NORTHWESTERN UNIVERSITY

Essays on the Hospital Industry

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Economics

By

Mariano Irace

EVANSTON, ILLINOIS

June 2018

Abstract

This dissertation studies the forces that drive the allocation of patients to hospitals in the United States. Even though it seems there is some market discipline in the hospital industry, we observe that many patients go to hospitals far from the quality frontier and we still do not have a full understanding of why that is the case. The chapters of this thesis study different forces that play a role in the allocative process in this industry.

Chapters 1 and 2 study persistence in hospital choices of patients in the State of New York. Specifically, I analyze the causal impact of the previous hospital choice of a patient on her current choice. Although the conventional wisdom is that patients are loyal to hospitals, there is actually little hard evidence supporting this idea. While inertia in consumer choice has been documented in many settings, there are very few studies on inertia in choice of medical provider. The first two chapters of this dissertation try to fill this gap by exploring the extent and determinants of patient loyalty toward hospitals.

Chapter 1 studies persistence in hospital choices of patients across different medical conditions. To disentangle state dependence from persistent unobserved heterogeneity, I exploit choice set variation across episodes due to emergency hospitalizations and temporary hospital closures due to Hurricane Sandy. I find that patients who are forced to switch hospitals by an exogenous shock are more likely to continue using the new hospital in the future than similar patients. After showing that there is state dependence in my setting, I study whether it stands in the way of the reallocation of patients to high quality providers in the context of hospital choice for heart surgery.

Chapter 2 is a complement to Chapter 1. It analyzes persistence in hospital choices of patients for childbirth. For identification, I follow a more traditional strategy than in Chapter 1 that relies on the existence of consumers who are new to the market (first-time mothers) and experienced consumers (mothers who had births before). I show that on average a mother is more likely to choose the same hospital used in the previous birth for her current birth than an observationally similar first-time

mother. The results are in line with the findings of Chapter 1, suggesting that patient loyalty is a general phenomenon.

Finally, Chapter 3 analyzes the impact of provider financial incentives on hospital behavior. It focuses on a major regulatory reform to the Long-Term Care Hospital industry that substantially reduced payments for about half the Medicare patient population treated in this setting. I study the strategies that Long-Term Care Hospitals have implemented to mitigate the impact of the new regulatory framework and the effects of these actions on access to care and treatment intensity. I find evidence that hospitals screen patients who are referred for long-term acute care more actively than before the reform and are more likely to reject those patients for which reimbursement decreased. In addition, many hospitals altered their discharge patterns in a way that is hard to explain given the way that financial incentives changed with the reform.

Acknowledgments

First of all, I would like to express my sincere gratitude to my advisors for their patience, motivation, and invaluable advice.

I am particularly indebted to Igal Hendel, who has been a wonderful mentor. His door was always open and his constant support was crucial during my graduate studies. Without his guidance, this work would not have been possible. I wish I could express in words my gratitude for all his help.

I am very grateful to Robert Porter for his generous advice and support. It was the second year graduate course on Industrial Organization that he teaches at Northwestern that led me to do research in this area.

I thank Leemore Dafny for providing invaluable insights during all the stages of this work. Even after she left Northwestern, she always had time to answer my questions and discuss the details of my work.

I would also like to thank Gaston Illanes, who offered his support at all times. He always was willing to look at the details of my work and his insights were always clarifying.

Special thanks go to Frank Limbrock, who guided me through the early process of my doctoral studies and provided critical suggestions that have made this work significantly better. This work has also benefited from conversations with Alex Torgovitsky, Martin Gaynor, Amanda Starc, Christopher Ody, Matthew Notowidigdo, and Elena Prager. I thank participants in the Industrial Organization Student Seminar Group and the Seminar in Industrial Organization at Northwestern University for helpful comments and suggestions.

I gratefully acknowledge support from the Becker Friedman Institute's Health Economics Initiative. I also want to thank Bernard Black, who helped me to get access to the data required for the different parts of this study.

I am grateful to my friends Matias Escudero and Gabriela Cugat, who helped me numerous times during my stay at Northwestern by discussing my work and attending informal presentations.

I would like to thank my professors at Universidad Nacional de Cordoba and at Centro de Estudios Monetarios y Financieros (CEMFI). Particular thanks go to Angel Enrique Neder, who encouraged me to pursue graduate studies in Economics from the very beginning of my academic career.

I thank all my family for their support and encouragement over the past years.

Finally, I thank my wife, Guadalupe, for being always by my side throughout this entire process.

Contents

List of Figures

List of Tables

Chapter 1

Persistence in hospital choices of patients across medical conditions

1.1. Introduction

In this chapter, I study persistence in hospital choices of patients. The conventional wisdom is that patients patronize one hospital and use it for all of their medical care needs. This idea of patient loyalty is consistent with views of industry analysts about the business practices of hospitals. For example, many view maternity services as loss leaders: hospitals offer these services not because maternity patients are profitable *per se*, but because they expect mothers and their families to go back to the hospital for more profitable services in the future. Another example concerns "data blocking" activities. It has been argued (Miller and Tucker, 2009; Miller and Tucker, 2014; HITPC, 2015; Desai, 2016) that hospitals hinder the sharing of patient data with other providers for competitive reasons: patients might find it easier to leave a hospital and seek treatment elsewhere once their clinical data can follow them across providers.

As pointed out by Heckman (1981), the empirical observation that a patient repeatedly uses the same hospital can be explained by either unobserved heterogeneity or state dependence. In the first case, the patient has strong and persistent latent preferences for the hospital. In the second case, previous choices have a causal impact on the current decision. The stories in the previous paragraph rest implicitly on the idea that hospital choices of patients are "sticky", but their implications depend on whether preference heterogeneity or state dependence drives the stickiness in patients' behavior. If the observed choice persistence is due to unobserved heterogeneity, hospitals cannot control the evolution of patients' preferences. In this case, there are no dynamic incentives to invest in unprofitable service lines (e.g. maternity services) in the hope of developing long-standing relationships with patients. Under state dependence, on the other hand, hospitals' investments in loss leader services influence future demand. Similarly, if the continued use of a hospital reflects patients' latent preferences for that provider, then data blocking activities do not have any impact on future demand.

Although the idea of patient loyalty towards hospitals is consistent with the views of people in the industry, there is little supporting evidence. Few studies document persistence in the hospital choices of patients or analyze the determinants of persistence. In particular, whether persistence results from stable preferences or state dependence remains an open question. In this chapter, I fill this gap by investigating empirically the determinants and implications of persistence in hospital choices of patients in the state of New York.

The study of persistence in hospital choices of patients is interesting for several reasons. Previous studies have documented the existence of state dependence in consumers' choices in a variety of industries. I provide additional insights about the sources of patients' preferences by analyzing whether patients display similar purchase patterns. Characterizing patient behavior is important to improve our understanding of the forces that drive the allocation of patients to hospitals and to inform the design of policies that influence patient demand. Although there is some market discipline in the hospital industry (in the sense that better performing hospitals have higher and increasing market shares), many patients choose hospitals that are far from the quality frontier (Chandra et al., 2016). I investigate whether state dependence plays a role in this process: does it stand in the way of the reallocation of patients to high quality providers? If so, then, for example, the benefits from better sorting of patients across facilities should be considered when evaluating policies aimed at achieving interoperability of hospitals' electronic health record (EHR) systems. Second, previous use of a hospital is a strong predictor of the current hospital choice of a patient. However, whether this preference is due to state dependence or unobserved heterogeneity matters for welfare analysis and policy evaluation. State dependence implies a less durable preference for a facility that has been used in the past: if the patient needs to switch hospitals, then she has to pay a one time cost, while unobserved heterogeneity implies that the utility loss from going to a less preferred alternative is permanent (Raval and Rosenbaum, 2017; Shepard, 2016). The distinction can affect policy conclusions. For example, if state dependence drives the persistence in patient behavior, the long-run welfare effect of excluding a hospital from an insurer's network will be lower than if persistence reflects unobserved heterogeneity.

In the first part of the chapter, the primary concern is to distinguish between persistence in choice due to state dependence and persistence in choice due to unobserved preference heterogeneity. It is difficult for researchers to separate the sources of persistence in a credible way. Most previous studies have relied, at least in part, on functional form assumptions about preference heterogeneity for identification. The concern is that parametric assumptions might lead to overestimates of the magnitude of state dependence. This is a natural concern in my context, where I expect unobserved heterogeneity to be empirically relevant, given that there are multiple attributes of patients and hospitals that I cannot control for (such as religious affiliation and location of the workplace of the patient, and hospital amenities). Given this, the burden of proof is to show that there is state dependence. I use an event-study approach that relies on transparent assumptions about the data generating process (dgp). In particular, I exploit quasi-exogenous shocks that force patients to visit a hospital that they would otherwise not have chosen: emergency hospitalizations and temporary hospital closures due to a natural disaster.

