

Does Hospital Leadership Matter?

Evidence from Pay-for-Performance

Hanna Glenn*

September 3, 2024

Abstract

Managers are often seen as drivers of firm decisions, strategies, and performance. Further, manager characteristics have been linked to employee wages, decisions of mergers and acquisitions, and financial performance. Health care policies in the US often focus on improving quality of care without increasing costs, and yet it is unclear whether hospital managers might contribute to this goal. In this paper, I estimate the differential response to such quality incentives by different types of nonprofit hospital executives: those with and without a clinical training background. I find that hospitals without clinically trained executives respond more to the financial incentives than those with clinically trained executives. I provide evidence that this may be due to differing preferences for profit vs. patient welfare. Thus, incentivizing clinically trained executives in hospitals could be an effective method to improve value in the US health care system.

*Department of Economics, Emory University, 1602 Fishburne Drive, Atlanta, GA 30322, hanna.glenn@emory.edu.

1 Introduction

In corporate environments, executives are considered crucial drivers of a company's practices and performance, as they are heavily involved in the culture, visions, and strategies of the firm. Executive heterogeneity has been linked to various actions and outcomes of firms; characteristics such as gender, age, and background are correlated with financial performance and behaviors such as employee diversity, wages, aggressiveness in mergers and acquisitions, and other corporate strategies.¹ However, there are still settings and relevant characteristics for which the relationship between management and firm behavior is not well understood.

In this paper, I study whether type of executive training affects firm behavior in the context of US hospitals. In the US, health care makes up almost 20% of spending (Medicare Medicaid Services 2024), and a major goal of policymakers is to improve the quality of health care without increasing costs. Yet, little is known about how hospital managers affect health care value, and it is unclear whether existing findings on executive characteristics can be extrapolated to the health care, and particularly hospital, context. Existing literature is generally limited to large, publicly traded firms, whereas a majority of US hospitals are nonprofit with presumably different objectives. In particular, health care executives have a particularly unique trade-off between profit and patient welfare, which varies from the types of firms studied in existing literature.

A characteristic relevant to this context is whether or not hospital executives have clinical training. Physicians can bring a unique combination of clinical and administration expertise, and they have insights that can potentially benefit their employees and lead to better quality (Stajduhar 2023, Ahmed 2022). However, hiring physician executives is controversial, as physicians are not always trained in business management, which can lead to poor practices and lower quality (Harvard Business Review 2018). Thus, I estimate how having clinically trained executives in hospital leadership affects hospital outcomes. I scrape executive names and titles of nonprofit hospitals in the US from publicly available Tax Form 990s. Combining this with other public hospital data, I construct a novel data set of nonprofit hospitals from 2010-2014 containing information on both executive team characteristics and quality metrics. While limited to nonprofits due to data availability, this sample of hospitals is an important driver of health care in the US. Private nonprofit hospitals make up 50% of all hospitals in the US, and staff, on average, 207 beds. For comparison, for-profits make up 36% of hospitals and staff 107 beds on average (ASPE 2023).

The non-random selection of leaders makes it difficult to interpret any direct comparison of dif-

¹Bertrand and Schoar 2003; Matsa and Miller 2011; Matsa and Miller 2011; Adams and D. Ferreira 2009; Ahern and Dittmar 2012; Flabbi et al. 2019; Benmelech and Frydman 2015; Custódio, M. A. Ferreira, and Matos 2013; Frydman 2019

fering executive teams as causal. Therefore, I leverage pay-for-performance policies enacted in the US as an exogenous shock to hospital incentives on quality, and compare the change in readmission and mortality rates of hospitals with and without clinically trained executives. Under the assumption that the two types of teams would have had similar trends in readmission and mortality rates absent the incentive change, I find that hospitals without clinically trained executives respond more to financial incentives on quality than hospitals with clinical executives. That is, while both types of hospital leadership teams decrease readmission rates after the policy change, those without clinical executives decrease readmissions at a faster rate than those with clinical executives. I find no evidence of selective patient practices driving this result.

I investigate whether the result might be explained by differing preferences on the trade-off between profit and patient welfare, where I model these hospital objectives and allow weight placed on each to vary with executive team type. Hospitals that place more weight on profit will naturally respond more drastically to a policy that adds quality directly into the profit function. Thus, a potential reason for the difference in response is that non-clinical executive teams are more profit driven. I investigate this empirically by comparing the response of hospitals with and without clinically trained executives to the response of for-profit hospitals, who have revealed that they are more profit driven by their ownership type. I find that nonprofit hospitals without clinically trained executives respond to the financial incentives on quality similarly to for-profits, while nonprofits without clinically trained executives respond differently than for-profits. For-profits and non-clinical executive teams change quality more in response to the incentive change than clinically trained executive teams.

Finally, the question remains whether clinically trained executives are a signal of unchanging underlying hospital preferences, or if they actively manage the hospital differently, effectively changing hospital objectives. I analyze this by leveraging changes in hospital propensity to hire clinical executives to decompose the full effect of clinical leadership into a signaling effect (clinical executives signal underlying hospital preferences), or managing effect (clinical executives manage the hospital differently). First, I show that hospital executive changes are not correlated with the change to financial incentives. For the signaling effect, I compare hospitals who ever have a clinically trained executive to hospitals who never have a clinically trained executive. For the managing effect, I limit to hospitals who ever have a clinically trained executive and compare those who have a clinically trained executive when the policies were announced to those who have clinically trained executives at other times. I find that the difference in response is driven by clinically trained executives managing the hospital differently.

Many researchers have documented a correlation between executive characteristics and firm

actions, starting with Bertrand and Schoar (2003), who show that manager fixed effects are an important driver of many firm behaviors and decisions. Having more female executives is correlated with female employee wages and corporate strategies (Flabbi et al. 2019; Matsa and Miller 2013); young male CEOs tend to be more aggressive in mergers and acquisitions, while those with military experience are less aggressive (Levi, Li, and Zhang 2010; Benmelech and Frydman 2015); CEOs with general ability tend to receive higher pay and perform better (Kaplan, Klebanov, and Sorensen 2012; Custódio, M. A. Ferreira, and Matos 2013; Adams, Akyol, and Verwijmeren 2018; Frydman 2019); older CEOs are associated with measures of corporate culture (Graham et al. 2022). Finally, Chief Diversity Officers are not found to have any effect on hiring more diversely in universities (Bradley et al. 2022).

Due to data limitations, many of these studies focus on large, publicly traded firms in the US. Brickley, Van Horn, and Wedig (2010) have the only study to my knowledge in the US nonprofit context considering hospital board of directors. Using exogenous variation in expected Medicare profits, the authors find that having an internal board of directors increases CEO compensation, and having physician board members decreases public donations (Brickley, Van Horn, and Wedig 2010). Additionally, two studies focus on hospital performance in other countries. Janke, Propper, and Sadun (2019) use data from England to study whether CEOs affect hospital production, and find no association (Janke, Propper, and Sadun 2019). However, Munoz and Otero (2022) investigate the role of CEOs in public hospitals in Chile, and finds an 8% decrease in mortality rates for hospitals with top managers. Further, a major contributor of this quality improvement is the exodus of older physicians as CEOs (Munoz and Otero 2022). I contribute to our knowledge of how executive team characteristics affect firm behaviors by (1) considering the important setting of nonprofit hospitals, which serve the majority of patients in the US, (2) considering a unique and relevant executive characteristic, and (3) observing firm behavior in the presence of an exogenous incentive change.

Additionally, I contribute to our understanding of how providers respond to pay-for-performance incentives in health care. The most in depth study of how hospitals respond to the Hospital Readmissions Reduction Program, one of the two large pay-for-performance initiatives, is Gupta (2021). This study finds that hospitals decreased readmissions by 5% and mortality rates by 2% on average as a result of the program, confirming prior studies (Mellor, Daly, and Smith 2017; Ziedan 2018; Ody et al. 2019; Gupta 2021). Around 40% of the decrease in readmissions is due to selective patient practices. Further, market concentration and hospital systems play a large role in how hospitals respond to the program (Kunz et al. 2024). Research on the other large pay-for-performance initiative, the Hospital Value Based Purchasing Program, generally finds that the program had no effect on underlying hospital quality (Office 2015; Norton et al. 2018; Friedson, Horrace, and Marier

2019). This paper contributes to our understanding of how characteristics of hospitals affect response to pay-for-performance incentives.

2 Setting and Data

I use several data sets to analyze the relationship between executives with clinical training and hospital response to financial incentives on quality. I construct novel information on individual hospital executives by scraping publicly available tax forms, and merge this data to publicly available data sets containing other hospital characteristics: the American Hospital Association Survey (AHA), the Centers for Medicare and Medicaid Services (CMS) Hospital Compare data and Case Mix Index files, and Healthcare Cost Report Information System (HCRIS) data. Ultimately, I employ this hospital-level data spanning from 2010-2014, which contains executive team characteristics, mortality and readmission rates as measures of quality, and other characteristics. In this section, I describe the institutional setting and data collection process. A more detailed description of the construction of this data can be found in Appendix A.

2.1 Nonprofit Hospital Executives

Nonprofit hospitals, as with most nonprofits, are typically governed by a board of directors, whose role is to set broad goals and strategies, and provide general oversight. The board selects executives, the highest level of management of a firm, to carry out day-to-day operations. An executive team usually consists of at least a Chief Executive Officer (CEO) and Chief Financial Officer (CFO), but there is variation in how firms organize executive teams. Some hospital executives specialize in health care settings by earning a degree in health care management, or an MBA specific to health care. There are executive positions that are often filled by someone with clinical training, such as a Chief Medical Officer (CMO), but medical doctors can also fill other executive positions. While some doctors earn additional degrees before stepping into an executive role, this is not a necessary condition to becoming a physician executive.

Hospital executive teams are understudied partly due to the inaccessibility of granular data. I construct a novel data set of nonprofit hospitals in the US which contains names, titles, and positions of all board members and executives tied to the hospital from 2009-2015. I gather this data from publicly available Tax Form 990s, which every sufficiently large nonprofit files each year in the US. To my knowledge, this is the first large scale gathering of executive names from these forms.² All

²Brickley, Van Horn, and Wedig (2010) collects compensation data for a select number of hospitals.

tax-exempt organizations in the US are required to file a Form 990 with Internal Revenue Services (IRS) each year. There are different types of forms, but any organization grossing over \$200,000 must file the most extensive Form 990. Sections of this form include a statement of revenue, statement of functional expenses, a balance sheet, and, as used in this project, a list of all key employees, executives, and board members.

For each hospital in the tax form data, I merge to hospital matches in the AHA data based on name and location.³ For all hospitals that can be matched to a general hospital in the AHA data, I use optimal character recognition (OCR) methods to extract the names, titles, and positions of each listed executive from its Tax Form 990. While board members are an important aspect of hospital organization, I do not include them in this analysis as they are beyond the scope of this paper. I extract text spanning 2009-2015, but I retain a much larger sample of hospitals by limiting to hospitals that have text information in 2010-2014. This yields 852 nonprofit hospitals that are matched to AHA data, and with complete text information in each year.

Finally, I verify clinical training of those who are identified as doctors, and obtain additional information about clinically trained executives by matching names to the National Plan and Provider Enumeration System (NPPES) database of all registered physicians. I present a table of means for various characteristics of doctor executives in the data in Table 1. In total, there are 813 clinically trained executives. On average, they are 52 years old, and 11% of clinically trained executives are female. While it is more likely they are a Chief Medical Officer than a Chief Executive Officer, still less than a third are CMOs. The most common specialty is internal medicine.

I focus on executive teams as a whole instead of specific positions because the organization of teams varies so drastically across hospitals. It is not uncommon for two hospitals to have the same position but the tasks performed are drastically different, or to have two distinct positions that essentially perform the same tasks. I create various hospital-level characterizations based on their executive team, such as the existence of a clinically trained executive, the number of clinically trained executives, the number of total executives, and whether the hospital has a Chief Medical Officer.

2.2 Pay-for-Performance Policies Targeting Readmission and Mortality Rates

The purpose of this paper is to investigate hospital behavior in the midst of financial incentives that change how hospitals receive payment for quality of care. Two programs were passed as part of

³I assess the observable differences between matched and non-matched hospitals in Appendix ??, and conclude that the samples are overall similar apart from an under-sampling of hospitals that belong to systems.

Table 1: Clinical Executive Summary Statistics

Variable	Mean
Age	52.5
Female	0.11
CEO	0.16
CMO	0.31
Specialty	
Internal Medicine	0.32
Surgery	0.09
Anesthesiologist	0.04
Family Medicine	0.17
Emergency Medicine	0.11
Other	0.28
Num. Clinical Executives	813

the Affordable Care Act that focused on pay-for-performance incentives for hospitals: the Hospital Readmissions Reduction Program (HRRP), and the Hospital Value Based Purchasing Program (HVBP). HRRP focused on penalizing hospitals with poor quality, and HVBP focused on rewarding hospitals with high quality or show improvement in quality.

In October 2011, the Center for Medicare and Medicaid Services (CMS) released a set of rules under HRRP mandating penalties for hospitals with above average readmission rates. The goal of HRRP is to lower readmissions through better care coordination, less initial stay complications, and better post-care instructions. Beginning in October 2012, hospitals with higher readmission rates than the national average in pneumonia, heart failure, or AMI (after adjusting for demographic characteristics) receive a fixed lower reimbursement rate for all Medicare patients seen in their hospital. Penalties are given in the form of a fixed rate reduction of 1-3% for every Medicare patient regardless of the condition. Excess readmission rates are calculated using a rolling look-back period of 3 years to determine whether the hospital is penalized. Therefore, hospitals had incentive to react immediately once details of the program were announced in October of 2011.

The HVBP Program instead rewards hospitals with high quality or significant improvement in quality. Specifically, CMS deducts Medicare payments by 2% from all eligible hospitals, collects this sum, and divides it among the rewarded hospitals. Several quality and cost measures surrounding safety, efficiency, cost reductions, clinical outcomes, and community engagement are combined to create a single score metric for each hospital. Hospitals are then compared to the average and are rewarded for being above average quality or for showing improvement (CMS 2023b).

