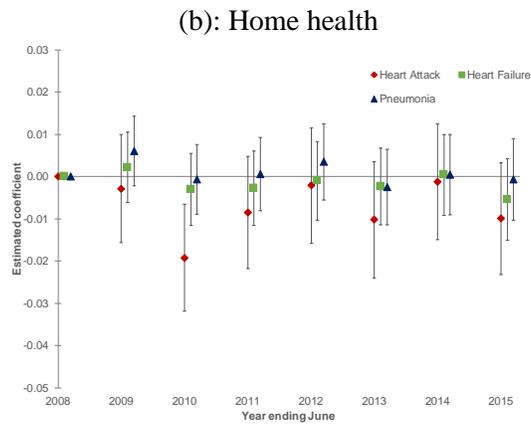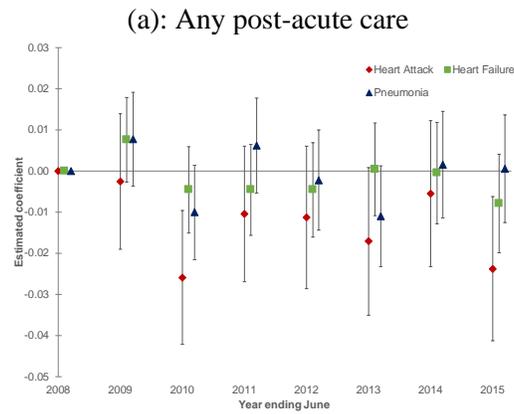
Figure A. 1: Estimated effects on extensive margin use

Note: This figure presents dynamic effects on differential use of any post-acute care (Panel a), home health (Panel b), and nursing home (Panel c) within 7 days following discharge from a hospital stay for heart attack, heart failure, and pneumonia for hospitals in the top tertile of penalty per dollar of revenue relative to other hospitals. To interpret magnitudes, note that the difference in mean penalty rate between top tertile and remaining hospitals was about 1.5x the standard deviation. Coefficients are estimated separately for each condition using Equation 5 on a 20% sample of Traditional Medicare beneficiaries. Models include hospital and year fixed effects. Standard errors are clustered by hospital.

Table A. 1: Descriptive statistics of patients by severity

Variables	Quartile 1 Least sick	Quartile 4 Most sick
Number of patients	205	230
A: Patient sickness		
Age	82	81
Female	0.54	0.60
White	0.84	0.72
Mean number of risks (score)	0.9	5.0
Risk for hospitalization: History of 2+ falls	0.01	0.49
Risk for hospitalization: 2+ hospitalizations	0.06	0.81
Risk for hospitalization: Recent decline in Mental	0.00	0.28
Risk for hospitalization: Take 5+ medications	0.76	0.97
P(Readmission) in 30 days	0.09	0.20
P(Readmission) in 7 days	0.03	0.05
B: Care inputs		
Total time spent with patient (min)	551	615
Number of visits	12.3	13.8
Episode length (Days)	28.9	31.7
P(Episode length > 7 days)	0.89	0.92
Total cost index ($)	1,199	1,355
C: First week only		
Total time spent with patient (min)	202	205
Number of visits	4.0	4.1
Total cost index ($)	417	429
D: After first week		
Total time spent with patient (min)	393	464
Number of visits	9.4	11.0
Total cost index ($)	881	1,049

Notes: This table presents descriptive statistics on key measures of resource intensity and readmissions for two groups of patients in the targeted conditions. This table uses patient data from the January through December 2012. The two groups of patients are from the lowest and highest quartiles according to a sickness severity measure. The measure is a score assigned to each patient by adding indicator values for different risk factors observed at the start of the episode – risks for hospitalization (2+ falls, prior history, recent decline in mental capacity, take 5+ medications), overall status (temporary, bad, or very bad), alcohol dependency, drug dependency, heavy smoking, obesity, and other indicators (disruptive behavior, impaired decision making, indwelling, intractable pain, memory loss, urinary incontinence). Some of these indicators are presented in Table 2 and in Panel A above. Panels C and D present statistics separately for the first week and beyond the first week, respectively.