In the first case, I find that patients who visit a new hospital during an emergency hospitalization are more likely to continue using that same facility in subsequent episodes than observationally similar patients. Patients who visit the same hospital they had been using before for the emergency hospitalization exhibit a higher repurchase rate, which indicates that unobserved heterogeneity is also empirically relevant. The observed patterns are similar across different types of emergencies, hospitals, and patients.

In the second case, I exploit the unexpected closures of three hospitals in New York City following Hurricane Sandy. I show that patients who needed hospital care during the time an affected hospital was closed for repairs were less likely to use the facility after its reopening than similar patients who did not have hospital visits during the unavailability window. Moreover, patients continue using the hospital they visited during the time their usual hospital was unavailable. I present this evidence in a transparent and simple way, without relying on complex estimators. Next, I cast this setting into the nonparametric framework of Torgovitsky (2016). In this model, each consumer has a type given by a vector of dynamic potential choices. The potential choices in a given period indicate the alternatives the consumer would choose under exogenous manipulations of the previous period choice. If I knew the distribution of types in the population, I could determine the proportion of consumers who exhibit state dependence. However, I do not know this distribution. Therefore, I consider all the distributions that could have generated the data. I restrict the set of admissible distributions to those that are consistent with the distribution of observables and satisfy an exclusion restriction (which takes the form of an independence restriction between preferences and the timing of hospital visits). Then, I search for the admissible distributions for which state dependence is lowest and highest. I am able to bound the proportion of patients who exhibit state dependence away from zero without imposing parametric assumptions about the nature of preference heterogeneity.

Given that there is state dependence in my setting, I next analyze its implications. The focus of most studies of switching costs has been to determine their impact on pricing, by analyzing whether the investment motive (reduce prices to expand the customer base) or the harvesting motive (raise prices to exploit locked-in consumers) dominates (Farrell and Klemperer, 2007). Given the institutional features of the US hospital industry, I expect dynamic pricing to be a less relevant issue in my context[1]. Instead, I focus on the allocative role of state dependence. Lock-in hinders the ability of patients to react to changes in the environment. In particular, lock-in might discourage a patient from seeking treatment in hospitals that are more suitable than her previous choice for treating her current medical condition: any quality differential must compensate for switching costs. Absent state dependence, more patients may obtain treatment at high-quality hospitals.

Hospital quality may be disease-specific. Moreover, quality measures are difficult to compare across different medical conditions. I therefore study the allocative role of state dependence in the context of hospital choice for a specific procedure: heart surgery (Coronary artery bypass grafting, CABG). Ideally, I would exploit the same quasi-exogenous shocks used in the first part of the

[1] Dynamic considerations could affect quality choices of hospitals. For example, hospitals could invest in the quality of certain services to attract patients, and then exploit locked-in patients with low quality for other services.

chapter to conduct this analysis. However, small sample sizes prevent me from doing so. I therefore rely on observational data to recover the parameters of interest, following a more traditional but less transparent identification strategy.

I compare the allocation of patients to hospitals in the actual state of the world and in a counterfactual scenario where there is no state dependence. This counterfactual scenario can be interpreted as arising from a combination of policies that target the different sources of state dependence. The allocative role of state dependence is large: previous use of a hospital increases the probability that the patient chooses that facility for CABG more than three times above the baseline. Absent state dependence, patients would switch to higher quality hospitals: in this scenario, ex-ante expected mortality would be 3% lower than the observed mortality rate.

State dependence might arise from a variety of sources. First, patients face monetary and/or time costs of transferring medical records between providers. These switching costs result in part from the data blocking activities mentioned above. Improvements in the interoperability of EHRs and the diffusion of health information exchanges (HIE) have the potential to reduce these costs. More generally, the existence of switching costs is associated with relationship-specific investments that cannot be transferred seamlessly across providers.

Second, state dependence might originate from search and evaluation costs. Choosing a hospital is a complex activity. Patients need to collect information about different alternatives. Hospitals are complicated objects to evaluate, so cognitive limitations might be substantial. Search costs may be such that inertia is the efficient way to deal with moderate or temporary changes in the environment (Stigler and Becker, 1977). Then, state dependence might capture the use of heuristics by patients for choosing a hospital. The presence of search costs also suggests that a patient might not consider all possible alternatives in a choice occasion. Hospitals used in the recent past are more likely to be included in the patient's consideration set (Samuelson and Zeckhauser, 1988; Andrews and Srinivasan, 1995). The asymmetric position in the decision-making process of recently used hospitals means that they are more likely to be chosen.

Third, the presence of state dependence can be explained by learning costs. Uncertainty about the quality of hospitals leads risk averse patients to remain with familiar hospitals. Then, state dependence arises from the premium that patients are willing to pay for greater familiarity with a facility.

It is likely that state dependence arises from a combination of the factors mentioned above. For example, in a two-stage decision process, previous use of a hospital might have an impact on both the consideration and the evaluation stages: hospitals used in the recent past are more likely to enter the consideration set, and patients pay a cost conditional on switching hospitals due to the transfer of medical records between providers.

My analysis focuses on the overall impact of state dependence on patients' hospital choices, without distinguishing between the potential underlying mechanisms. However, disentangling the various sources of state dependence is important in order to craft policies to overcome inertia. Moreover, the welfare implications of eliminating state dependence will be different if it results from a tangible cost, as opposed to something that only affects choices. Data limitations prevent me from decomposing the sources of state dependence in a credible way in the current setting, but this is an interesting avenue for future work.

Previous studies have documented the presence of state dependence in consumers' choices of a variety of products (orange juice and margarine (Dubé et al., 2010), internet portals (Goldfarb, 2006a), pension funds (Luco, 2016; Illanes, 2016), health insurance plans (Nosal, 2012; Handel, 2013; Ericson, 2014; Polyakova, 2016; Ho et al., 2017), among others). There is a more limited number of studies of persistence in patients' choices of medical providers. The papers most closely related to my study are Jung et al. (2011), Shepard (2016), and Raval and Rosenbaum (2017). Jung et al. (2011) study the factors that affect hospital choices of employees at a large self-insured company. They use stated preference data from a survey: employees were asked to indicate the hospitals they would be most likely to consider if they needed to be hospitalized for a surgical procedure. The authors find that prior use of a hospital and patient satisfaction with a facility from prior experiences have a large effect on future hospital choices; the effect of prior use is smaller in cases where the previous admission occurred through the emergency department. While their analysis is based on hypothetical future choices, I study actual sequences of choices; moreover, I study what drives the observed persistence in choices. Shepard (2016) provides evidence of adverse selection against health insurance plans covering prestigious and costly hospitals. These plans attract consumers with strong preferences for these type of providers, particularly consumers who have used them in the past. These consumers are likely to choose these hospitals for all of their medical care needs, driving up costs for the insurer, which leads to exclusion of the facilities from the network. In this setting, previous use of a provider is useful for identifying patients with strong preferences for the hospital: whether patient loyalty is due to state dependence or to durable preference heterogeneity is irrelevant, and the empirical analysis does not attempt to separate them. Shepard (2016) emphasizes the effect of choice persistence on medical costs, while I consider how persistence might prevent a patient from switching to a hospital that is better at treating her current medical condition. Raval and Rosenbaum (2017) analyze patients' hospital choices for childbirth in Florida. They use a panel data fixed effects estimator to separate persistence in choice due to switching costs and persistence in choice due to unobserved preference heterogeneity. They consider women who have three children and switch hospitals between their first and second births. For identification, they compare the hospital choices (for the third birth) of women who attended the same two hospitals for the first two births but in different order. They find that approximately 40% of choice persistence reflects switching costs. The current work differs from their study in two dimensions. First, I do not restrict attention to a particular medical condition, but study persistence in hospital choices of patients more generally (in particular, patients might seek treatment for different medical conditions over time); persistence in hospital choices of mothers for childbirth is analyzed in Chapter 2. Second, I use a different identification strategy to separate the sources of choice persistence.