I focus on outcomes of readmission and mortality rates, which are publicly available at the hospital-condition level in CMS Hospital Compare. I combine rates for pneumonia, AMI, and heart failure (the relevant HRRP conditions) into a weighted average based on the number of patients in each condition.

2.3 Summary Statistics

In Table 2, I present means of relevant variables for several hospital samples used throughout the paper. First, in column (1), are means for hospitals who ever have a clinical executive at any point in the sample. In column (2) are means for hospitals who have a clinical executive for the entire sample period, and column (3) shows means for hospitals who never have a clinical executive in the sample period. The main analysis compares hospitals in columns (2) and (3), and supplemental analyses leverage hospitals in column (1).

Comparing columns (2) and (3), hospitals that never have a clinical executive are smaller, less likely to be an academic medical center, and less likely to be affiliated with a system. Further, they have smaller executive teams on average by 1-2 people, and are less likely to have anyone acting as a designated CMO. Finally, in regards to exposure to pay-for-performance programs, hospitals who never have clinical executives are less likely to receive benefits from the HVBP Program, but seem more similar to hospitals with clinical leadership in terms of penalties from HRRP.

Next, I plot weighted average readmission rates, mortality rates, and patient case mix index across time for different hospital types in Figure 1. The effective start year of pay-for-performance incentives is shown as a dotted line in 2012. Average readmission rates trend similarly prior to 2012, and all hospital types exhibit a decrease in raw readmission rates after 2012. However, there is a faster rate of decrease for hospitals who have a clinically trained executive at some point, or who never have a clinically trained executive. When considering mortality, there is not a drastic change in 2012 for any hospital type, but hospitals who always have a clinical executive continue on an increasing trend in mortality where the other hospitals types remain more flat. Finally, hospitals that never have clinical executives consistently have less complex patients, but there is no observable change in raw case mix index over time.

Table 2: Summary Statistics by Leadership Team Type

Variable	(1) Ever Clinical Exec	(2) Always Clinical Exec	(3) Never Clinical Exec
Hospital Characteristics			
Academic Med. Center	0.41	0.52	0.24
Number Beds	231	222	116
Physician Owned	0.01	0.02	0.02
System Affiliated	0.58	0.52	0.44
Executive Team Characteristics			
Number Executives	4.85	5.25	3.11
Fraction Clinical Execs	0.20	0.26	0
Fraction Int. Medicine Execs	0.06	0.10	0.0
Has a CMO	0.53	0.45	0.15
Has Clinical CEO	0.25	0.11	0.0
Penalty/Payment Variables			
Ever Received HVBP Incentive	0.79	0.75	0.62
Penalized for AMI	0.05	0.04	0.03
Penalized for HF	0.05	0.07	0.05
Penalized for Pneumonia	0.08	0.09	0.05
Penalized for AMI + HF	0.06	0.05	0.04
Penalized for AMI + Pneumonia	0.03	0.04	0.02
Penalized for HF + Pneumonia	0.41	0.39	0.34
Penalized for All Conditions	0.30	0.27	0.22
Num. Hospitals	391	56	445

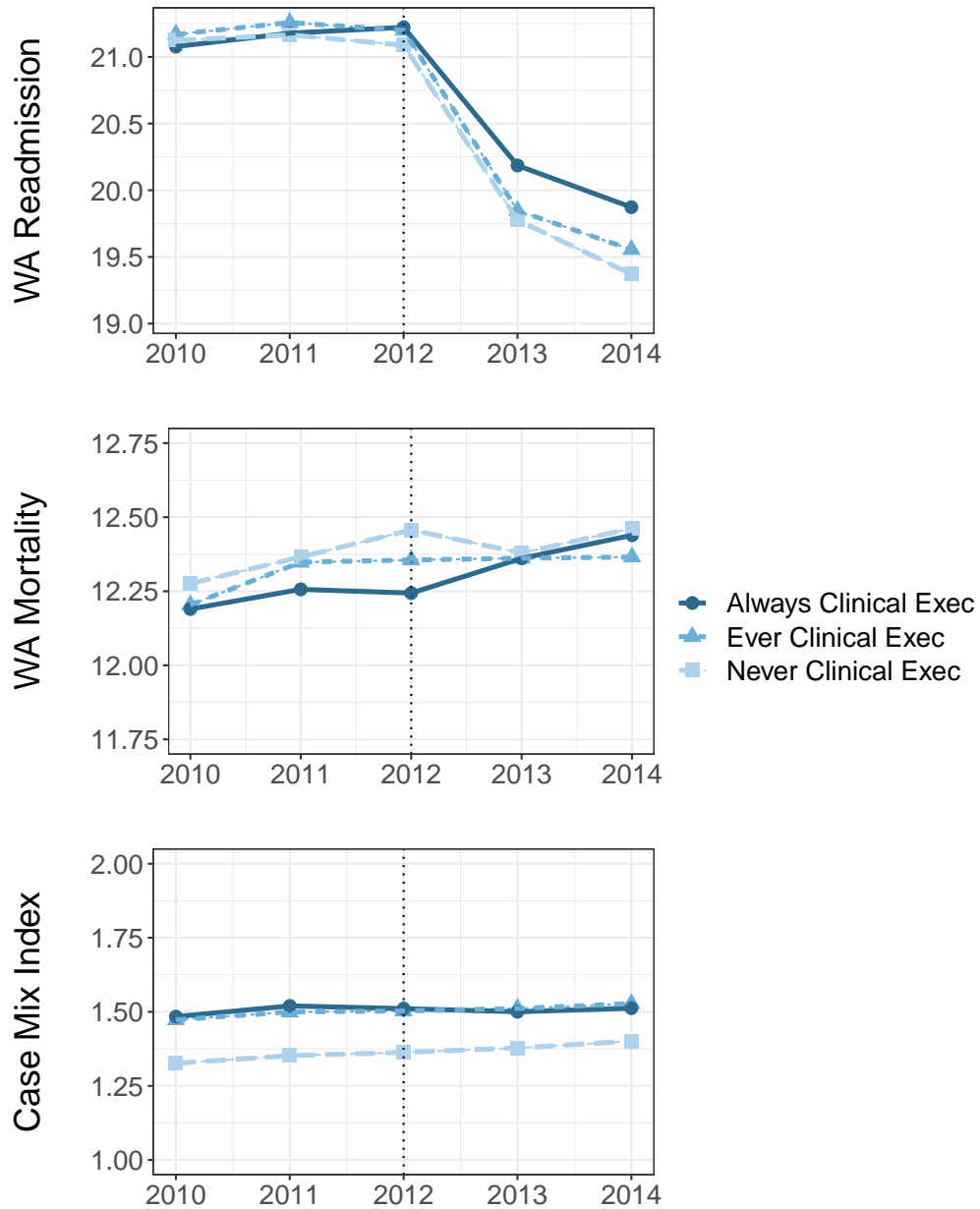
3 Effect of Clinically Trained Executives on Response to Financial Incentives on Quality

I test empirically whether having clinically trained executives matters in how hospitals respond to the pay-for-performance policies effective in 2012. The quality metrics targeted by the programs are readmission and mortality rates, so I use these as outcomes. I limit the sample to hospitals who always have clinical executives or never have clinical executives. That is, the hospitals in this analysis do not change their propensity to hire a clinical executive during the sample period.⁴

I ultimately employ a synthetic difference-in-differences estimation strategy that mirrors two-

⁴An important assumption for this analysis is that this sample selection is not correlated with the program enactment. I investigate this assumption in Section 5.1, and find no evidence that hospitals are more or less likely to change their propensity to hire a clinical executive because of the pay-for-performance programs.

Figure 1: Outcomes Over Time



way-fixed effects, where the key independent variable is the interaction between having clinically trained executives and being in years after pay-for-performance took effect. All hospitals had incentive to improve quality in response to the policy change, as penalties and payments are determined by comparing quality among hospitals. Thus, I leverage the timing of the policy change, rather than the policies themselves.

One may be concerned about potential differences in quality trends between clinical and non-clinical executive team hospitals leading to a violation in the parallel trends assumption. This is why I implement a synthetic difference-in-differences estimation strategy (Arkhangelsky et al. 2021). In particular, I weight more heavily any non-clinical team hospitals with pre-trends similar to clinical team hospitals in readmission, mortality, and baseline case mix. Further, I more heavily weight time periods that balance pre- and post-program outcomes for non-clinical teams. The control group consists of approximately 240 hospitals, and I show the fairly uniform distribution of their weights in Appendix A.4.

I present the unweighted two-way-fixed-effects specification equation for clarity, and the identification assumptions are conceptually similar for synthetic difference-in-differences. In Appendix C.4, I show that the findings are robust to using two way fixed effects, matching, or synthetic control.

$$y_{ht} = \beta (\text{clinical training x post 2012})_{ht} + \gamma_h + \delta_t + \epsilon_{ht} \quad (1)$$

The variable y_{ht} is one of the outcome variables capturing readmission or mortality rates discussed in Section 2.2, and γ_h and δ_t are hospital and time fixed effects, respectively.

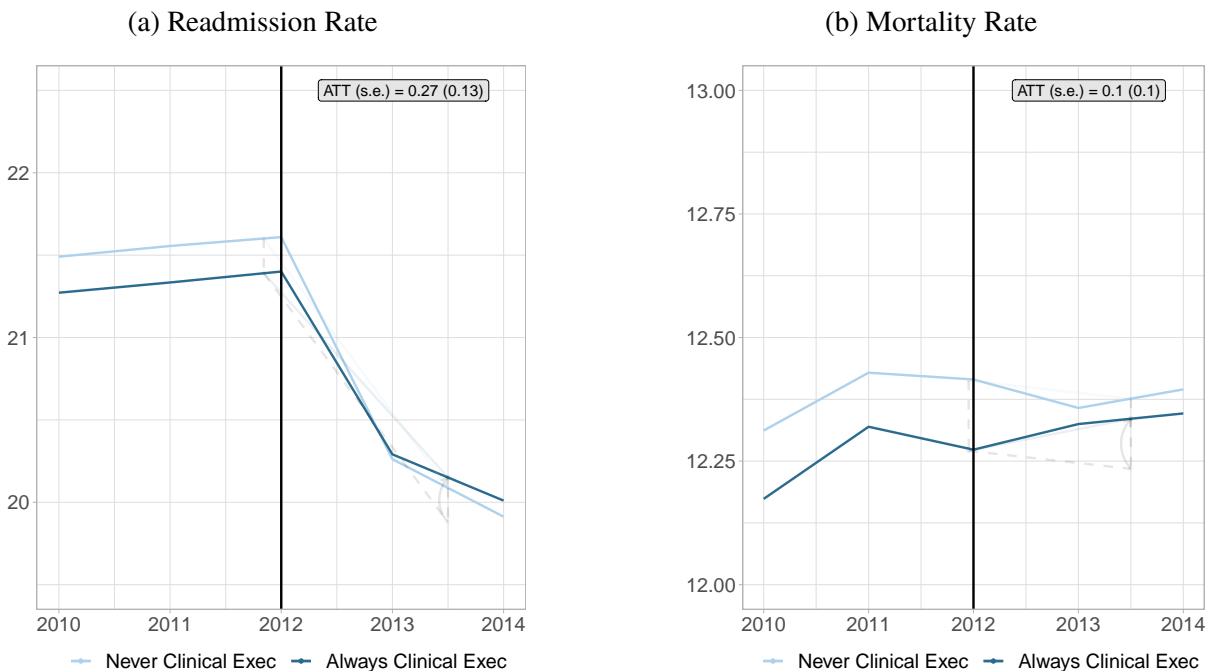
Once incorporating the weights used in the synthetic differences-in-differences method, we ensure parallel pre-trends in the outcome variables. Thus, I assume that under the weighted clinical and non-clinical hospital composition, trends would have continued to be parallel absent the incentive change in 2012. This includes an implicit assumption that no other event occurred in 2012 correlated with both the outcomes and composition of leadership team. Finally, I assume that there is no anticipation of the policy change before 2012. The rules of the policies were announced in October of 2011, so 2012 was the first full year that hospitals could meaningfully respond to the policies.

The estimated difference in readmission and mortality rates between hospitals with and without clinical executives are shown in Figure 2. Estimate β is represented in the difference in slope after 2012 when pay-for-performance incentives began. Panel 2a shows the estimated difference in readmission rates, and reveals that non-clinical teams decrease readmissions by .3 (standard error .13) more than clinical teams. This is a relatively small magnitude given an average readmission rate of 21, but reveals a meaningful difference in behavior given that both hospital types significantly decrease readmissions. In Appendix C.3, I show that these results are primarily driven by pneumonia patients.

In panel 2b, I show the estimated difference in mortality. While the magnitude of the difference is similar to that for readmission rates (relative to the mean), mortality is a noisy measure, and I cannot rule out null effects. I investigate whether the estimated difference in readmission rate

response could be due to differential participation in selective patient practices in Appendix C.2, and find that this is not the case.