Table A. 2: Proprietary data versus national sample

Variables	(1) All states	(2) Sample states	(3) Sample hospitals
Panel A: Hospital stays			
Heart attack	100,994	39,160	17,028
Heart failure	256,030	99,992	38,968
Pneumonia	211,997	76,049	26,302
Total	569,021	215,201	82,298
Post-Acute Care within 7 days			
% any PAC	38.3	40.7	42.8
% home health	15.6	16.8	18.9
Stays not sent to PAC			
Reimbursement ($)	10,810	10,605	11,754
30-day readmission rate	16.3	16.2	16.2
Stays sent to PAC			
Reimbursement ($)	13,720	13,235	14,776
30-day readmission rate	23.4	23.4	23.8
Stays sent to Home health			
Reimbursement ($)	13,126	12,531	13,798
30-day readmission rate	20.4	20.3	20.2
Panel B: Home health episodes			
Heart attack	12,883	5,523	2947
Heart failure	38,189	16,375	8150
Pneumonia	24,109	9,708	4381
Total	75,181	31,606	15,478
% Female	58.2	57.0	57.3
% White	85.9	87.5	85.3
Mean age	81.3	81.4	81.6
Panel C: HRRP penalty			
Number of hospitals	3,271	1,093	297
Heart attack	0.048	0.059	0.088
Heart failure	0.031	0.034	0.045
Pneumonia	0.029	0.030	0.042
Hospitals not penalized (%)	31.1	27.3	19.9

Notes: This table presents descriptive statistics on hospital stays (Panel A), home health episodes (Panel B) and HRRP penalties applicable in 2012 (Panel C). The descriptive statistics in Panels A and B were computed over the period July 2008–June 2011, the benchmark period for the first year of penalty. Panel A compares key variables for all Medicare hospital stays in the targeted conditions (Col. 1), with those in the 16 states (Col. 2), and the hospitals (Col. 3) represented in the sample provided by the home health firm. All values in Panel A (except volume) are simple averages across the three conditions. Reimbursement values are deflated to be in 2017 dollars. Panel B presents the number of home health episodes starting within 7 days of discharge from a hospital stay during the same period and basic demographic means. Panel C reports penalty per dollar of revenue (as described in Section II.B) across hospitals in the three groups.

Table A. 3: Effects on episode cost categories

	(1)	(2)	(3)	(4)
	Log(Visit cost)		Log(Personnel cost)	
	Overall	First week	Overall	First week
1(Heart Attack)$_i$	-0.133	-0.137	-0.143	-0.182
	(0.131)	(0.112)	(0.148)	(0.142)
x P$_{hjc}$	-0.170	0.819***	-0.135	1.326***
	(0.277)	(0.187)	(0.302)	(0.242)
1(Heart Failure)$_i$	0.001	0.021	-0.033	-0.043
	(0.104)	(0.065)	(0.118)	(0.073)
x P$_{hjc}$	0.709***	0.319	0.903***	0.540*
	(0.193)	(0.217)	(0.212)	(0.291)
1(Pneumonia)$_i$	-0.013	-0.085***	0.008	-0.072***
	(0.036)	(0.020)	(0.042)	(0.027)
x P$_{hjc}$	-0.751	0.112	-1.659	-0.636
	(1.341)	(1.251)	(1.529)	(1.542)
Mean dep. var. ($)	762	255	485	140

Notes: This table presents regression results on different elements of (log) episode cost for the home health firm obtained using equation 2. The coefficient of interest is on the interaction of indicator for each condition and penalty salience, P_{hjc}. For each outcome we present effects over the entire episode (Columns 1 and 3), and in the first week only (Columns 2 and 4). We disaggregate total cost into two elements—visit cost and personnel cost. To interpret the coefficients, note that mean (s. d.) of penalty salience was 0.02 (0.04) across targeted conditions. About 8% of episodes end within the first week. Standard errors are presented in parentheses and are clustered by hospital and HHA office. The sample size for all regressions is 8,968.