The remainder of the chapter proceeds as follows. Section 1.2 describes the data used in the empirical analysis. Section 1.3 discusses the empirical challenges present in my setting. Sections 1.4 and 1.5 provide evidence of state dependence in hospital choices of patients by exploiting shocks that shift the loyalty state of patients: emergencies and temporary hospital closures, respectively. Section 1.6 quantifies the impact of state dependence on health outcomes in the context of hospital choice for cardiac surgery. Section 1.7 provides concluding remarks.

1.2. Data

For the empirical analysis, I use detailed patient level data on the universe of visits to hospitals in the state of New York for inpatient and outpatient care. The dataset was obtained from the Statewide Planning and Research Cooperative System (SPARCS). It includes data on inpatient discharges (IP) (1995-2015), ambulatory surgery visits (AS) (1995-2015), and emergency department visits (ED) (2003-2015). In this section, I provide a brief overview of the data. Details about the specific samples used for the different applications are discussed in the corresponding sections.

Each record in the dataset is a hospital visit. The data includes an encrypted patient identifier that allows me to track patients' visits over time, hospital and physician identifiers, patient demographics (age, gender, race, ethnicity, zip code of residence), admission and discharge dates, type and source of admission, discharge status, diagnosis and treatment information, primary payer, charges, length of stay, and indicators for mortality within 7, 15, 30, 180, and 365 days of the discharge date[2]. The patient identifier is missing in 1.7% of records; I exclude these observations in the analysis.

The dataset is therefore a panel that follows the hospital choices of patients in New York. Given that data on hospital visits is more complete for the period from 2005 to 2015 and that information about hospitals characteristics before 2005 is scarcer, I analyze hospital choices of patients in the period from 2005 to 2015. However, I use all the available data to create individual histories of hospital visits.

[2]A complete list of variables can be found here: `https://www.health.ny.gov/statistics/sparcs/datadic.htm`

For the empirical analysis, I want different hospital visits of a patient to correspond to different episodes of care. I refer to the initial hospital visit of a patient to treat a certain condition as the index event, and I treat readmissions or visits for follow-up care as part of the same episode rather than as different episodes. In the latter case, persistence would be inflated by counting a visit for follow-up care to the same hospital where the patient originally received treatment as a repurchase. The general criterion I follow is to aggregate visits by a patient to the same hospital within a short period of time for related medical conditions into a single episode of care. If a visit is erroneously categorized as a readmission using this criterion, I expect the choice situation to be similar to the (erroneously) associated index visit and I do not want persistence to be driven by these cases. If anything, I prefer to err on the side of understating, rather than overstating, the extent of persistence. I tried different time windows (30, 60, 90, and 120 days) to identify readmissions; the qualitative results are robust to the use of different specifications. To identify hospital visits for related medical conditions, I use the Multi-level Clinical Classifications Software (CCS) developed by the Agency for Healthcare Research and Quality[3]. This classification system groups diagnoses (ICD-9-CM) into 18 Level 1 CCS categories (these are broad condition categories such as "Diseases of the circulatory system", "Diseases of the digestive system", and "Diseases of the respiratory system"). Visits assigned to the same Level 1 CCS category based on principal diagnosis are considered to be related to the same medical condition for aggregation purposes.

I combine the patient data with data on hospital characteristics from SPARCS, the New York State Department of Health, Institutional Cost Reports, Hospital Compare, and the American Hospital Association (AHA). I used Google Maps to calculate the driving distance from the geographic centroid of a patient's zip code of residence to the street address of each hospital; as some zip codes could not be located, information on travel distance is not available for a small fraction of records.

[3] https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp

1.3. Conceptual framework

As mentioned in the introduction, several studies have documented, in a variety of contexts, that consumers who have purchased a product in the past are more likely to choose that same product in the current choice occasion than consumers who have bought an alternative product before. As Heckman (1981) points out, there are two explanations for this empirical regularity. First, current choices might change relevant elements of the choice environment (such as prices, preferences, or choice sets) for future purchase occasions. This is referred to as state dependence. Second, consumers might have serially correlated unobserved preferences that make them choose the same alternative over time. If these unobserved preferences are not adequately controlled for, then past purchases and current choice probabilities would be linked even if past choices do not modify the current choice environment as in the first case. This is referred to as spurious state dependence. Therefore, persistence in choices is not enough proof of the presence of state dependence. In the context of hospital choice, if I observe a patient who chooses the same hospital each time she needs medical care, it might be the case that she evaluates the characteristics (quality, convenience, etc) of different alternatives in each occasion and then decides to visit the same hospital.

In order to provide a reference point for the empirical analysis, I consider the identification problem in the context of a model of hospital choice. The discussion only intends to illustrate the empirical challenges that I face and the possible strategies to deal with them; each section of the empirical analysis will use a particular framework to address the questions of interest. Patients experience health shocks over time. These shocks determine the medical conditions for which a patient needs medical care. In each episode, a patient visits a hospital. I treat the incidence and timing of hospital visits as exogenous. In each episode, the patient chooses the hospital with the highest utility[4]. Let $h_{it} = j$ denote patient i's admission to hospital j in episode t. The utility that

[4] I assume that there is no outside option.

patient i obtains from alternative j in episode t is given by:

$$u_{ijt} = \alpha_{it} D_{ijt} + \beta_{it} Z_{jt} + \gamma_{it} \mathbb{I}(s_{it} = j) + \varepsilon_{ijt} \tag{1.1}$$

where D_{ijt} is the travel distance from the patient's home to the hospital in episode t, Z_{jt} is a vector of hospital attributes in episode t, $\mathbb{I}(x)$ takes the value one if x is true and zero otherwise, s_{it} is the loyalty state of the patient in episode t (which summarizes the history of her past hospital visits), and ε_{ijt} captures (possibly persistent) intrinsic preferences of the patient for the hospital. I consider first order state dependence. The loyalty state of the patient is determined by the hospital used in the previous episode: if the patient visited hospital j in episode $t-1$, then $s_{it} = j$. The parameters of the model $\theta_{it} = (\alpha_{it}, \beta_{it}, \gamma_{it})$ might depend on characteristics of the patient, some of which could change across episodes and/or be unobserved by the researcher.

The representation of the patient choice process largely follows the prior literature on hospital choice: the main determinants of hospital choices of patients are convenience (captured by the distance from the patient's home to the hospital) and hospital quality. However, the evaluation of an alternative also depends on whether the patient used that hospital in the previous episode. In particular, a consumer receives a utility premium γ (which I refer to as the state dependence parameter) from visiting the same hospital as in the previous episode[5]. This effect should be interpreted in a causal sense: if a patient was exogenously assigned to a certain hospital in episode $t-1$, then she is more likely to choose that same facility in episode t than an otherwise similar patient. The formulation of utility therefore means that if health shocks or other events made otherwise similar consumers (in terms of their characteristics in the current episode) gravitate towards different hospitals in the past, then their current choices will be different.

Although this formulation of utility is quite general and is similar to models used in previous work, it is restrictive in many dimensions. First, it assumes that patients are myopic: the evaluation

[5] An equivalent interpretation is that the patient has to pay an incremental cost conditional on switching hospitals.

of different alternatives only depends on the characteristics of the current episode. In particular, patients do not take into account that current choices will impact future decisions due to lock-in. Second, the loyalty state of the patient could be a more complicated function of past choices, not just of the choice made in the last episode. There is a trade-off between capturing all of the dynamics, and keeping the analysis transparent and tractable. While switching costs in my setting might have both learning and transactional aspects (Farrell and Klemperer, 2007), the model above only captures the latter: a patient who switches from hospital A to hospital B would have to pay the switching cost if she later goes back to A. I focus on first-order state dependence because: i) Many of the factors that drive of state dependence operate through the choice made in the immediately preceding episode, and; ii) I expect that the effect of the immediately preceding episode is stronger than the effects of more distant episodes. Third, some of the drivers of state dependence discussed in the introduction are associated with channels that do not operate through the utility function *per se*. For example, state dependence could arise as the hospital used in the previous episode is more likely to enter the consideration set of the patient in the current episode, but with no utility premium from choosing this hospital over other alternatives in the consideration set. In this case, Equation 1.1 is a reduced form representation of the decision process of the patient. Fourth, the switching cost faced by a patient is symmetric across alternatives: the loyalty premium that a patient gets from sticking to hospital A is equal to the premium from staying with any other hospital B (in other words, γ is independent of j). Fifth, there are no complementarities between different hospitals. However, it might be the case that the cost of switching from hospital A to hospital B is different than the cost of switching from hospital C to hospital B. In some parts of the empirical analysis, I will be able to relax some of these assumptions.