Figure 2: Effect of Clinical Training on Readmission and Mortality Rates



3.1 Intensive Margin

Additionally, I estimate the intensive margin of clinical leadership: whether the fraction of clinically trained executives matters for response to financial incentive changes. I categorize hospitals with clinically trained executives into two bins: above the median and below median fraction of clinically trained executives. Again, I present the unweighted version of the synthetic difference-in-differences that I ultimately employ:

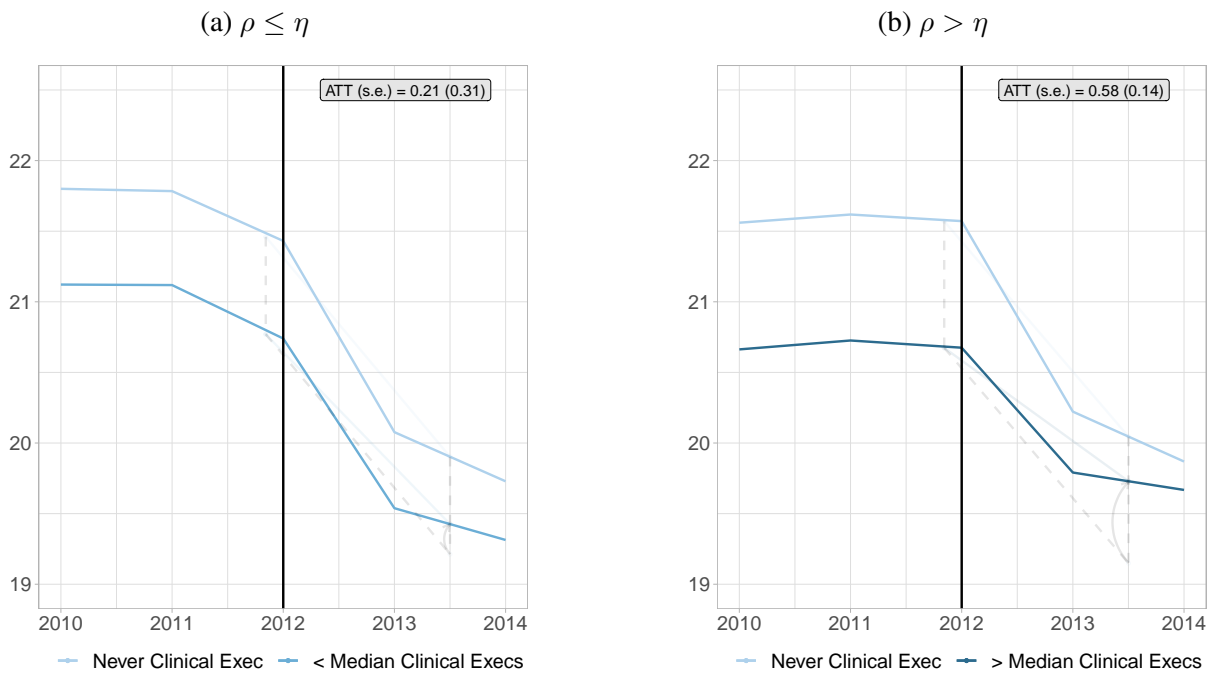
$$y_{ht} = \beta (\mathbb{1}\{\rho \leq \eta\} \times \text{post 2012})_{ht} + \gamma_h + \delta_t + \epsilon_{ht}, \text{ and} \quad (2)$$

$$y_{ht} = \beta (\mathbb{1}\{\rho > \eta\} \times \text{post 2012})_{ht} + \gamma_h + \delta_t + \epsilon_{ht}, \quad (3)$$

where ρ is the fraction of clinical executives on hospital h 's executive team, η is the median fraction of clinically trained executives, which is .2 in this sample, and γ and δ are hospital and time fixed effects.

I present the results of this estimation with readmission rates as the outcome in Figure 3. In panel 17a, I compare executive teams with no clinical training to those with a fraction of clinically trained executives up to the median value, 20%. With an average treatment effect of .21 (standard error .31), hospitals having this range of clinically trained executives are not driving the main finding. However, in panel 17b, I compare hospitals with no clinically trained executives to hospitals with a fraction of clinically trained executives over the median of 20%. I find that hospitals without clinically trained executives decrease readmissions by .6 (standard error .14) more than hospitals with clinically trained executives. Thus, the readmission rate response is driven by having a higher proportion of clinical executives relative to no clinical executives.⁵

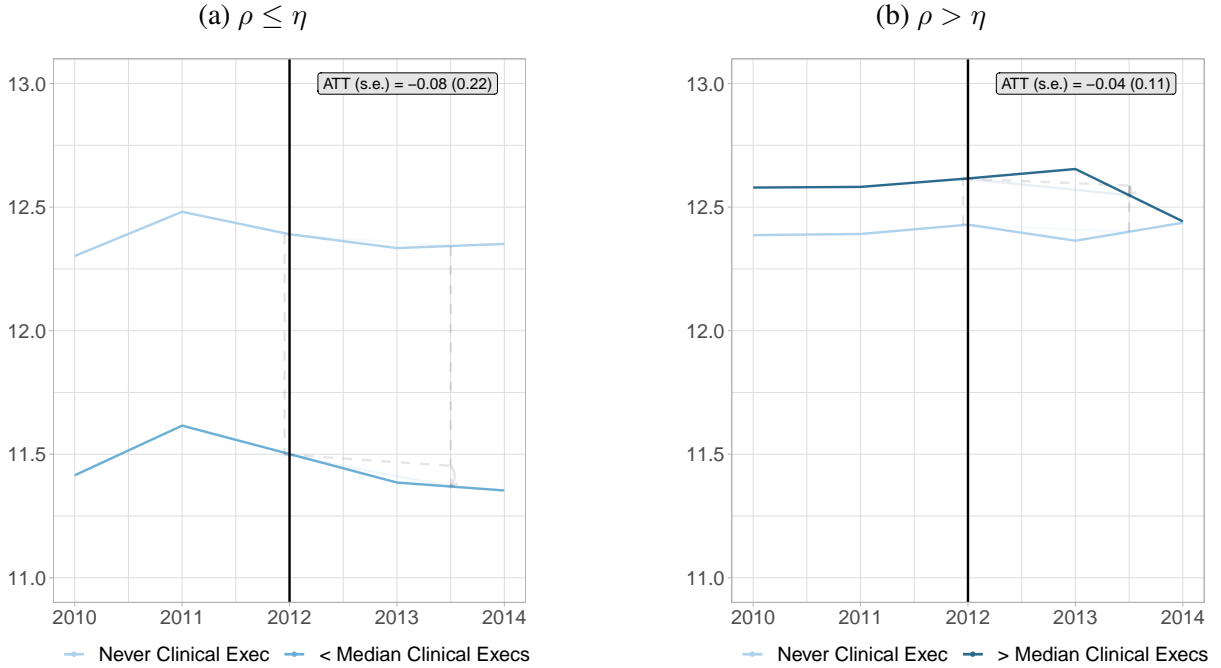
Figure 3: Effect of Clinical Training on Readmission Rates, Binned Treatment



I present the estimation results for mortality as an outcome in Figure 4. Unsurprisingly given the noisiness in the main findings for mortality, it is unclear whether there is any intensive margin effect for mortality. Both average treatment effect magnitudes are lower than that for the main result with high standard errors. Thus, the intensive margin does not seem to matter as much as it does for decreasing readmission rates.

⁵This result holds even when limiting the sample to hospitals with a similar total executive team sizes (Appendix C.1)

Figure 4: Effect of Clinical Training on Mortality Rates, Binned Treatment



4 Are Clinical Executives Less Profit Driven?

One hypothesis that could explain the differential response due to clinical leadership is how profit-driven a clinician is compared to other training backgrounds. To demonstrate this, consider hospital behavior under the enactment of pay-for-performance policies, where hospitals face a trade-off between profit and societal/patient benefit. Pay-for-performance incentives essentially incorporate quality directly into a hospital's profit function (Dranove 2011), whereas a fee-for-service system only incentivizes quantity. A hospital that places all value on profit will only care about quality when it enters into the profit function, and therefore will respond more drastically to pay-for-performance than a hospital who already cared about quality through societal benefit before the incentives.

To formalize this intuition, I specify the hospital objective function as a weighted average of profit and societal (or patient) outcomes and derive characteristics of the optimal quality. For simplicity, I abstract away from quality having a direct effect of quantity and price. Without any financial incentives to increase quality, hospital quality is decided by:

$$\max_{\theta} \alpha\pi(\theta) + (1 - \alpha)u(\theta),$$

where $\pi(\theta)$ is a profit function depending on θ , and the relationship between profit and θ depends

on pay-for-performance incentives. Additionally, $u(\cdot)$ is extra utility gained from quality, which is increasing and concave in θ , and $\alpha \in [0, 1]$ captures the weight placed on profit vs. societal benefit. There is an implicit stay-open condition throughout the model. Taking the first order condition yields

$$(1 - \alpha)u'(\theta) = -\alpha\pi'(\theta),$$

where marginal benefit equals marginal cost of increasing quality. Solving for $u'(\theta)$ and differentiating with respect to α :

$$\frac{du'(\theta)}{d\alpha} = \frac{-1}{(1 - \alpha)^2} \pi'(\theta)$$

First, assume the hospital is in a strictly fee-for-service setting, where quality benefits patients, but does not directly provide profit to the hospital through reimbursements. Then, assuming there is a positive cost to increasing quality, $\pi(\theta)$ is decreasing in θ . Then $\frac{du'(\theta)}{d\alpha} > 0$, revealing that in a fee-for-service setting, the more weight placed on profit, the lower quality of care chosen by the hospital.

Second, assume the hospital is in a strictly pay-for-performance setting, where quality benefits patients and directly increases profit. Then, assuming the marginal financial benefit of increasing quality is greater than the marginal financial cost, $\pi(\theta)$ is increasing in θ . Then $\frac{du'(\theta)}{d\alpha} < 0$, revealing that in a pay-for-performance setting, the more weight placed on profit, the higher quality of care chosen by the hospital. I combine the results found in each scenario into one response function that depends on α , where θ_{p4p} is quality under pay-for-performance incentives and θ_{ffs} is quality under fee-for-service only. That is,

$$\begin{aligned} \frac{d\Delta\theta}{d\alpha} &= \frac{d(\theta_{p4p} - \theta_{ffs})}{d\alpha} \\ &= \frac{d\theta_{p4p}}{d\alpha} - \frac{d\theta_{ffs}}{d\alpha} \\ &> 0. \end{aligned}$$

Hence, this model predicts that change in quality depends on how much weight the hospital places on profit vs. societal benefit. Particularly, hospitals with more weight on profit respond more than hospitals with more weight placed on societal benefit. In reality, α is unknown, except in the case of for-profits, who have revealed that they have a high α by their ownership status.

Clinically trained executives could also affect this α , as doctors commonly take a hippocratic oath to put patients first.

4.1 Estimation

The hypothesis driven from the theory described in Section 4 is that clinically trained executives put more weight on quality of care before the financial incentives on quality began. To investigate this hypothesis empirically, I merge all general for-profit hospitals in the AHA survey to my sample of nonprofit hospitals. Then, I compare readmission and mortality rates of always clinical executives and never clinical executives to for-profit hospitals after the incentive change. While I do not have information on leadership teams of for-profit hospitals, on average, they have revealed through their ownership status that they more heavily weigh profit. Thus, comparing different executive team behavior to for-profit behavior helps reveal how profit driven certain leadership teams are. Thus, I estimate the weighted synthetic difference-in-differences version of

$$y_{ht} = \beta (\text{for-profit x post 2012})_{ht} + \gamma_h + \delta_t + \epsilon_{ht} \quad (4)$$

where y_{ht} is one of the readmission or mortality rate outcomes discussed in Section 2.2, for-profit_{ht} is an indicator for having for-profit ownership status, and γ and δ are hospital and time fixed effects. In my estimation, nonprofits in the comparison group are more heavily weighted if their pre-trends are more similar to for-profits, and time periods are weighted to balance pre and post trends for for-profit hospitals. These weights follow the synthetic difference-in-differences approach, which is more robust to the parallel trends assumption.

I create two subsets of data to compare the response of for-profits to the response of the subsets of nonprofits based on executive team. First, I limit only to for-profits and nonprofits that always have clinically trained executives. Second, I limit to for-profits and nonprofits that never have clinically trained executives.⁶ I show summary statistics of relevant variables for for-profit hospitals in Appendix A.4. Again, I employ a synthetic difference-in-differences strategy to rely less heavily on a strong parallel trends assumption for different types of hospitals.

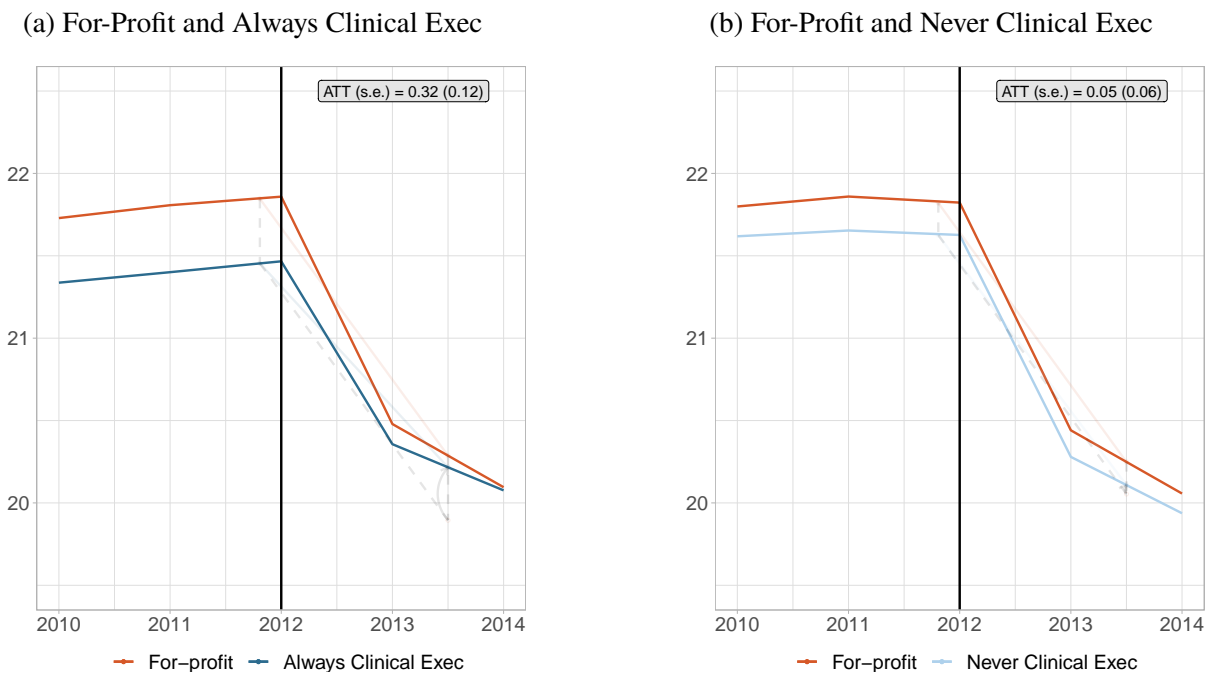
The identification assumptions I make here are largely similar to those in Section 3, only now incorporate for-profit hospitals. First, I assume that, absent the change in incentives, the weighted composition of for-profits and the relevant comparison group would have had parallel trends in outcomes. I also assume that there is no anticipation of the policy changes prior to 2012, which is

⁶I assume that the sample selection of stable leadership teams is not correlated with the policy changes. Please see Section 5 for the analysis and discussion of this assumption.

reasonable as the rules of the policies were announced in October of 2011. Finally, I assume no other changes correlated with both hospital type (for-profit, nonprofit with and without clinically trained executives) and quality occurred in conjunction with the 2012 pay-for-performance policies.