Table A. 4: Changes in PAC use by hospitals

	(1) Heart Attack	(2) Heart Failure	(3) Pneumonia
A. Any PAC	-0.0081	-0.0246	0.0081
	(0.019)	(0.029)	(0.024)
Y-Mean	0.36	0.40	0.39
B. Home health	0.0229	-0.0189	0.0093
	(0.017)	(0.023)	(0.019)
Y-Mean	0.15	0.17	0.14
C. Skilled nursing	-0.0347**	-0.0116	0.0014
	(0.016)	(0.020)	(0.020)
Y-Mean	0.16	0.18	0.20
D. Hospice	0.0013	-0.0096	-0.0039
	(0.008)	(0.011)	(0.010)
Y-Mean	0.04	0.02	0.03
E. Long-term care	0.0046	0.0151***	0.0061
	(0.004)	(0.004)	(0.005)
Y-Mean	0.01	0.01	0.02
Penalty per dollar of revenue			
Mean	0.05	0.04	0.05
S. d.	0.10	0.06	0.07
Observations	238,323	606,790	482,428

Note: This table presents estimated D-D effects on the use of different types of post-acute care, obtained independently for each condition using equation 4 (coefficient β). Each cell presents coefficients from a different regression. Models were estimated using the 20% sample of Medicare claims data. Penalty per dollar of revenue is the key RHS variable and constructed as described in Section II.B. All models include hospital and year fixed effects. Standard errors are clustered by hospital. Mean values pertain to the period July 2008–June 2011 (pre-HRRP). Table A. 5 presents corresponding triple difference effects where the third dimension is patient risk based on the history of hospital use.

Table A. 5: Heterogeneous effects on extensive margin use

	(1)	(2)	(3)	(4)	(5)	(6)
	Heart Attack		Heart Failure		Pneumonia	
A. Any PAC						
Pnlt rate x Post	-0.0081	-0.0061	-0.0246	-0.0222	0.0081	0.0195
	(0.019)	(0.022)	(0.029)	(0.038)	(0.024)	(0.030)
History of use		0.0307***		0.0291***		0.0654***
		(0.004)		(0.002)		(0.003)
Triple diff.		-0.0125		-0.0059		-0.0217
		(0.043)		(0.046)		(0.045)
Y-Mean	0.36		0.40		0.39	
B. Home health						
Pnlt rate x Post	0.0229	0.0186	-0.0189	-0.0325	0.0093	0.0032
	(0.017)	(0.019)	(0.023)	(0.032)	(0.019)	(0.025)
History of use		-0.0265***		-0.0346***		-0.0281***
		(0.003)		(0.002)		(0.002)
Triple diff.		0.0086		0.0233		0.0108
		(0.030)		(0.038)		(0.032)
Y-Mean	0.15		0.17		0.14	
C. Skilled nursing						
Pnlt rate x Post	-0.0347**	-0.0218	-0.0116	0.0125	0.0014	0.0271
	(0.016)	(0.018)	(0.020)	(0.028)	(0.020)	(0.023)
History of use		0.0403***		0.0446***		0.0642***
		(0.003)		(0.002)		(0.002)
Triple diff.		-0.0363		-0.0376		-0.0466
		(0.039)		(0.041)		(0.039)
Y-Mean	0.16		0.18		0.20	
Penalty per dollar of revenue:						
Mean	0.05		0.04		0.05	
S. d.	0.10		0.06		0.07	
Observations	238,323		606,790		482,428	

Notes: This table presents estimated effects on use of different types of post-acute care (within 7 days of discharge from the hospital) obtained using a 20% sample of Traditional Medicare claims described in Table 1. This table repeats the average effects first presented in Table A. 4 (cols. 1, 3, and 5) obtained independently for each condition using equation 4. It also presents triple difference coefficients for patients with recent history of hospital use (Cols. 2, 4, and 6). Penalty per dollar of revenue is calculated as described in Section II.B. All models include hospital and year fixed effects. Standard errors are clustered by hospital. Mean values pertain to the period July 2008–June 2011 (pre-HRRP).