Suppose that I estimate this model from data on patients' actual choices. The identification problem arises from the potential endogeneity of the loyalty state variable. The fact that patient i chooses hospital j in episode t implies that $E[\varepsilon_{ijt}|d_{it} = j] > E[\varepsilon_{ijt}]$ [6]. If the random component of

[6] More precisely, it implies that $\delta_{ijt} + \varepsilon_{ijt} \geq \max_{k \neq j} \delta_{ikt} + \varepsilon_{ikt}$, where $\delta_{ijt} = \alpha_{it} D_{ijt} + \beta_{it} Z_{jt} + \gamma_{it} \mathbb{I}(s_{it} = j)$. The

utility is correlated across episodes, then $d_{it} = j$ implies that the error term associated to alternative j for patient i will likely be high in episode $t+1$. As a consequence, the estimated state dependence coefficient captures the underlying unobserved propensity of the consumer to choose alternative j, and not just the structural effect of the previous choice on current utility. This is a standard selection problem. Note that the serial correlation of the error term can arise from several sources, such as misspecification of the distribution of the taste coefficients α_{it} and β_{it}, omitted variables, and measurement error. This concern seems particularly well founded in my setting, where there are many attributes of patients and hospitals that I do not observe (religious affiliation of the patient, details about amenities of the hospital, etc). Therefore, I cannot take a positive value of the estimated γ as conclusive evidence of the presence of state dependence.

To deal with this issue, the ideal design would randomly assign patients to hospitals in episode $t-1$ and analyze their choices in episode t; in this case, the loyalty state variable in episode t is uncorrelated with the preferences of the patient[7]. Given that most studies rely on observational data for the analysis, identification has typically relied on both functional form assumptions about the nature of heterogeneity and choice set variation across choice occasions (Sudhir and Yang, 2014). As Torgovitsky (2016) points out, the first strategy addresses the issue of identification, at least from a mathematical point of view. We could make distributional assumptions about the parameters and the error term in the utility function[8] and estimate the resulting model via maximum likelihood. The problem with this approach is that its validity depends on correctly specifying the distribution of unobserved heterogeneity. As Dubé et al. (2010) point out, any persistent preference heterogeneity not captured by the model will be loaded onto the econometric error term, leading us to conclude incorrectly that there is state dependence. For example, Dubé et al. (2010) show that allowing for a

expression in the main text better illustrates the empirical challenge faced by the researcher.

[7]This allows us to deal with the selection problem and therefore identify state dependence given the structure imposed by Equation 1.1. However, random assignment is not enough to point identify state dependence in a more general sense. See Subsection 1.5.4 for a discussion about this issue.

[8]For example, if we assume that ε_{ijt} has an extreme value type I distribution, and that $\theta_{it} \sim F(\xi; X_{it})$, where X_{it} are observable characteristics of the patient and F is a c.d.f. indexed by ξ, then we have a random coefficients logit model.

flexible pattern of heterogeneity can lead to different conclusions than more traditional approaches. Most studies rely (at least in part) on parametric assumptions to separately identify the sources of persistence. Exceptions are Torgovitsky (2016) and Illanes (2016), who recover the values of the parameters that are consistent with the identifying restrictions under different distributions of preference heterogeneity.

Exploiting choice set variation across episodes provides more transparent and credible evidence on the determinants of persistence in choice. The idea is to break the link between the previous choice and unobserved preferences of the consumer, so the selection problem is eliminated or at least attenuated. Heckman (1981) argues that to distinguish state dependence from latent heterogeneity I need a sufficiently large variation in the choice set to induce purchases that would not have been made otherwise. Dubé et al. (2010) exploit temporary price changes to identify state dependence in the context of choice of branded products. Suppose that consumers are induced to switch away from their preferred products by price discounts on other goods. If consumers continue purchasing the "less-preferred" products after prices return to normal levels, then this points out to state dependence as the source of choice persistence. Goldfarb (2006b) studies consumers' website choices and exploits product unavailability (caused by Internet denial of service attacks) to identify lock-in. Sudhir and Yang (2014) exploit the mismatch between previous choice and previous consumption created by free upgrades (which are mainly due to inventory shortages) in the context of car rentals to separate state dependence from unobserved heterogeneity. Israel (2005) points out that I can compare two individuals that face the same decision today and have identical loyalty states: one that was forced to use a certain alternative j in the previous episode by an exogenous shock, and another one who chose that same product voluntarily; under selection, the exogenous shock would produce a relatively high number of suboptimal matches among the affected population, which will make consumers depart from alternative j once we return to the usual choice environment.

In the next two sections, I follow this strategy to provide credible evidence about the sources of persistence in hospital choices of patients. In particular, I exploit quasi-exogenous shocks that shift the loyalty state of a patient: emergency hospitalizations (Section 1.4) and temporary hospital closures (Section 1.5).

1.4. Emergencies

In this section, I exploit emergencies as a quasi-exogenous source of variation in the loyalty state of patients. The strategy is to analyze the hospital choice of a patient in episode t following an emergency hospitalization in episode $t-1$. By emergency, I mean an episode in which the patient needs immediate medical care and there is little scope for choosing a particular hospital. I consider emergencies that induce the patient to try a hospital other than the one she had been using. I analyze whether the patient continues using this "new" hospital in the future. As the loyalty state of the patient is initiated by an emergency hospitalization, and to the extent that the new hospital choice is responsive to her preferences, repurchase behavior reflects the extent of state dependence.

The first step is to define emergencies. The identifying assumption is that the facility visited for an emergency hospitalization is determined by factors other than the preferences of the patient: the ambulance transport decision and the location of the patient at the time of the health shock (Doyle et al., 2015). This assumption would be violated if: 1) The patient or any of her surrogates requests transportation to a particular hospital during the emergency episode; 2) Patients choose where to live based on health status, so the locus of treatment during an emergency was "chosen" prior to that episode[9,10]. Since the assumption that the hospital used in an emergency is not determined by the preferences of the patient is not verifiable, I take several steps to make the assumption more

[9]For example, patients with heart disease might decide to locate close to their preferred hospital, so in the event of a heart attack they are likely to be taken to that facility.

[10]Even if the patient is not involved in the choice process, the emergency hospital could reflect her preferences. Consider two patients who live on opposite sides of the same zip code. I only observe the zip code of residence of a patient, but not her exact address. There are two hospitals, one on each side of the zip code. If patients are taken to the closest facility in an emergency, then the two patients go to different hospitals in that episode. Then, the locus of treatment in an emergency indicates which facility is more convenient for the patient. If the disutility of travel is high, repurchase behavior could reflect unobserved preferences for the emergency hospital (actual distance to the emergency hospital is smaller than in the data). This should be less of a concern the smaller the zip code. I control for this consideration in the analysis below.

plausible.

Ideally, I would define emergencies as episodes where the patient suffers a severe and unexpected health shock and arrives to the hospital by ambulance. Unfortunately, I cannot distinguish in my data whether a patient arrived to the hospital by ambulance or self-transport. Given this limitation, I define emergencies as episodes in which the patient is admitted to the hospital through the emergency department (ED) for a severe medical condition that requires immediate care. These non-deferrable conditions correspond to admitting diagnoses (ICD-9-CM) with similar admission rates through the ED on weekdays and weekends (Card et al., 2009). These conditions represent 6% of all ED admissions, and are extremely acute and often life-threatening. Table 1.1 shows the most common non-deferrable conditions in the full dataset and characteristics of these episodes. Finally, I exclude emergencies in which the patient is admitted to the hospital from a health care facility (e.g. a skilled nursing facility). As the patient is likely to have chosen a health care facility close to her preferred hospital, the use of a hospital during the emergency reflects strong preferences for the facility. In summary, I am confident that the emergencies that I consider are episodes with limited scope for the patient to choose the hospital.

Once I identify an emergency according to the criteria outlined above, I analyze the hospital choice of the patient in the first episode following the emergency (I refer to this episode as the current episode)[11]. In this episode, the loyalty state of the patient is determined by the hospital used for the emergency hospitalization. For example, if the patient used hospital A in the emergency episode, then the patient is loyal to hospital A at the time of the next episode. In the analysis, I consider two situations:

[11] As explained in Section 1.2, I drop readmissions from the working sample. Therefore, I analyze whether the patient continues going to the emergency hospital for episodes not directly related to the emergency itself. In the main specification, I use a 90 days window to identify readmissions. To ensure robustness, I also performed the analysis using other time windows (30, 60, and 120 days), without any substantial change in the nature of the results. These results are available upon request.