The estimates and graphical representation of the effect of being for-profit is shown in Figure 5. In panel 5a, the average treatment effect shows the effect of being for-profit compared to nonprofit with clinical executives on readmission rates after pay-for-performance changes. I find that for-profits decrease readmission rates by .3 (standard error .12) more than nonprofits with clinically trained executives. That is, for-profits respond more aggressively to the incentive change than nonprofits with clinical leadership. This difference is nearly identical to the difference between clinical and non-clinical teams shown in Section 3. So, unsurprisingly, panel 5b shows no such difference when comparing for-profits to nonprofits without clinically trained executives. Not only is the estimate not statistically significant, but the magnitude indicates a precise zero. That is, a lack of clinically trained executives leads nonprofits to act like for-profits in terms of readmission rates when responding to this incentive change.

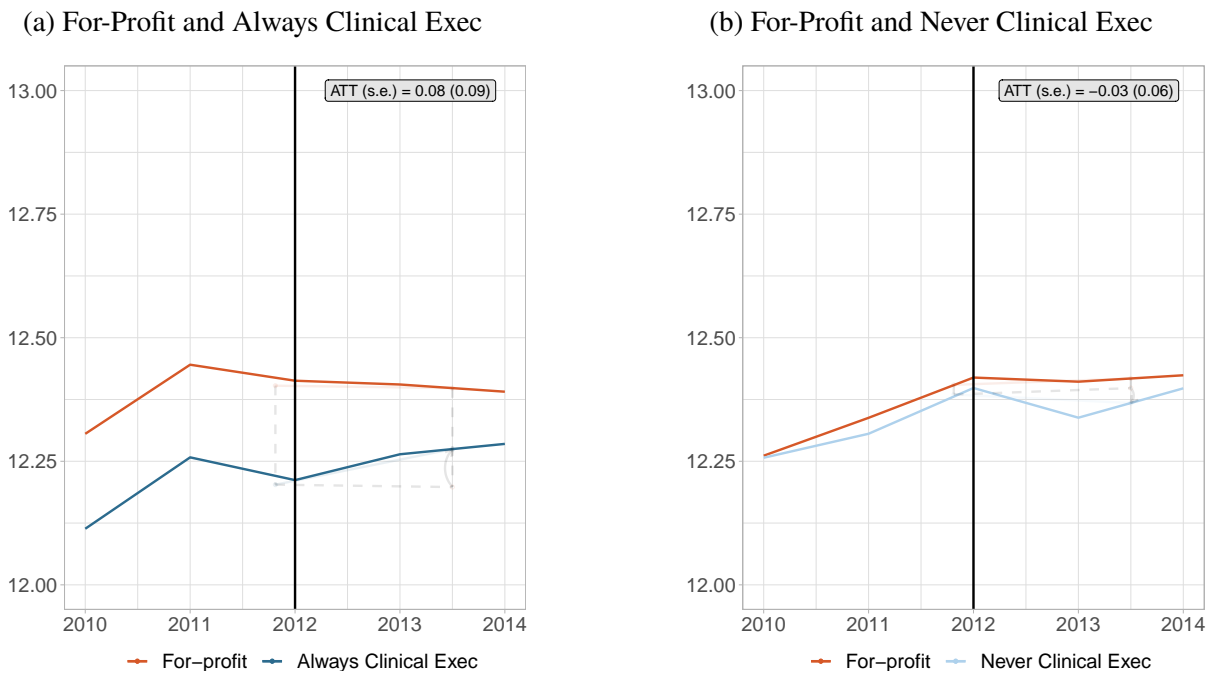
Figure 5: Comparison to For-Profit: Readmission



Next, I present estimated differences in mortality rates between different hospital types in Figure 6. In panel 6a, the estimated difference is between for-profits and nonprofits with clinically trained

executives. The magnitude of the difference in mortality rates is .08, where clinical team hospitals continue on an increasing trend and for-profits remain stable in mortality rate. While I cannot rule out a lack of differential response, the documented noisiness of mortality measures (Mackenzie et al. 2016) may mean that the true effect is different from zero. However, the estimated difference in mortality between for-profits and nonprofits without clinically trained executives is certainly zero with a magnitude of .03, shown in panel 6b.

Figure 6: Comparison to For-Profit: Mortality



5 Signaling vs. Managing Decomposition

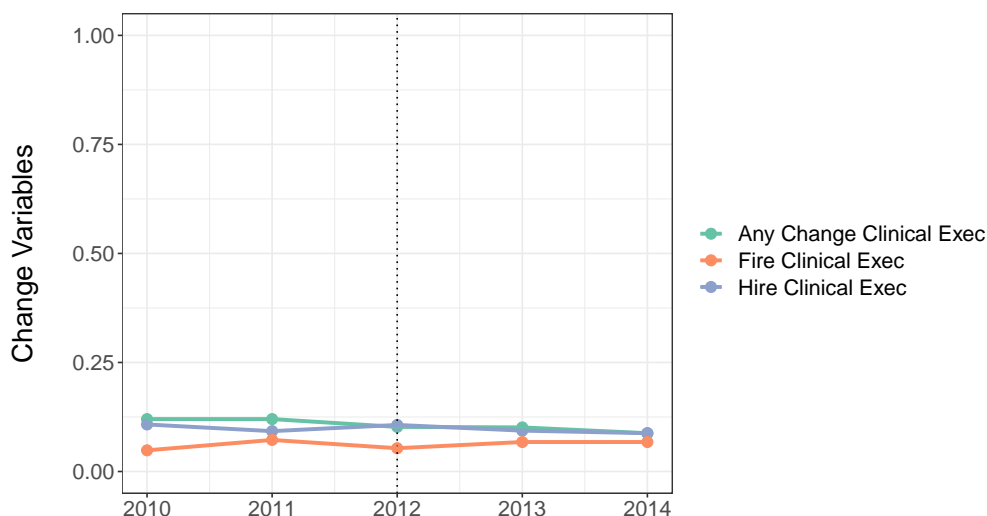
Given that having clinically trained executives does affect hospital response to financial incentives on quality, an important distinction is whether the existing clinical leaders *reveal* the underlying objectives of the hospital, or *manage* the hospital differently. Up to this point, I have limited the sample of nonprofit hospitals to those that don't change their propensity to hire a clinically trained executive over the sample period, which combines any signaling and managing effects. However, under the assumption that hiring or firing these types of executives is not correlated with the pay-for-performance policy changes, such changes can be leveraged to decompose the effect into signaling

versus managing. Since this decomposition depends heavily on the assumption that clinical executive hiring is not endogenous, I first investigate leadership changes in response to the policy change. I provide details of the decomposition estimation in Section 5.2

5.1 Analysis of Executive Team Changes

Decomposing the observed difference in response based on timing is dependent on executive team changes not being correlated with the financial incentive changes. First, I present raw means over time that represent the fraction of hospitals who hire a clinical executive, fire a clinical executive, or have any change in their propensity to have a clinical executive in Figure 7. There do not seem to be any drastic changes in 2011 or 2012 when the policies were becoming relevant to hospitals.

Figure 7: Leadership Change Means Over Time



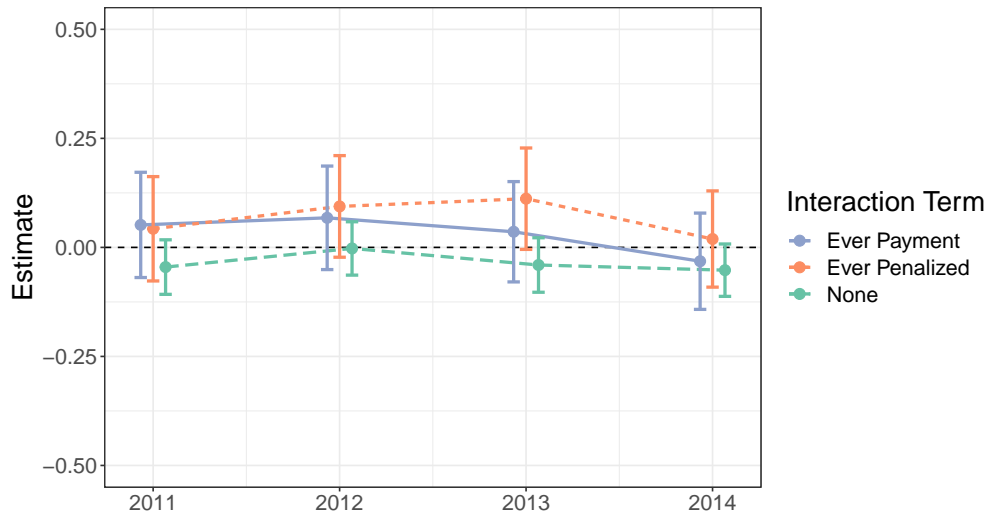
Next, I assess whether this assumption is reasonable by analyzing how executive team changes are occurring over time for certain groups of hospitals. I estimate the following two-way-fixed-effects specification:

$$\text{change}_{ht} = \sum_{j=2011}^{2014} \beta_j (\mathbf{1}\{t = j\} \times \text{Program Exposed})_{ht} + \alpha_h + \epsilon_{ht}, \quad (5)$$

where the variable Program Exposed_h has three definitions. First, an indicator for whether the hospital was ever penalized under HRRP. Second, an indicator for whether the hospital ever received payments under HVBP. And finally, I leave out the Program Exposed variable to simply see how

changes are correlated with given years for all hospitals. The outcome variable change_{ht} is an indicator for whether the hospital changes the number of clinical executives in a given year. The estimates from this analysis are presented in Figure 8. All estimates are not statistically different from zero, indicating that expectations of program exposure do not lead to changes in propensity to hire a clinical executive. Thus, it seems unlikely that endogenous team formation is biasing the estimates.

Figure 8: Leadership Analysis Results



5.2 Estimation

To disentangle clinical executives as signals of underlying hospital preferences or a different type of hospital manager, I carefully define the “treat” and comparison group of hospitals in the following estimating equation. For clarity, I also present the specification details in Figure 9. If clinically trained leaders are simply signals of underlying objectives, it shouldn’t matter when hospitals have hired a clinical executive. For example, take two hypothetical hospitals. Hospital A has a clinically trained executive from 2010-2011. Hospital B has a clinical executive from 2010-2013. If clinically trained executives are a signal of underlying objectives, these two hospitals should respond similarly to pay-for-performance objectives despite only one of them having a clinical executive in 2012. However, if clinically trained leaders manage the hospital in a unique way, these two hospitals would respond differently since only one of them has clinical leadership in 2012 when the policy changes occurred.

$$y_{ht} = \beta (\text{treat x post 2012})_{ht} + \gamma_h + \delta_t + \epsilon_{ht}, \quad (6)$$

Thus, in one specification I compare hospitals who ever have a clinically trained executive to hospitals who never have a clinically trained executive. This captures the signaling effect. In another specification, I remove all hospitals who never had a clinically trained executive, and compare hospitals who had a clinically trained executive in 2012 to those who had a clinically trained executive in any year except 2012, capturing a managing effect.

Figure 9: Decomposition Model Specification Details

Signaling:		Managing:	
Treat Group:	Never MD	Treat Group:	No MD in 2012
Comparison Group:	Ever MD	Comparison Group:	Ever MD (not in 2012)

As in the previous analyses, I estimate the synthetic difference-in-differences version of each of these specifications. That is, I weight control group units more heavily when their pre-trends are similar to the treated group units, and I weight more heavily time periods that balance pre and post period trends for the control group. I have discussed the assumptions of parallel trends, no anticipation, and no other confounding events in both Sections 3 and 4. These assumptions are also necessary here. I also assume that executive team changes are not endogenous with the policy changes, which I explore in detail in Section 5.1.

The estimates of each decomposition specification are shown in Table 3. The readmission rate estimates are in columns (1) and (2), and mortality rate estimates are in columns (3) and (4). The first row compares ever clinical executive hospitals to never clinical executive hospitals, and the second row compares hospitals with a clinical executive in 2012 to hospitals who have a clinical executive in a year other than 2012. While there is no statistically significant estimate, the coefficients of the managing effect are very similar to the full effect found in Section 3. The coefficients for the signaling effect are both statistically insignificant and close to zero in magnitude. Thus, I can rule out that clinical executives are simply a signal of underlying hospital objectives.

6 Heterogeneity by Physician Specialty

While most doctors take a hippocratic oath to provide care with integrity and compassion, doctors with training in certain specialties may be more equipped to improve care for certain patients.

Table 3: Decomposition of Readmissions and Mortality Synthetic DiD Results

	Weighted Avg. Readmission Rate		Weighted Avg. Mortality Rate	
	(1)	(2)	(3)	(4)
Signaling Effect	-0.01 (0.07)		0.03 (0.06)	
Managing Effect		0.17 (0.1)		0.12 (0.09)
Observations	3,005	1,595	3,005	1,595

Results from Equation with readmission and mortality rates as outcome variables.