Table 1.1: Most common non-deferrable conditions

Description	ICD-9-CM	Weekend	Intensive care	Emergency (SPARCS)	30 days mortality	Length of stay	Admission hour		
							p25	p50	p75
Intertrochanteric fx-cl	820.21	0.28	0.15	0.96	7.9%	7.1	10	15	19
Cellulitis of face	682.0	0.28	0.04	0.98	0.9%	3.7	11	15	18
Cardiac arrest	427.5	0.29	0.81	0.97	75.2%	8.1	9	14	18
Poison-medicinal agt NOS	977.9	0.29	0.45	0.97	1.5%	3.4	7	14	19
Rhabdomyolysis	728.88	0.28	0.15	0.98	4.3%	6.2	10	15	19
Epistaxis	784.7	0.29	0.17	0.98	5.1%	4.1	8	13	19
Ac alcohol intox-contin	303.01	0.29	0.11	0.98	0.5%	3.5	7	15	19
Closed fracture of pubis	808.2	0.29	0.08	0.96	4.2%	5.3	10	15	19
Subarachnoid hemorrhage	430	0.28	0.74	0.98	20.0%	11.8	8	14	19

Notes: This table shows the most common non-deferrable conditions among emergency department (ED) admissions in the full inpatient data. To determine if a diagnosis is considered non-deferrable, I test whether the fraction of ED admissions that occur during the weekend is statistically different from 2/7. The reported statistics correspond to episodes in which the patient is admitted to the hospital through the ED. Intensive care refers to whether the patient spent time in an intensive care unit during the hospital stay. The emergency variable from SPARCS indicates an episode in which "the patient requires immediate medical intervention as a result of severe, life threatening, or potentially disabling conditions." The other two major categories used by SPARCS to categorize hospital admissions are: 1) Urgent: The patient requires immediate attention for the care and treatment of a physical or mental disorder. Generally the patient is admitted to the first available and suitable accommodation; 2) Elective: The patient's condition permits adequate time to schedule the admission based on the availability of a suitable accommodation.

1. The current loyalty state was initiated by the emergency. Moreover, the emergency hospital had never been used by the patient before[12]. Therefore, I exclude emergencies in which the patient goes to a hospital different from the last one she used before the emergency but that she used at some point in the past. As a result, the emergency produces a strong shift in the loyalty state of the patient: her choices before the emergency reveal that she does not have strong preferences for the facility used in that episode. Therefore, the repurchase behavior of the patient in the current episode reflects the extent of state dependence.

2. The current loyalty state was initiated before the emergency: the emergency hospital is the same facility that the patient had been using before. In this case, there is a selection issue: the choices of the patient before the emergency reveal that she has strong preferences for that facility. Therefore, the repurchase behavior of the patient in the current episode captures both state dependence and persistent preference heterogeneity.

The repurchase rate is the fraction of patients who in the current episode choose the same hospital used for the emergency hospitalization. Given that the distribution of choice probabilities for a patient might change across episodes, the raw repurchase probability is not very informative about the impact of previous choices on current behavior. For example, in cases where the previous hospital does not offer the type of surgery required in the current episode, the repurchase rate would be zero even in the presence of state dependence. If the hospital used during the emergency is the best hospital for treating the current medical condition of the patient, then a high repurchase rate reflects both state dependence and the quality of the match between the patient and the facility. Therefore, I compare the repurchase rate with the marginal probability of choosing the emergency hospital based on characteristics of the current episode. This way, I measure the likelihood that a patient who used a given hospital j in episode $t-1$ chooses that hospital in episode t relative to an

[12]More precisely, the patient did not visit the hospital between 1995 (the first year for which I have patient data) and the day of the emergency. Because I only consider emergencies that take place on or after 2005, this restriction means that the patient had not used the facility for at least 10 years before the emergency.

observationally similar patient. To calculate the patient's marginal choice probability, I assume that it is proportional to the market share of the hospital within the group of similar patients (see Raval et al. (2017a) and Carlson et al. (2013)). More precisely, I use the following process:

1. Using the full dataset, I define cells of equivalent episodes based on zip code, diagnosis, type of visit, and admission year. Note that a patient might transition across different cells over time (for example, if she moves or if she seeks hospital care for different medical conditions). I define these cells to be as precise as possible while maintaining sufficient sample sizes to determine hospital market shares within them. In the main specification, I keep cells that have at least 20 observations. Denote the market share of hospital j within cell k by s_{jk}.

2. I assign each episode following an emergency hospitalization to the corresponding cell. Denote the hospital used in episode i by h_i and the hospital used by the same patient in her previous episode (the emergency hospitalization) by h_i^b.

3. For each episode, the marginal probability of choosing the emergency hospital is the market share of this facility within the episode's cell. Denote the marginal probability of choosing the emergency hospital in episode i by $p(i) = s_{wk}$, where $i \in k$ and $w = h_i^b$.

4. For each $p \in [0,1]$, the corresponding excess repurchase probability is given by the mean of $x_i = \mathbb{I}\{h_i = h_i^b\} - p(i)$ over episodes with $p(i) = p$.

The idea is to look at patients with different loyalty states (they used different hospitals for the emergency episode) but with the same probability p of choosing the emergency hospital in the current episode. Consider episodes where the loyalty state was initiated by the emergency. If previous choices do not have an impact on current decisions, then we should expect the repurchase rate to be p (so the excess repurchase probability is zero). If the excess repurchase probability is positive, I take that as evidence of state dependence. For episodes where the loyalty state was initiated before the emergency, the excess repurchase probability captures both state dependence

Table 1.2: Episodes following an emergency hospitalization

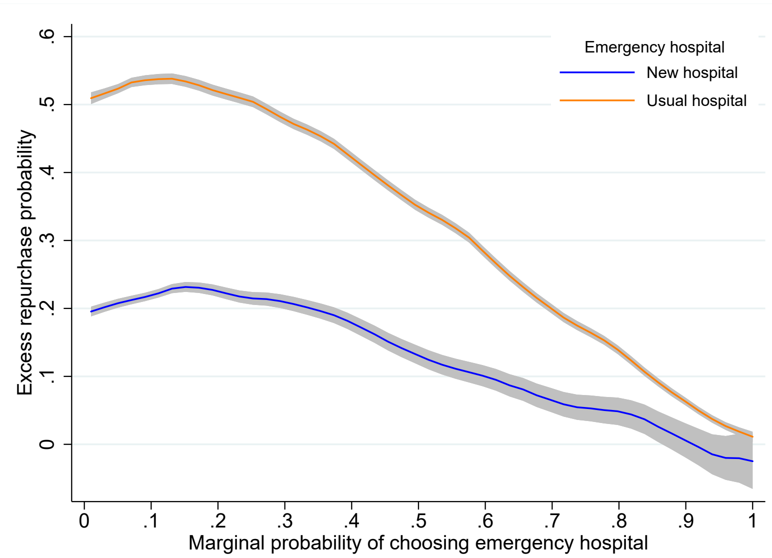
	Emergency hospital	
	New	Usual
Number of cases	40,549	105,390
Episodes before emergency:		
Number of episodes	1.6	3.8
Propensity to use emergency hospital	0	0.82
Episode after emergency:		
Days since emergency	393	349
Repurchase rate	0.42	0.78
Marginal choice probability	0.24	0.39
Inpatient	0.27	0.31
Emergency Department	0.59	0.57
Ambulatory Surgery	0.14	0.12

Notes: This table shows summary statistics of the episodes following an emergency hospitalization. I distinguish episodes based on whether the emergency hospital was being used by the patient before the emergency or is a new hospital. The prior propensity to use the emergency hospital is defined as the proportion of the patient's hospital visits prior to the emergency episode that were to the emergency hospital. The repurchase rate is the proportion of patients who chose the emergency hospital in the episode following the emergency. Inpatient, Ambulatory Surgery and Emergency Department categorize hospital visits. The marginal choice probability is the probability of choosing the emergency hospital based on characteristics of the patient in the current episode (the next episode after the emergency hospitalization).

and unobserved heterogeneity. Doing the same exercise for all possible values of p, I recover the excess repurchase probability schedule. In practice, the latter is obtained by a locally weighted regression of x_i on $p(i)$. In the main specification, I only consider cases where the marginal choice probability is higher than 0.01. If the marginal probability is lower than this value, it most likely corresponds to a case where the patient will not consider the facility for hospital care.

Table 1.2 shows summary statistics of the current episodes used in the analysis. I distinguish episodes depending on whether the emergency hospital had never been used before the emergency or was the usual hospital of the patient. There are two differences between these groups to point out. First, patients in the second group had on average more episodes before the emergency

Figure 1.1: Excess repurchase probability, all cases



Notes: This figure shows the excess repurchase probability as a function of the marginal probability of choosing the emergency hospital. Included: episodes following an emergency hospitalization. There are separate schedules for episodes following emergencies that shifted and did not shift the loyalty state of the patient. The shaded areas represent 95% confidence intervals.

hospitalization than patients in the first group. Second, patients who experience a shift in their loyalty state as a result of the emergency episode have lower probability of returning to the emergency hospital based on their observable characteristics.

I construct the excess repurchase probability schedule pooling across all emergencies, hospitals, and current episodes. The results are shown in Figure 1.1. There are separate schedules for episodes following emergencies that shifted (blue) and did not shift (orange) the loyalty state of the patient. There are two main points to be noted.

First, for cases where the loyalty state of the patient was initiated by the emergency hospitalization, the excess repurchase probability is positive over the unit interval. This indicates that a patient who, due to an emergency, visited a certain hospital j in episode $t-1$ for the first time is more likely to choose hospital j in episode t than a patient with the same characteristics. Moreover, the impact of the previous choice on the current decision is large: for example, if based on current characteristics there is a 20% probability that the patient chooses hospital j, then the choice probability increases to more than 40% for patients who used that same hospital in the previous episode. This is evidence of the presence of state dependence in hospital choices of patients. In general, a positive excess repurchase probability could also signal the presence of unobserved heterogeneity; however, as explained above, I address this concern by focusing on loyalty states initiated by emergencies.

Second, the excess repurchase probability is higher for episodes following an emergency that did not shift the loyalty state of the patient. As explained above, in this case repurchase behavior not only reflects state dependence, but it also captures the latent propensity of the patient to choose the facility used in previous episodes. Therefore, the higher repurchase rate for these cases reflects the presence of substantial unobserved preference heterogeneity. As a result, a naive analysis that does not take into account the endogeneity of previous choices will overstate the magnitude of state dependence.

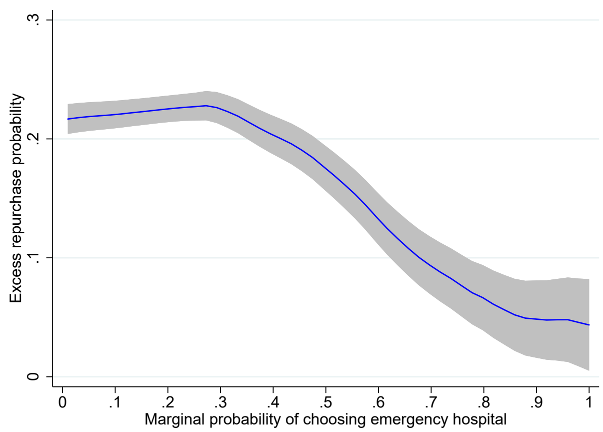
To assess the robustness of my findings, I construct the excess repurchase probability schedule for episodes following different types of emergencies. The results are shown in Figure 1.2. The qualitative results are the same as in the main analysis.

The previous results could mask heterogeneity in persistence across different types of hospitals. In particular, there might be differences in loyalty towards high and low quality hospitals. The pattern of heterogeneity might provide insights about the determinants of state dependence. I restrict attention to episodes following an emergency that induced a shift in the loyalty state of the patient, so repurchase behavior reflects state dependence. I distinguish cases based on the quality of the emergency hospital. Quality is not observable, so I use teaching status as a proxy for high quality[13]. Figure 1.3 shows that the excess repurchase probability schedules of teaching and non-teaching hospitals are similar. Although not conclusive evidence, this suggests that learning is not the primary driver of the observed persistence: patients are equally loyal to high and low quality hospitals.

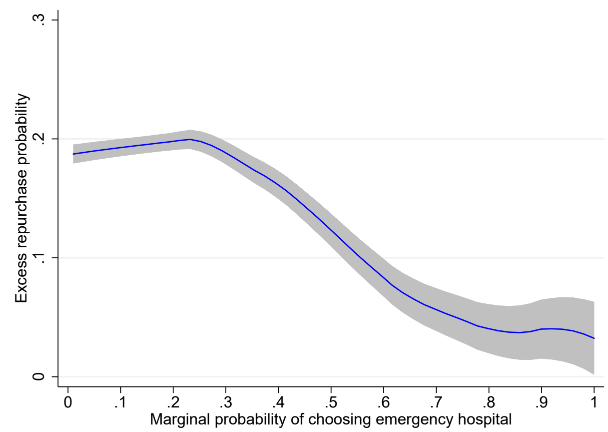
[13]Teaching status is obtained from the AHA Annual Survey Database and refers to hospital membership in the Council of Teaching Hospitals (COTH).

Figure 1.2: Excess repurchase probability, by type of emergency

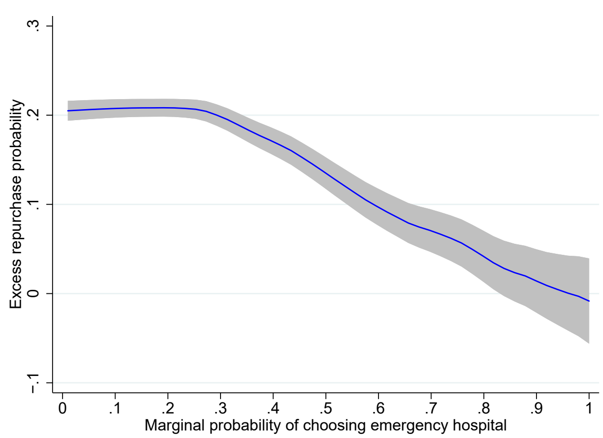
(a) Intensive care



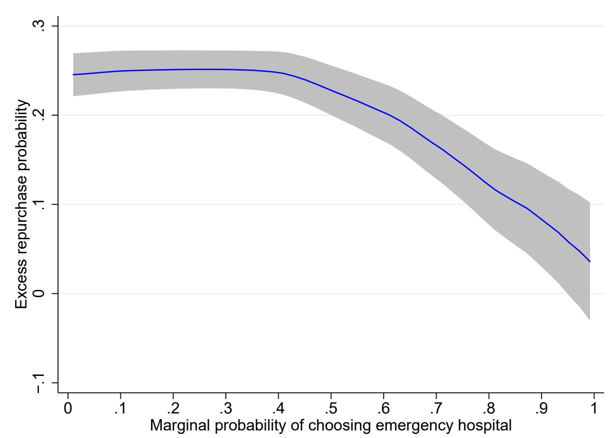
(b) Injury or poisoning



(c) Night admission



(d) Heart attack and stroke

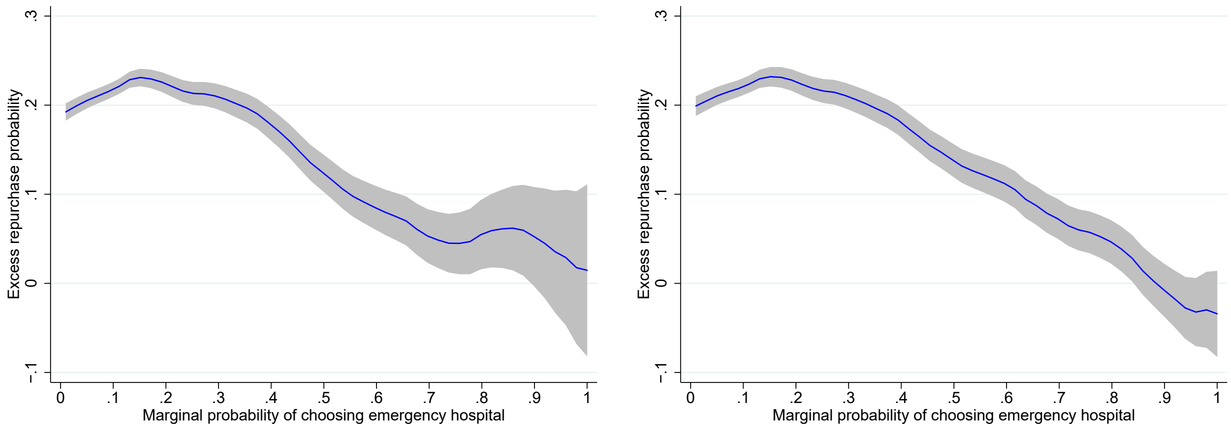


Notes: These figures show the excess repurchase probability as a function of the marginal probability of choosing the emergency hospital. Included: episodes following an emergency hospitalization that shifted the loyalty state of the patient. Different figures correspond to different types of emergencies: a) During the emergency hospitalization, the patient spent time in an intensive care unit; b) The emergency was coded as injury or poisoning; c) The emergency admission took place during the night (between 9PM and 5AM), and; d) The emergency was coded as heart attack or cerebrovascular accident. The shaded areas represent 95% confidence intervals.