Standard errors are clustered at the hospital level.

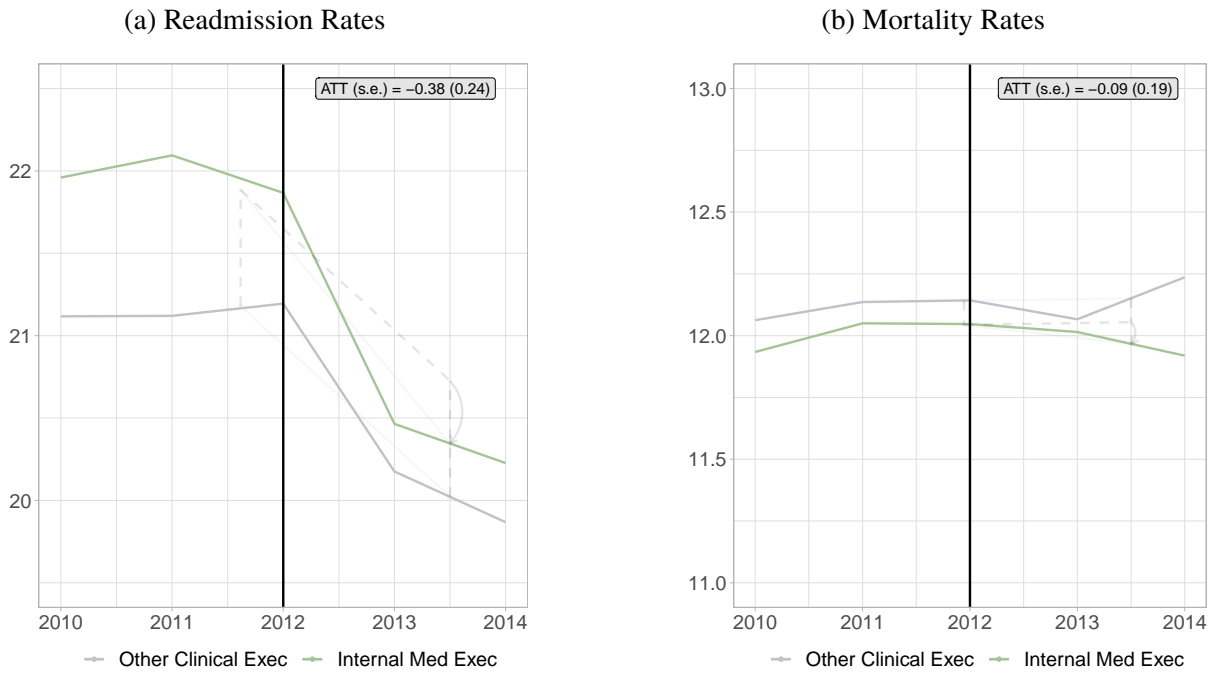
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Since the pay-for-performance incentive programs directly target three conditions, physicians with knowledge of these conditions may be especially suited to respond to the policies. Specifically, a pulmonologist typically treats pneumonia, and a cardiologist typically treats heart failure and heart attack patients, which are the three conditions considered in the HRRP penalties. Both of these specialties are under the umbrella of internal medicine, which is the most granular specialty type in my data.

I limit the sample of hospitals to those who always have some type of clinical executive over the sample period. Then, I estimate the difference in readmission and mortality rates after pay-for-performance among hospitals who do and do not have an internal medicine clinical executive. I lose a significant fraction of the sample with this limitation: there are 60 hospitals that consistently have an internal medicine executive or only a different specialty executive. The summary statistics for these samples are in Appendix A.4. I present the results from this comparison in Figure 10.

While I am limited in power due to the small sample size of hospitals in this analysis, I provide evidence that hospitals with internal medicine executives may change readmission rates differentially compared to teams with clinical executives of other specialty types. The estimated effect of having an internal medicine executive is -0.38 (standard error $.24$), where internal medicine executives decrease readmissions at a faster rate than other specialty types. The results for mortality rates are similar (still not statistically significant), but smaller in magnitude and smaller relative to the main results. This is probably due to the specialty-specific conditions being more relevant for the readmission rate penalties than the quality rewards.

Figure 10: Effect of Internal Medicine Training on Quality



7 Conclusion

While previous literature has documented correlations between executive characteristics and firm behaviors/performance, research is still limited on this topic in the health care sector. Health care managers are tasked with maintaining a business while also prioritizing patient care, which is why existing findings on manager effects may not extend to this setting. This paper investigates whether having executive leaders with clinical training impacts hospital behaviors and outcomes, particularly in response to a policy change that financially incentivizes quality. I utilize a unique data set of nonprofit US hospitals that captures the proportion of clinical executives, specialties of clinical executives, hospital characteristics, and hospital quality metric outcomes.

In 2012, two pay-for-performance policies went into effect that either penalized hospitals with high readmission rates, or gave payments to hospitals with low readmission/mortality rates. These policies were relevant to hospital finances, and gave tangible incentive for hospitals to increase quality. Thus, I leverage this exogenous policy change and estimate how clinically trained executives affect readmission and mortality rates in response to the change using a synthetic difference-in-differences strategy. I find that hospitals without clinically trained executives improve quality more than hospitals without clinically trained executives after the incentive change.

I investigate several hypotheses that arise given the estimated difference in response. First, this result is not driven by non-clinical teams adjusting their patient complexity. However, comparing non-clinical teams to for-profit hospitals reveals that it could be underlying preferences for profit vs. patient benefit that drive the difference. Finally, I leverage exogenous changes in executive teams to show that clinically trained executives are not simply a signal of underlying preferences, but manage the hospital differently.

In US health care, a very common policy goal is to improve quality of care without increasing costs. However, up to this point, it was unclear whether hospital leaders might play a role in attaining this. In this paper, I show that having hospital managers with clinical training does affect the way hospitals respond to incentives. Importantly, my results point to clinically trained executives already prioritizing quality before incentives were put in place. My findings have implications for policy makers considering financial incentives on quality, and even point to the potential of incentivizing the presence of clinically trained hospital leaders.

References

- Bertrand, Marianne and Antoinette Schoar (2003). “Managing with style: The effect of managers on firm policies”. In: *The Quarterly journal of economics* 118.4, pp. 1169–1208.
- Adams, Renée B and Daniel Ferreira (2009). “Women in the boardroom and their impact on governance and performance”. In: *Journal of financial economics* 94.2, pp. 291–309.
- Azoulay, Pierre, Joshua S Graff Zivin, and Jialan Wang (2010). “Superstar extinction”. In: *The Quarterly Journal of Economics* 125.2, pp. 549–589.
- Brickley, James A, R Lawrence Van Horn, and Gerard J Wedig (2010). “Board composition and nonprofit conduct: Evidence from hospitals”. In: *Journal of Economic Behavior & Organization* 76.2, pp. 196–208.
- Levi, Maurice, Kai Li, and Feng Zhang (2010). “Deal or no deal: Hormones and the mergers and acquisitions game”. In: *Management Science* 56.9, pp. 1462–1483.
- Dranove, David (2011). “Health care markets, regulators, and certifiers”. In: *Handbook of health economics*. Vol. 2. Elsevier, pp. 639–690.
- Matsa, David A and Amalia R Miller (2011). “Chipping away at the glass ceiling: Gender spillovers in corporate leadership”. In: *American Economic Review* 101.3, pp. 635–639.
- Ahern, Kenneth R and Amy K Dittmar (2012). “The changing of the boards: The impact on firm valuation of mandated female board representation”. In: *The quarterly journal of economics* 127.1, pp. 137–197.
- Kaplan, Steven N, Mark M Klebanov, and Morten Sorensen (2012). “Which CEO characteristics and abilities matter?” In: *The journal of finance* 67.3, pp. 973–1007.
- Custódio, Cláudia, Miguel A Ferreira, and Pedro Matos (2013). “Generalists versus specialists: Lifetime work experience and chief executive officer pay”. In: *Journal of Financial Economics* 108.2, pp. 471–492.
- Matsa, David A and Amalia R Miller (2013). “A female style in corporate leadership? Evidence from quotas”. In: *American Economic Journal: Applied Economics* 5.3, pp. 136–169.
- Benmelech, Efraim and Carola Frydman (2015). “Military ceos”. In: *Journal of financial Economics* 117.1, pp. 43–59.
- Office, US Government Accountability (2015). *Hospital Value-Based Purchasing: Initial Results Show Modest Effects on Medicare Payments and No Apparent Change in Quality-of-Care Trends (Vol. GAO-16-9)*.
- Mackenzie, Simon J et al. (2016). “Measuring hospital-wide mortality—Pitfalls and potential”. In: *The Journal for Healthcare Quality (JHQ)* 38.3, pp. 187–194.

- Mellor, Jennifer, Michael Daly, and Molly Smith (2017). “Does it pay to penalize hospitals for excess readmissions? Intended and unintended consequences of Medicare’s Hospital Readmissions Reductions Program”. In: *Health economics* 26.8, pp. 1037–1051.
- Adams, Renée B, Ali C Akyol, and Patrick Verwijmeren (2018). “Director skill sets”. In: *Journal of Financial Economics* 130.3, pp. 641–662.
- Harvard Business Review, Article (Oct. 2018). URL: <https://hbr.org/2018/10/why-doctors-need-leadership-training>.
- Norton, Edward C et al. (2018). “Moneyball in medicare”. In: *Journal of health economics* 61, pp. 259–273.
- Ziedan, Engy (2018). “Essays on How Hospitals Responded to Pay for Performance Incentives”. PhD thesis. University of Illinois at Chicago.
- Flabbi, Luca et al. (2019). “Do female executives make a difference? The impact of female leadership on gender gaps and firm performance”. In: *The Economic Journal* 129.622, pp. 2390–2423.
- Friedson, Andrew I, William C Horrace, and Allison F Marier (2019). “So Many Hospitals, So Little Information: How Hospital Value-Based Purchasing Is a Game of Chance”. In: *Southern Economic Journal* 86.2, pp. 773–799.
- Frydman, Carola (2019). “Rising through the ranks: The evolution of the market for corporate executives, 1936–2003”. In: *Management Science* 65.11, pp. 4951–4979.
- Janke, Katharina, Carol Propper, and Raffaella Sadun (2019). *The impact of ceos in the public sector: Evidence from the english nhs*. Tech. rep. National Bureau of Economic Research.
- Ody, Christopher et al. (2019). “Decreases in readmissions credited to Medicare’s program to reduce hospital readmissions have been overstated”. In: *Health Affairs* 38.1, pp. 36–43.
- Arkhangelsky, Dmitry et al. (2021). “Synthetic difference-in-differences”. In: *American Economic Review* 111.12, pp. 4088–4118.
- Gupta, Atul (2021). “Impacts of performance pay for hospitals: The readmissions reduction program”. In: *American Economic Review* 111.4, pp. 1241–1283.
- Ahmed, Shakeel (July 2022). *Council post: Why doctors make the best healthcare CEOs*. URL: <https://www.forbes.com/sites/forbesbusinesscouncil/2022/07/27/why-doctors-make-the-best-healthcare-ceos/?sh=267bebeb3b6b>.
- Bradley, Steven W et al. (2022). “The impact of chief diversity officers on diverse faculty hiring”. In: *Southern Economic Journal* 89.1, pp. 3–36.
- Graham, John R et al. (2022). “Corporate culture: Evidence from the field”. In: *Journal of financial economics* 146.2, pp. 552–593.

- Munoz, Pablo and Cristóbal Otero (2022). “Managers and public hospital performance”. In: *Working Paper*.
- ASPE (2023). *Ownership of Hospitals: An analysis of newly-released federal ...* URL: <https://aspe.hhs.gov/sites/default/files/documents/582de65f285646af741e14f82b6df1f6/hospital-ownership-data-brief.pdf>.
- CMS, Agency (2023a). *Hospital readmissions reduction program (HRRP)*. URL: <https://www.cms.gov/medicare/payment/prospective-payment-systems/acute-inpatient-pps/hospital-readmissions-reduction-program-hrrp#:~:text=The%20Hospital%20Readmissions%20Reduction%20Program,in%20turn%2C%20reduce%20avoidable%20readmissions..>
- (2023b). *Hospital value-based purchasing program*. URL: <https://www.cms.gov/medicare/quality/initiatives/hospital-quality-initiative/hospital-value-based-purchasing>.
- Stajduhar, Tony (Nov. 2023). *Should doctors run healthcare organizations?* URL: <https://www.newsweek.com/should-doctors-run-healthcare-organizations-1847200#:~:text=Physicians%20make%20great%20leaders%20because,lead%20to%20better%20patient%20outcomes..>
- Kunz, Johannes S et al. (2024). “Assessing the quality of public services: For-profits, chains, and concentration in the hospital market”. In: *Health Economics*.
- Medicare Medicaid Services, Centers for (2024). *National Health Expenditure Data: Historical*. Accessed: 2024-08-05. URL: <https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expendituredata/historical#:~:text=Pages%20in%20this%20section&text=U.S.%20health%20care%20spending%20grew,spending%20accounted%20for%2017.3%20percent..>

A Data

A.1 Gathering Hospital Leadership Names

I use the API of ProPublica’s Nonprofit Explorer to access the archive of nonprofit Tax Form 990s.⁷ I extract the Employee Identification Number (EIN) for all nonprofits categorized with a National Taxonomy of Exempt Entities (NTEE) code of E20 (hospital), E21 (community health system), and E22 (general hospital). There are 5,588 EINs total in this list.