Figure 1.3: Excess repurchase probability, by type of hospital

(a) Teaching

(b) Non-teaching



Notes: These figures show the excess repurchase probability as a function of the marginal probability of choosing the emergency hospital. Included: episodes following an emergency hospitalization that shifted the loyalty state of the patient. Different figures correspond to different types of emergency hospital. The shaded areas represent 95% confidence intervals.

1.5. Hurricane Sandy

1.5.1. Setting

Hurricane Sandy hit the New York Metropolitan area at the end of October 2012. Damage from the storm led to the temporary closure of three hospitals in New York City: NYU Langone Medical Center, Bellevue Hospital Center, and Coney Island Hospital. Bellevue and NYU Langone are located in Manhattan next to each other, while Coney Island Hospital is located in Brooklyn. Table 1.3 provides a basic description of the affected hospitals. The facilities differ along various dimensions, so there is heterogeneity in the settings that I analyze. NYU Langone is an academic medical center with a high proportion of privately insured patients. The other two hospitals are part of the city's Health and Hospitals Corporation and attract mostly Medicare, Medicaid, and uninsured patients. NYU Langone and Bellevue are large hospitals (more than 900 beds), while Coney Island Hospital is a medium size facility (371 beds). In terms of service offerings and designations, NYU

Table 1.3: Characteristics of affected hospitals

	Coney Island Hospital	Bellevue	NYU Langone
Visits:			
Ambulatory Surgery	4,487	7,697	17,710
Emergency Department	60,621	93,538	36,475
Inpatient	17,580	26,763	33,095
Demographics, inpatient:			
Age	50.6	42.7	43.6
Female	0.53	0.39	0.58
Medicare	0.35	0.17	0.27
Medicaid	0.48	0.49	0.08
Private	0.09	0.08	0.62
Other insurance	0.08	0.25	0.03
Certified beds:			
Total	371	912	987
Services:			
Perinatal Designation	Level 2	Regional Center	Regional Center
CABG	No	Yes	Yes
Transplant Center	No	No	Yes

Notes: This table describes the hospitals affected by Hurricane Sandy. The number of visits and the demographic profile of patients correspond to the 12 month period preceding the storm (November 2011 - October 2012).

Langone is the most sophisticated, followed by Bellevue. Although Coney Island Hospital offers a wide range of services, it does not provide the most complex services.

The affected hospitals remained closed for repairs and renovations during several weeks, forcing patients usually served by these facilities to find an alternative hospital for their medical care needs. Consider, for example, a patient who receives medical care at Coney Island Hospital and had never gone elsewhere before. If the patient needed medical care during the time this facility was closed, she would have had to go to another hospital, such as Maimonides Medical Center. In this section, I study whether the temporary unavailability of the affected hospitals had a long-lived impact on patients' preferences: Does the patient in the example continue going to Maimonides

Medical Center once Coney Island Hospital reopens or does she return to her usual hospital? Do affected patients become long-term patients of the new facilities? Anecdotal evidence suggests that the possibility of permanently losing patients was a concern for administrators at the shuttered hospitals[14].

This natural experiment is particularly well suited to study the dynamics of hospital choice. First, the type of choice set variation produced by the storm is ideal given how competition takes place in the hospital industry. Second, the hospital closures were unexpected and unrelated to patients' preferences for different facilities, thus providing a quasi-exogenous source of variation in hospital choice. Third, the data allows me to identify the patients most likely to have been affected by the temporary hospital closures: patients with strong preferences for an affected facility who needed hospital care during the unavailability window. Therefore, I can identify patients who would have chosen one of the affected hospitals had it been available, but were forced to choose a second-best option. There are two main limitations of my analysis. First, I only observe hospital choices of patients for less than three years after the affected facilities reopened. Second, I analyze a specific empirical context, which places limitations on the external validity of my conclusions. In particular, the analysis is not designed to provide estimates of the extent of state dependence in other settings.

Raval et al. (2017b) exploit unexpected hospital closures in different markets following a natural disaster to analyze the substitution patterns predicted by different models of hospital choice. One of the natural disasters that they consider is Hurricane Sandy. In one of their specifications, the authors identify patients who used the affected hospitals in the pre-storm period as those most likely to have experienced the closure of their preferred hospital. However, whether their continued preference for the shuttered facilities is due to switching costs or unobserved heterogeneity is irrelevant for their analysis, so they do not attempt to separate the channels. The main objective of my analysis is to separately identify state dependence from persistent unobserved heterogeneity, while the nature of

[14] http://www.nytimes.com/2012/12/04/nyregion/with-some-hospitals-closed-after-hurricane-sandy-others-overflow.html?mcubz=1

substitution patterns is not a primary concern.

The analysis proceeds in two steps. First, I show that patients who needed hospital care while a hospital was closed (treatment group) are less likely to visit that facility in the future than patients who did not have hospital visits during the unavailability window (control group). Moreover, non-returning patients in the treatment group favor the facility used during the unavailability window. I show this in a clear way, without relying on complex estimators. Then, I cast the setting into a nonparametric framework. By imposing an independence restriction between preferences and the timing of hospital visits, I can reject the hypothesis of no state dependence under minimal assumptions about the nature of unobserved preference heterogeneity.

1.5.2. Sample construction

For each case study[15], I construct a panel of patients' hospital choices. As explained in Section 1.2, readmissions and visits for follow-up care are excluded from the analysis, so hospital visits correspond to different episodes of care. I divide episodes into three periods based on the date of the patient's admission to the hospital: period 0 corresponds to the pre-storm period, period 1 is the time window during which the affected hospital remained closed for repairs, and period 2 goes from the reopening of the shuttered facility trough the end of the sample period. For each patient, I only consider episodes that took place while the patient was living in the service area of the affected hospital: I want to use information on those episodes where the patient is likely to consider this facility for hospital care. In the case of CIH, this step would remove, for example, episodes that take place while the patient lives in Buffalo or Manhattan - so CIH is not viewed as a practical alternative - but would keep hospital visits by that same patient while she lives in southern Brooklyn.

To construct the service area of a hospital, I follow the methodology of Raval et al. (2017b). I identify the smallest collection of zip codes that accounted for 90% of inpatient discharges from

[15] In what follows, I emphasize the case of Coney Island Hospital (CIH, henceforth), because the definitions of service area and the competitive set are more transparent for CIH than for Bellevue and NYU Langone. However, the methodology outlined below applies to all three cases.

the facility in the year prior to the storm. The resulting area might include zip codes where the hospital is competitively insignificant. Therefore, I exclude zip codes where the facility had a market share below 4% in the year before the storm. Finally, I make some adjustments to ensure the geographic contiguity of the resulting service area. Figures A.1 through A.3 in Appendix A show the service area of each of the affected hospitals. I repeated the analysis using alternative thresholds to determine the inclusion of zip codes in the service area and the qualitative results (not reported) are consistent with the baseline results discussed here.

I construct the working sample to exploit the features of the data that allow me to identify the effects of interest. I keep patients from the full sample who: 1) Had at least one hospital visit in both periods 0 and 2[16], and; 2) Did not have any hospital visits while the affected facility was closed or visited only one hospital during that time window. For each individual in the resulting sample, I keep the last episode of period 0, all episodes (if any) in period 1, and the first episode of period 2. Therefore, for each patient I know: 1) The hospital chosen in the last episode before the storm; 2) Whether she needed hospital care while the affected facility was closed and, if so, which hospital she used, and; 3) The hospital chosen in the first episode of period 2. I distinguish patients based on whether they visited a hospital in period 1. I refer to the set of patients who needed hospital care while the affected hospital was closed as the treatment group, while the other patients constitute the control group.