I query more detailed information from the API on each hospital EIN. I save the name, secondary name, state, and zip code, all of which do not vary by year. I also record URL links to the Tax Form 990 PDFs, which are unique in each year. For the sake of a comprehensive data set, I keep years 2006-2020 (I later limit to 2010-2014 when focusing on pay-for-performance initiatives in 2012). Thus, I finish this step with a panel data set of EIN characteristics and PDF locations. Importantly, there are multiple types of Tax Form 990s depending on the size of the nonprofit. In many cases, one nonprofit has at least two different forms filed in a given year. I filter out any EIN-years for which there are no PDFs available. There are numerous foundations or auxiliary firms with the purpose of raising funds for the hospital, but do not provide services to patients. I filter out any not-for-profit with “foundation” or “auxiliary” in the name. I also filter out various specialty centers such as hospice or cancer centers.

For my analysis, I need to link these nonprofits with other sources of data to recover penalties from HRRP, payments from HVBP, bed size, and outcomes of interest. I match to the American Hospital Association (AHA) Survey, which contains hospital characteristics and Medicare ID number, and will easily merge to other data sets used for hospital information. An EIN to AHA ID crosswalk does not exist. Therefore, I take a conservative approach to matching EINs to AHA ID based on hospital name and location.

In the AHA data, I keep general acute care hospitals in the contiguous US, Alaska, and Hawaii (excluding places like Puerto Rico) that are classified as nonprofit or state/community. I remove hospitals who weren’t present in the data or change system ID in 2009-2015, meaning they either closed or were acquired. Due to the survey nature of this data, a hospital name may look slightly different from one year to the next. For example, “Waldo County General Hospital” is also “Waldo County General Hospital Maine Health”. Further, zip codes often change by one or two digits, making them unreliable to match based on. To deal with this, I first keep only unique AHA ID, name, zip, state, and system name combinations. Then, I convert the data from long to wide so

⁷At the time of writing this, information on using version 2 of the API can be found at <https://projects.propublica.org/not-for-profits/api>.

that each AHA ID occurs only once, but may have multiple names, zip codes, or system names associated with it.

I proceed matching based on names in several steps. I focus on exact string matches, so I remove all spaces and common characters that could cause mismatches such as &, ', -, and "inc". I loop through each AHA ID, limit to nonprofits in the same state, and extract any exact name matches. When an exact match is found, I record the link between AHA ID and EIN. In this first layer of matching, 860 hospitals in the AHA data are linked to an EIN, equivalent to 31% of nonprofit hospitals in the sample. In the next layer of matching, I remove common words such as "healthcare", "regional", "hospital", etc. That way if there are subtle differences in names, removing common words may allow for an exact match. Again, I take each AHA hospital name and look for exact matches in the nonprofits in the same state. This adds an additional 90 hospital matches, accounting for a total of 34.5% of AHA hospitals. Finally, I manually search through unmatched EINs to identify any matches. From google searches, I identify an additional 300 EIN-AHA ID matches, yielding a total of 1200 EINs.

While it would be ideal to analyze all nonprofit hospitals, a sample size of 852 firms is relatively large in the hospital executive literature. I present a comparison of in-sample and out-of-sample nonprofits. I present means and standard deviations of these measures in Table ???. The in-sample nonprofits are slightly larger in terms of beds and patients seen. They are also slightly more likely to be penalized or receive payments under the pay-for-performance incentives. However, average readmission and mortality rates are very similar, as well as case mix index. The largest difference in the two samples is that I under-represent nonprofit hospitals in systems. This is due to the nature of the tax form 990s, where systems often file one tax form for the entire system, making it difficult to discern the true managers of a specific hospital. Because of this, I drop hospitals from the sample who only have leadership information from a system tax form.

I then extract the names of board members and executives from the Tax Form 990 PDFs of matched hospitals. In the data set of hospital PDF URLs that I collected earlier, I limit to the hospitals with solid matches described above. I then loop through each EIN, downloading PDFs locally and using the tesseract package in R to extract text from the relevant pages of the PDF using OCR text extraction methods. In particular, I loop through each page of the PDF, look for the title associated with leadership names: "Officers, Directors, Trustees, Key Employees, and Highest Compensated Employees", and save all the text from any pages where this title is found. I save the text to a list of all EIN, years present.

One aspect of the NonProfit Explorer API is that, only in some cases, if two forms are present for an EIN, year, only the first one (which is typically not the one with the relevant information) is

Table 4: NFP Sample Comparison

Variable	In-sample NFP		Out-of-sample NFP	
	Mean	Std. Dev.	Mean	Std. Dev.
Hospital Characteristics				
Academic Med. Center	0.31	0.46	0.25	0.43
Number Beds	169.09	203.57	143.11	186.70
Physician Owned	0.01	0.12	0.01	0.07
System Affiliated	0.50	0.50	0.63	0.48
Executive Team Characteristics				
Number Executives	3.92	2.75		
Fraction Clinical Execs	0.09	0.17		
Fraction Int. Medicine Execs	0.03	0.09		
Has a CMO	0.33	0.47		
Has Clinical CEO	0.12	0.33		
Penalty/Payment Variables				
Ever Received HVBP Incentive	0.70	0.46	0.61	0.49
Penalized for AMI	0.04	0.19	0.03	0.17
Penalized for HF	0.05	0.21	0.06	0.24
Penalized for Pneumonia	0.07	0.25	0.04	0.20
Penalized for AMI + HF	0.05	0.22	0.03	0.18
Penalized for AMI + Pneumonia	0.02	0.14	0.01	0.11
Penalized for HF + Pneumonia	0.37	0.48	0.33	0.47
Penalized for All Conditions	0.26	0.44	0.20	0.40
Num. Hospitals	843		2,222	

pulled. Therefore, for some hospitals, a couple years will have gaps in text extraction data. I locate EIN, years where this problem is occurring, and a team of RAs locates and downloads the correct forms manually. I extract text from these manually downloaded forms in the same manner as above.

The form of the extracted text data is a data frame with one column, where each line of text is saved in a different row. Typically on the same page as the names and positions is a list of the highest compensated employees and their compensation. In order to not record extra names, I filter out any rows after the start of this section. I then remove any digits, parentheses and brackets, other punctuation, letters that occur by themselves, two letter “words” that have no meaning, and excess space between words. I then split up the phrase into individual words, so one phrase with 5 words is broken up into 5 variables. I write a text cleaning function that locates names, positions, titles, and indications of resigning. I flag name rows using the Census name list data for the year 2000.

The columns with the most flags are then identified as name columns. I then extract any text that indicates a doctor title and link it to the name located the closest to it. Similarly, I extract text of all potential positions such as CEO, CFO, CMO, president, board member, etc., and link it to the name most closely located to it.

I then create a name-level data set that only includes executives. That is, I remove all board members from the data. I now seek to match people in this data set to existing physicians in the National Plan and Provider Enumeration System (NPPES) data, where all physician National Provider Identifiers (NPI) are recorded. For each executive (not just those identified as doctors), I record the number of name matches found in the NPPES data. There are several steps to identifying physicians with a true match. First, if an executive self-identifies as a doctor and has a unique match in NPPES, I record the unique NPI number for that executive. Second, if an executive self-identifies as a doctor and has multiple name matches in NPPES, I manually search online for information on the executive, and find the unique NPI most likely associated with the executive. Finally, for any executives who do not self-identify as a physician, but have name matches in the NPPES data, I manually research them and record the unique NPI number if they are a physician.

The data is now an executive-year level data set with information on clinical training and specialty when relevant. I then create hospital level indicators based on the names, titles and positions associated with the hospital: the number of clinically trained executives, the number of executives with particular specialties, the number of total executives, and whether the hospital employs a CMO.

A.2 Executive Team Changes

For the main analysis, I limit to hospitals who either always have a clinically trained executive or never have a clinically trained executive. In a supplemental analyses, I leverage changes to hospital propensity to hire a clinically trained executive. I explore the exogeneity of such changes in Section 5.1. Additionally, I present the number of hospitals who change their propensity to hire a clinically trained executive at different times in Table ???. I break the timing up into pre-2012, 2012, and post-2012. I show the number of hospitals who have clinically trained executives in different time combinations. The number of hospitals that only have an MD executive in 2012, but at no other time, is only 6. The majority of hospitals (187) who ever have a clinical executive have one in every time period.

Table 5: Timing of MD Exec Changes

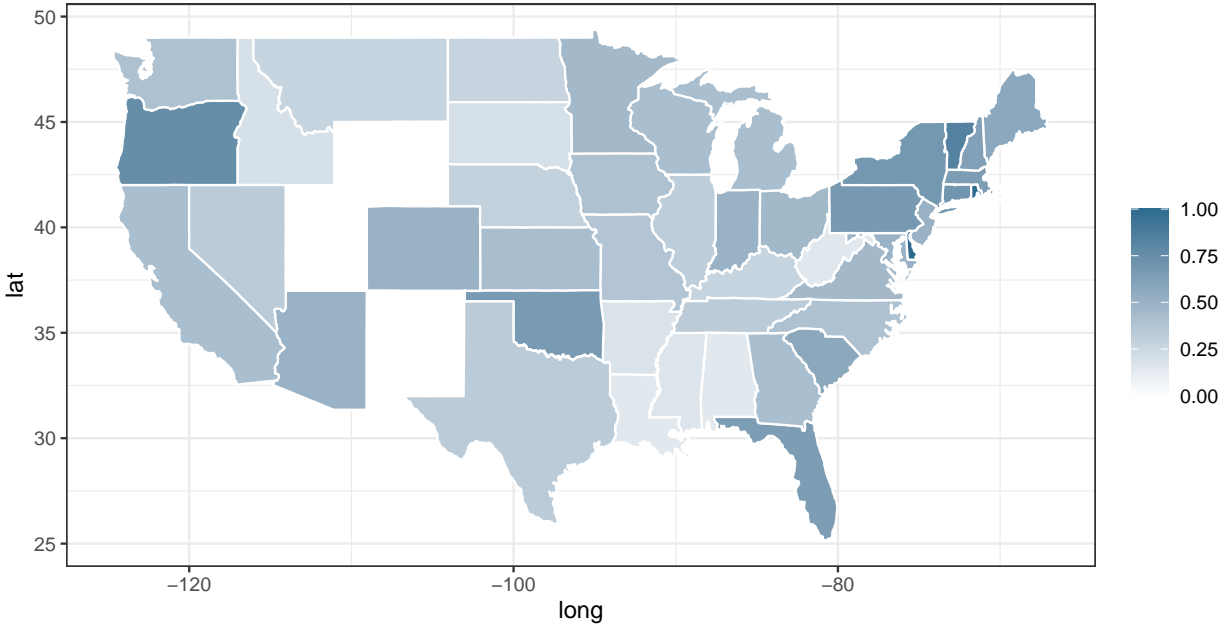
Has MD Pre-2012	Has MD 2012	Has MD Post 2012	Num. Hospitals
0	1	1	35
1	1	0	24
0	1	0	8
1	1	1	197
0	0	1	60
1	0	0	43
0	0	0	452
1	0	1	24

A.3 Merging to Other Hospital Data

Using the hospital Medicare ID number found in the AHA survey, I merge the executive information to various publicly available hospital data sets. For information on penalties or payments given by the pay-for-performance policies, I use the Hospital Cost Report Information System (HCRIS) data. For 30-day readmission rates, 30-day mortality rates, and the number of patients in the relevant conditions, I merge in the CMS Hospital Compare data. Finally, to investigate the role of selective patient practices I use a case mix index variable, which comes from the CMS Case Mix Index files.

A.4 Additional Summary Statistics

Figure 11: Percent of Hospitals with Clinical Executives by State



Notes: This graph shows the percent of hospitals with a clinically trained executive in each state.

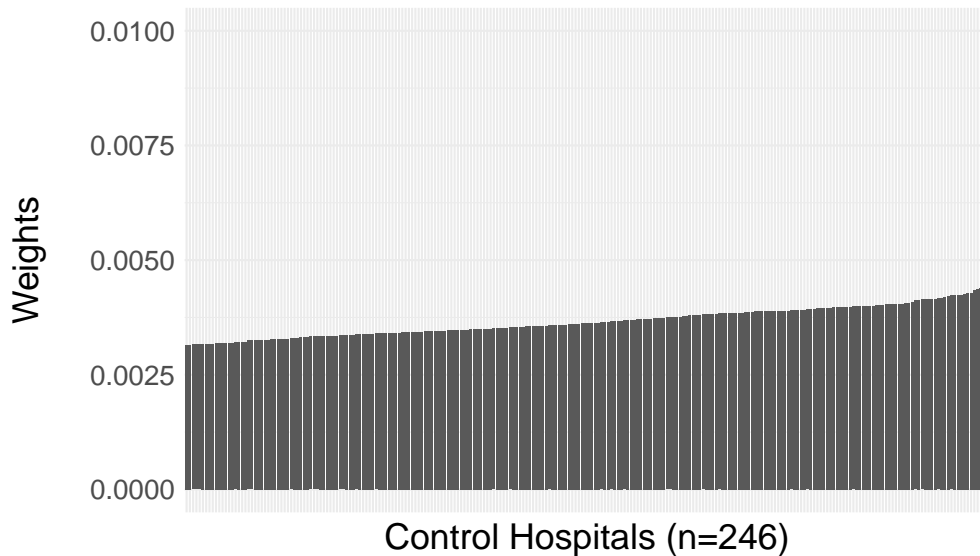
Table 6: Summary Statistics by Hospital Type

Variable	Ever Clinical Exec	Always Clinical Exec	Never Clinical Exec	For-Profit
Hospital Characteristics				
Academic Med. Center	0.41	0.52	0.24	0.17
Number Beds	231	222	116	157
Physician Owned	0.01	0.02	0.02	0.19
System Affiliated	0.58	0.52	0.44	0.89
Executive Team				
Number Executives	4.85	5.25	3.11	
Fraction Clinical Execs	0.20	0.26	0.00	
Fraction Int. Medicine Execs	0.06	0.10	0.00	
Has a CMO	0.53	0.45	0.15	
Has Clinical CEO	0.25	0.11	0.00	
Penalty/Payment Variables				
Ever HVBP Incentive	0.79	0.75	0.62	0.87
Penalized for AMI	0.05	0.04	0.03	0.03
Penalized for HF	0.05	0.07	0.05	0.06
Penalized for Pneum.	0.08	0.09	0.05	0.06
Penalized for AMI + HF	0.06	0.05	0.04	0.06
Penalized for AMI + Pneum.	0.03	0.04	0.02	0.01
Penalized for HF + Pneum.	0.41	0.39	0.34	0.55
Penalized for All Conditions	0.30	0.27	0.22	0.36
Num. Hospitals	391	56	445	680

Table 7: Summary Statistics by Specialty

Variable	Always Int Med Exec	Always Other Exec
Hospital Characteristics		
Academic Med. Center	0.61	0.49
Number Beds	343.84	181.61
Physician Owned	0.00	0.02
System Affiliated	0.57	0.58
Executive Team Characteristics		
Number Executives	5.89	5.13
Fraction Clinical Execs	0.27	0.27
Fraction Int. Medicine Execs	0.24	0.00
Has a CMO	0.60	0.39
Has Clinical CEO	0.20	0.10
Penalty/Payment Variables		
Ever Received HVBP Incentive	0.83	0.75
Penalized for AMI	0.07	0.06
Penalized for HF	0.07	0.06
Penalized for Pneumonia	0.03	0.12
Penalized for AMI + HF	0.07	0.06
Penalized for AMI + Pneumonia	0.07	0.02
Penalized for HF + Pneumonia	0.57	0.35
Penalized for All Conditions	0.43	0.20
Num. Hospitals	30	35

Figure 12: Control Group Weights, Readmission and Baseline Case Mix



B HRRP and HVBP Programs

In October 2011, the Center for Medicare and Medicaid Services (CMS) released a set of rules under HRRP mandating penalties for hospitals with above average readmission rates. The goal of HRRP is to lower readmissions through better care coordination, less initial stay complications, and better post-care instructions. Beginning in October 2012, hospitals with higher readmission rates than the national average in pneumonia, heart failure, or AMI (after adjusting for demographic characteristics) receive a fixed lower reimbursement rate for all Medicare patients seen in their hospital. In 2015, CMS also included chronic obstructive pulmonary disease, coronary artery bypass graft surgery, and elective primary total hip arthroplasty and/or total knee arthroplasty as conditions which go into the penalty calculation (CMS 2023a).

Penalties are given in the form of a fixed rate reduction of 1-3% for every Medicare patient regardless of the condition. Further, CMS does not distinguish a necessary readmission from an avoidable readmission; any repeat hospital visit is included in the penalty calculation. Excess readmission rates are calculated using a rolling look-back period of 3 years to determine whether the hospital is penalized. Therefore, hospitals had incentive to react immediately once details of the program were announced in October of 2011.

The HVBP Program instead rewards hospitals with high quality or significant improvement in quality. Specifically, CMS deducts Medicare payments by 2% from all eligible hospitals, collects this sum, and divides it among the rewarded hospitals. Several quality and cost measures surround-

Figure 13: Time Weights, Readmission and Baseline Case Mix

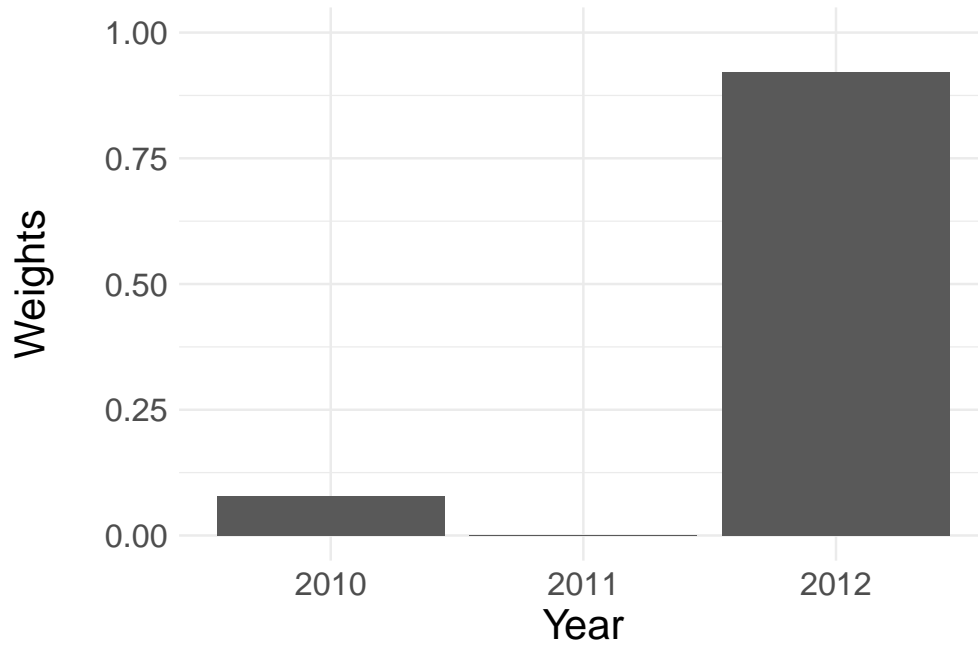


Figure 14: Control Group Weights, Mortality and Baseline Case Mix

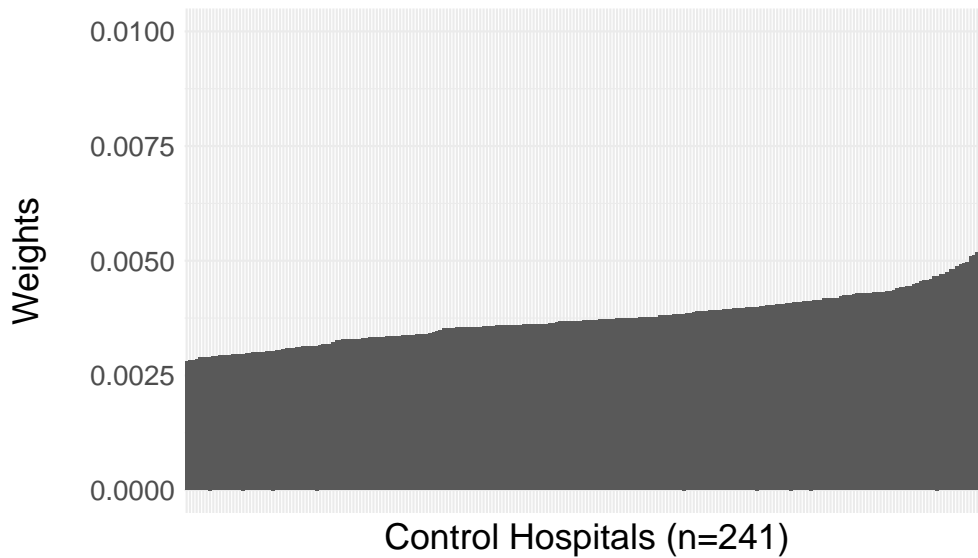
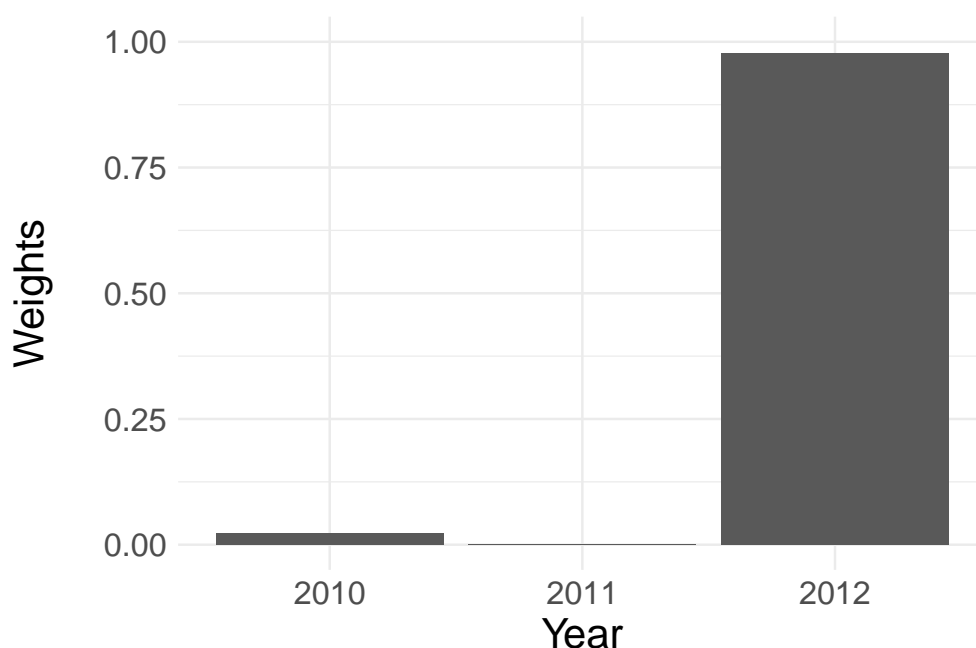


Figure 15: Time Weights, Mortality and Baseline Case Mix



ing safety, efficiency, cost reductions, clinical outcomes, and community engagement are combined to create a single score metric for each hospital. Hospitals are then compared to the average and are rewarded for being above average quality or for showing improvement (CMS 2023b).

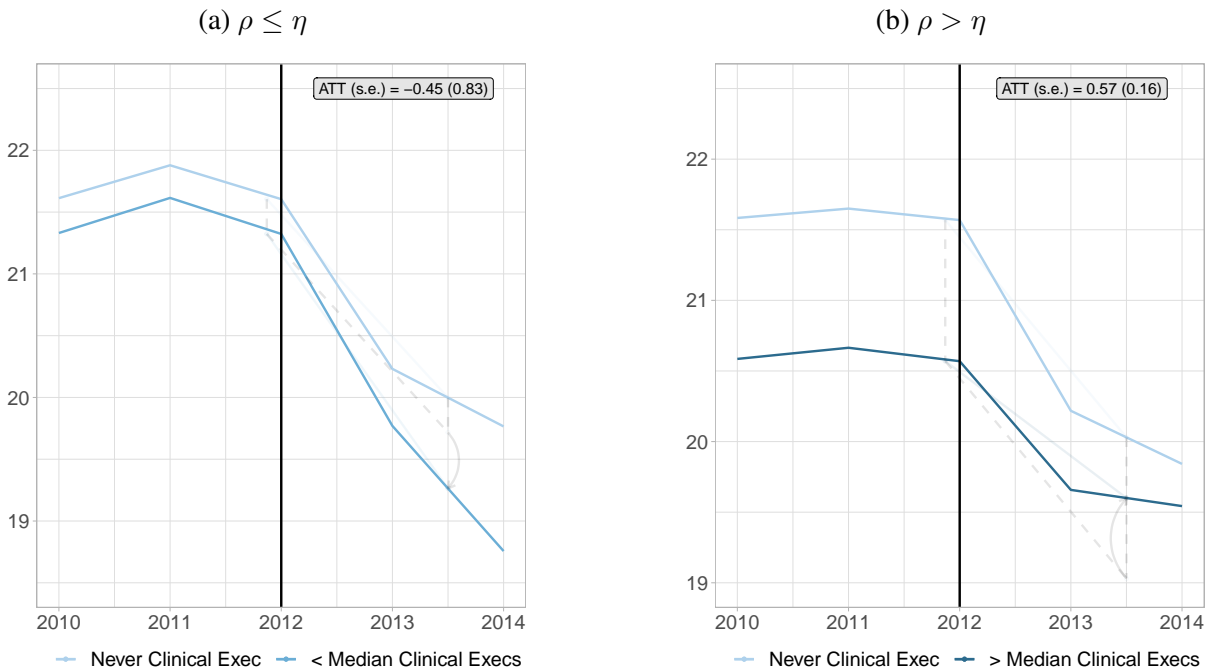
C Supplemental Analyses

C.1 Binned Treatment

In Section 3, I present intensive margin results by binning hospitals according to the fraction of their executive team with clinical experience. There, when comparing hospitals with greater than the median (20%) to less than the median fraction of clinical executives, I find that the difference in response is driven by hospitals with a fraction greater than the median. One concern could be that this result is driven by a difference in the total size of hospital executive teams, the denominator of that fraction. In this section, I use the same estimating equation, but I limit to only nonprofit hospitals with 7 or less executives in total, removing outliers of executive team size. The results are shown in Figure 16, and show that the response is still driven by hospitals with a fraction of clinical executives above the median.

I also present results with mortality as the dependent variable, shown in Figure 17. These

Figure 16: Effect of Clinical Training on Readmission Rates, Binned Treatment



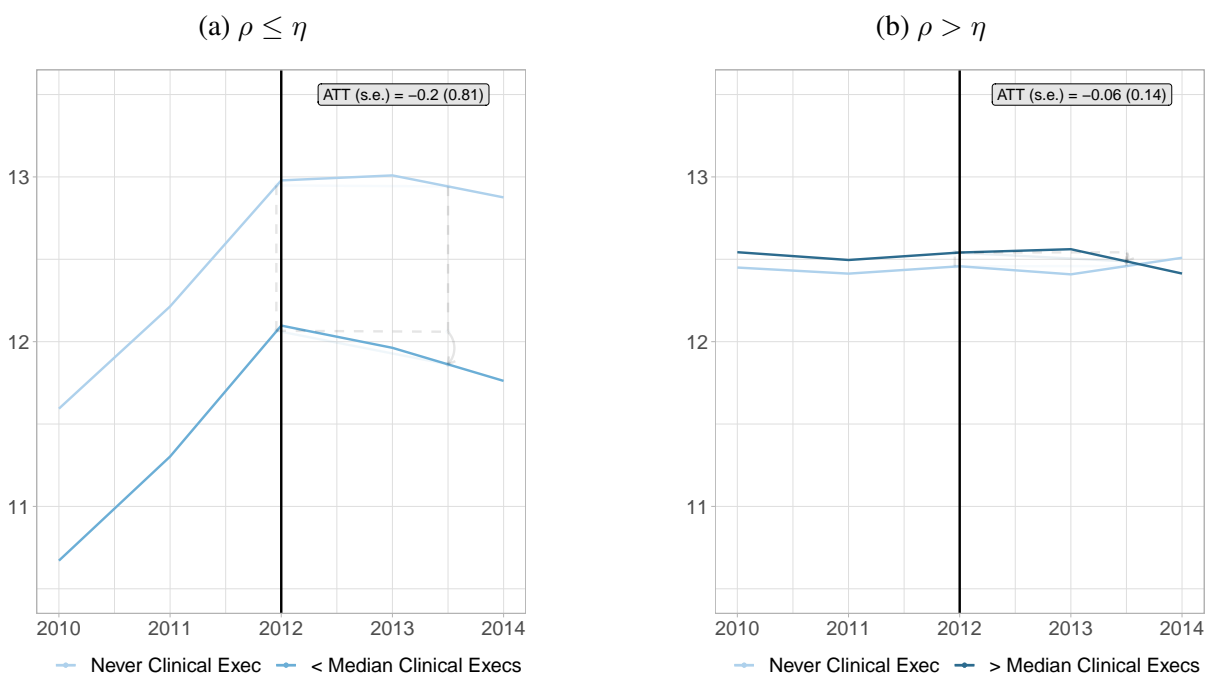
results do not differ from the main finding in that there is no statistical difference in response for hospitals with different types of leadership teams with above or below the median clinically trained executives.

C.2 Selective Patient Practices

A method hospitals have been shown to employ in order to decrease readmission rates is being more selective with which patients to admit to the hospital (Gupta 2021), as only admitting healthy patients mechanically decreases readmission rates. One hypothesis behind a more drastic quality decrease by non-clinical teams is that they may be more willing to participate in such practices, while executives with clinical training are less likely to turn patients away.

To test this, I employ case mix index as the outcome in specification 1 to determine whether one type of hospital differentially changes patient composition after pay-for-performance incentives, where the same identification assumptions hold. I present the estimated difference in Figure 18, comparing hospitals that never have a clinical executive to hospitals that always have a clinical executive. Neither type of hospital drastically changes case mix index after the incentive change, and there is no estimated differential response in complexity of patients seen in the hospital. Thus,

Figure 17: Effect of Clinical Training on Mortality Rates, Binned Treatment



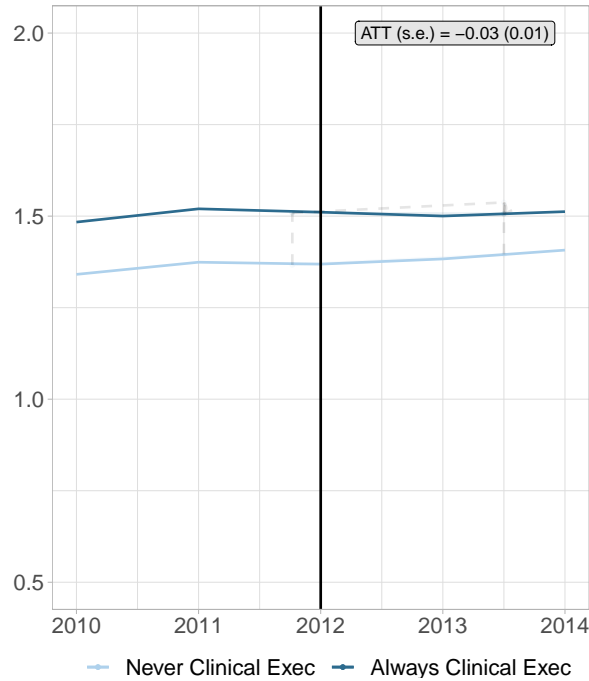
the estimated difference in readmission rate response between clinical and non-clinical executives is not due to differentially changing case mix.

C.3 Results By Condition

In the main body of the paper, the main dependent variables I consider are weighted averages of readmission and mortality rates for several conditions relevant to the policy changes, weighted by the number of patients seen with each condition. Here, I investigate whether the findings are driven by any particular condition. The conditions I consider are heart attack (AMI), heart failure, and pneumonia, which are the conditions that go into the calculation for HRRP penalties. Thus, I consider the effect of clinically trained executives on readmission and mortality rates for each condition individually. I estimate equation 1 with these outcome variables.

I present the resulting estimates and standard errors in Table 8. Columns (1)-(3) show results for readmission rate outcomes in each condition. The main finding, that clinically trained executives lead hospitals to have higher readmission rates after pay-for-performance policies, is primarily driven by patients with pneumonia. While magnitudes are positive for all conditions, pneumonia has the largest magnitude and is the only statistically significant estimate. Columns (4)-(6) present

Figure 18: Effect of Clinical Executive on Case Mix Index



results for mortality rates of each condition. Similarly to the results for weighted average mortality, there are no statistically significant results here. However, pneumonia has the largest positive magnitude, just like pneumonia readmissions.

Table 8: Effect of Clinically Trained Executive on Condition-Specific Readmission and Mortality Rates

	Readmission Rates			Mortality Rates		
	AMI	HF	Pneum.	AMI	HF	Pneum.
	(1)	(2)	(3)	(4)	(5)	(6)
Always Clinical Exec x Post Program	0.25 (0.18)	0.14 (0.18)	0.46** (0.14)	-0.13 (0.19)	0.03 (0.13)	0.3 (0.19)
Hospital FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	1,135	1,640	1,635	1,300	1,640	1,635

Standard errors are clustered at the hospital level.

Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1

C.4 Robustness to Other Estimations

In analyses presented in the main body of the paper, I use a synthetic difference-in-differences approach to identify the effect of clinically trained executives. I choose this as the main estimation strategy due to potential unobserved differences correlated with executive choice, where synthetic DiD assures parallel trends by choosing the optimal composition of control and treatment hospitals. As a robustness check, I now present results under three different estimation techniques for robustness: classic two way fixed effects, two-way fixed effects with manual weights that balance case mix index and penalties of hospitals with different leadership teams, and synthetic control.

For the manual weighting strategy, I apply the Coarsened Exact Matching (CEM) procedure as described by Azoulay, Graff Zivin, and Wang (2010). This approach involves grouping hospitals based on observable characteristics such as bed size, the number of patients in each relevant condition, and eventual HRRP penalty status in each relevant condition. Hospitals are then placed into bins according to these characteristics, and matches are made exactly between hospitals within the same bin. Hospitals that do not find an exact match in this process are excluded from the sample.

C.4.1 Effect of Clinically Trained Executives on Response to Financial Incentive on Quality

In Section 3, I first establish the effect of clinically trained executives on response to change in incentives on quality using a synthetic difference-in-differences approach. Now, I present results using various estimation techniques in Table 9 in order to assess to robustness of this finding. Columns (1)-(3) present estimates and standard errors for the three estimations: TWFE, TWFE with matched sample, and synthetic control, respectively. For both TWFE and TWFE with matching, the magnitude is larger than that for synthetic DiD (0.27) and statistically significant. The estimate and standard error for synthetic control are very similar to those for synthetic DiD. In columns (4)-(6), I present the results for mortality rates, which confirm no differential response. This exercise supports the finding that clinically trained executives affect the readmission rate response of hospitals to change in incentives on quality.

Additionally, I estimate an event study version of the two way fixed effects specifications. The estimates and 95% confidence intervals are shown in Figure 19. These confirm the main findings with slightly noisier estimates, and show that the difference in response in readmission rates occurs in 2013 and 2014, one and two years after the policy changes. The mortality results confirm the main findings, that there is no difference in response in mortality rates.

Table 9: Effect of Clinically Trained Executive, Various Estimations

	Readmission Rate			Mortality Rate		
	TWFE	Matched	SC	TWFE	Matched	SC
	(1)	(2)	(3)	(4)	(5)	(6)
Ever Clinical Exec x Post Program	0.42** (0.15)	0.65* (0.3)	0.25 (0.14)	0.11 (0.15)	0.03 (0.4)	0.06 (0.12)
Hospital FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	2,387	1,025	1,640	2,381	1,023	1,640

Standard errors are clustered at the hospital level.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

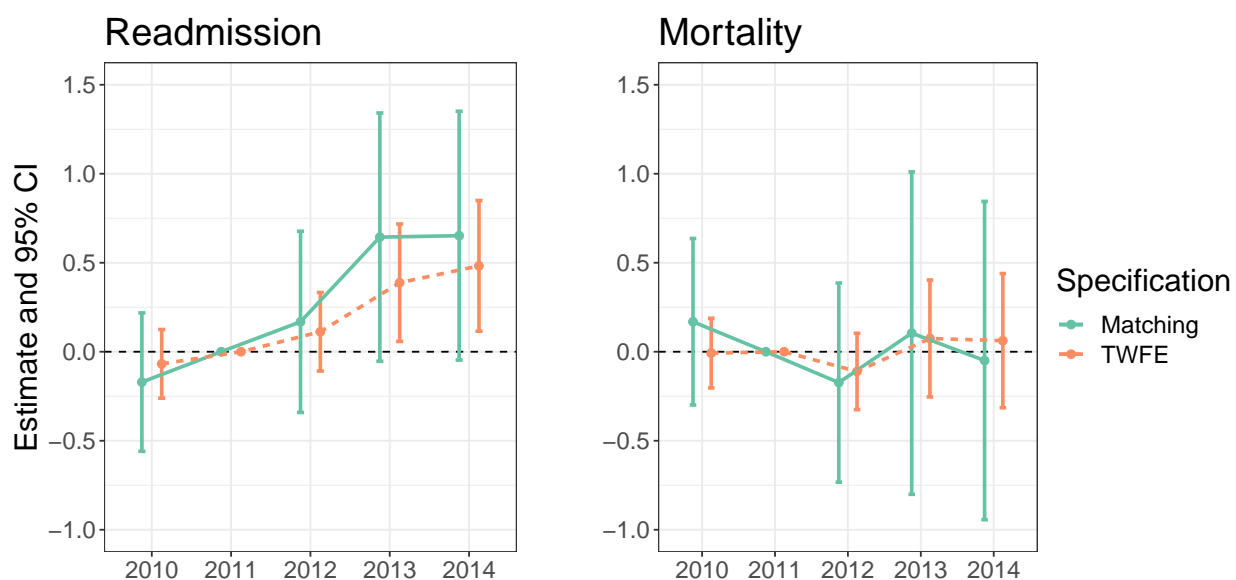


Figure 19: Effect of Clinical Executives; Event Study Results

C.4.2 Are Clinical Executives Less Profit Driven?

In Section 4, I test empirically whether nonprofits with different leadership teams respond differently to the incentive change than for-profit hospitals using a synthetic difference-in-differences methodology. Here, I present results from various other estimation strategies. In Table 10, I present estimates and standard errors with weighted average readmission rate as the outcome. These results confirm what is shown in the main body of the paper, that nonprofits without clinically trained

executives act similarly to for-profits, but nonprofits with clinically trained executives do not.

Table 10: Compare For-Profit to Types of Leadership (Readmission Rates), Various Estimations

	Compare to Always Clinical Exec			Compare to Never Clinical Exec		
	TWFE	Matched	SC	TWFE	Matched	SC
	(1)	(2)	(3)	(4)	(5)	(6)
For-Profit x Post Program	-0.37* (0.15)	-0.28 (0.44)	-0.43* (0.19)	-0.03 (0.07)	-0.08 (0.12)	0.09 (0.1)
Hospital FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	2,660	660	2,660	3,820	1,790	3,820

Standard errors are clustered at the hospital level.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

While the results in the main body of the paper for mortality rates as the dependent variable do not indicate differences in response, I still present the estimates from other estimations in Table 11. These results do not contradict what I show in the main body of the paper, as I still see small and statistically insignificant estimates.

Table 11: Compare For-Profit to Types of Leadership (Mortality Rates), Various Estimations

	Compare to Always Clinical Exec			Compare to Never Clinical Exec		
	TWFE	Matched	SC	TWFE	Matched	SC
	(1)	(2)	(3)	(4)	(5)	(6)
For-Profit x Post Program	0.03 (0.14)	0.09 (0.45)	-0.16 (0.2)	0.06 (0.07)	0.05 (0.11)	0.03 (0.1)
Hospital FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	2,660	660	2,660	3,820	1,790	3,820

Standard errors are clustered at the hospital level.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

C.4.3 Signaling vs. Managing Decomposition

In Section 5, I test empirically whether clinical executives make a difference because of a signaling or managing effect. That is, I leverage the timing of hiring clinically trained executives. Here, I present results from the same specification, but various estimation strategies. In Table 12, I present estimates and standard errors with weighted average readmission rate as the outcome. These results confirm what is shown in the main body of the paper, that the managing effect is driving the findings.

Table 12: Signaling vs. Managing Decomp. (Readmission Rates), Various Estimations

	Signaling Effect			Managing Effect		
	TWFE	Matched	SC	TWFE	Matched	SC
	(1)	(2)	(3)	(4)	(5)	(6)
Estimate	0.03 (0.08)	0.18 (0.16)	-0.04 (0.08)	0.25* (0.11)	0.26 (0.22)	0.18 (0.11)
Hospital FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	3,005	670	3,005	1,595	320	2,040

Standard errors are clustered at the hospital level.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

While the results in the main body of the paper for mortality rates as the dependent variable do not indicate differences in response, I still present the estimates from other estimation strategies in Table 13. These results do not contradict what I show in the main body of the paper, as I still see small and statistically insignificant estimates. Therefore, there is no opposing signaling vs. managing effect.

Table 13: Signaling vs. Managing Decomp. (mortmission Rates), Various Estimations

	Signaling Effect			Managing Effect		
	TWFE	Matched	SC	TWFE	Matched	SC
	(1)	(2)	(3)	(4)	(5)	(6)
Estimate	-0.01 (0.08)	-0.23 (0.18)	0.02 (0.06)	0.14 (0.11)	-0.05 (0.28)	0.11 (0.08)
Hospital FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	3,005	670	3,005	1,595	320	2,040

Standard errors are clustered at the hospital level.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1