For the analysis, I focus on the hospital choices of patients in the first episode of period 2. I exclude a small number of cases where the affected hospital does not seem to be in the market for the type of medical care required by the patient in that episode. I infer provision of the required services by aggregating all records in the original data into cells defined by unique combinations of claim type, diagnosis category and semester, and observing the number of patients with those characteristics receiving care at the corresponding facility. I then use a three visit threshold to

[16]The restriction that patients have at least one episode in period 0 has the benefit that I can use information on actual hospital choices to infer the strength of their preferences for the affected hospital.

determine if the hospital is a feasible alternative for the patient. This step removes approximately 5% of patients.

I refer to the sample that results from the selection steps outlined above as sample #1. I define two subsamples that I use in the subsequent analysis. Sample #2 contains those patients who were loyal to the affected hospital at the moment of the storm (they visited that facility in the last episode of period 0), while sample #3 contains those patients who chose either the shuttered hospital or one of its main competitors in all the three episodes considered. To identify the main competitors of an affected hospital, I compute the diversion ratio from that facility to other hospitals in the year before the storm. I define the six hospitals with the highest diversion ratios as the main competitors of the affected facility. Tables A.1 through A.3 in Appendix A show information on the demographic profile and other characteristics of patients in the three samples. There are two main differences between control and treatment groups. First, the treatment group is older and sicker: it has a higher proportion of inpatient episodes and Medicare/Medicaid patients than the control group. Second, patients in the treatment group used the affected hospital in the pre-storm period less frequently than patients in the control group. Although I control for these differences in the analysis below, the imbalance along the latter dimension points to potential threats to identification.

1.5.3. Reduced form evidence

I now discuss patterns in the data that help identify the presence of state dependence. Consider two patients with similar underlying preferences for CIH (idiosyncratic tastes or unobserved characteristics - separate from state dependence - that make them gravitate towards this hospital) and who were loyal to that hospital at the moment of the storm. The first patient needs hospital care in period 1 and therefore goes to a hospital other than CIH (given that this facility is closed), while the second patient does not need hospital care in period 1. I compare the hospital choices that these patients make in the first episode of period 2. In this episode, the patients have different loyalty states: the patient who got sick during period 1 is not loyal to CIH, while the patient who did not seek hospital care in period 1 remains loyal to CIH. If the temporary unavailability of CIH did not change the underlying preferences of patients for hospitals, then differences in choice probabilities between the two patients capture the impact of state dependence.

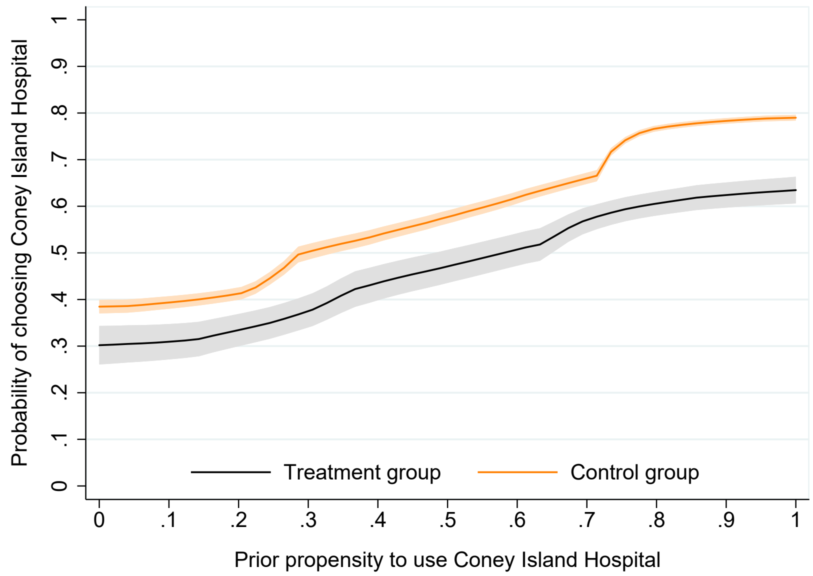
The identifying assumption is that the timing of hospital visits is exogenous. In other words, whether a patient seeks hospital care in period 1 or not is independent of her tastes for different facilities. In this case, the distribution of preferences is the same for patients in the treatment and control groups. The assumption seems a priori reasonable in my setting, given that patients most likely visit a hospital due to a health shock. However, there are ways in which the assumption could be violated. I discuss the potential threats to the validity of my approach at the end of the section.

I start by showing that patients who needed hospital care while an affected hospital was closed were less likely to return to that hospital in the first episode of period 2 than similar patients who did not have hospital visits in period 1. I use sample #2 for the analysis, so all the patients considered were loyal to the affected hospital at the moment of the storm.

Figure 1.4 shows, for both the treatment and control groups, the probability of choosing CIH in the first episode of period 2 as a function of the prior propensity to use the facility. This propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to this facility[17]. There are two things to note. First, the probability of choosing CIH in period 2 increases with the prior propensity to use that facility, for both the treatment and control groups. This is not surprising, given that the pre-storm propensity to use the affected facility captures the strength of the patient's preferences for that hospital. Second, patients in the treatment group are less likely to return to CIH after its reopening than patients in the control group with similar prior propensity. The gap in choice probabilities reflects a lasting effect of the temporary unavailability of the affected hospital on patients' choices. Figures 1.5 and 1.6 show the results for Bellevue and NYU Langone, respectively. For NYU Langone, the difference between treatment and control groups is not as clear

[17] Consider, for example, a patient who had five hospital visits in the pre-storm period. If three of these visits were to the affected hospital, then the prior propensity to use this facility is 0.6 (3/5).

Figure 1.4: Probability of choosing Coney Island Hospital in period 2



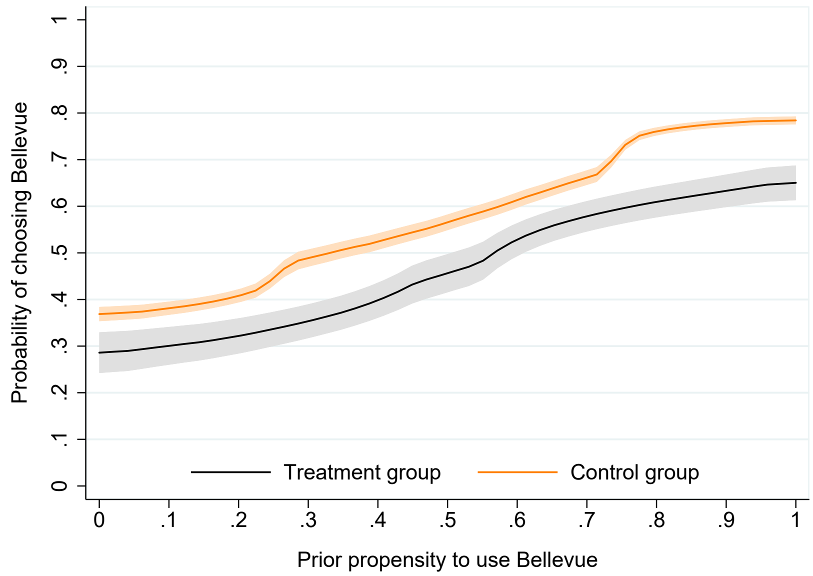
Notes: This figure shows the probability of choosing Coney Island Hospital in the first episode of period 2 as a function of the prior propensity to use the facility. This propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to Coney Island Hospital. Included: patients who were loyal to Coney Island Hospital at the moment of the storm. Kernel-weighted local regression estimates. The shaded areas represent 95% confidence intervals.

as in the other two cases; however, the analysis below shows that the difference in patient behavior is significant in this case.

To control for preference heterogeneity related to observables, I estimate a probit of choosing the affected hospital in the first episode of period 2 on the prior propensity to use the facility, zip code fixed effects, diagnosis fixed effects, a spline of the number of days between the reopening of the shuttered facility and the episode, and an indicator for treatment status. The results are displayed in Table 1.4. In the case of NYU Langone, having a hospital visit in period 1 is associated with a decrease of 6.5 percentage points in the probability of using the facility in period 2; the effect is statistically different from zero at usual significance levels. For CIH and Bellevue, the marginal effect of treatment status is -9.1 and -10.5 percentage points, respectively[18].

[18] Two possible explanations for the lower effect found in the case of NYU Langone are the following. First, NYU Langone might derive more loyalty from patients given its reputation and higher quality. Second, the hospital might have actively engaged in regaining patients who visited other hospitals during the time it was closed (for example, by

Figure 1.5: Probability of choosing Bellevue in period 2

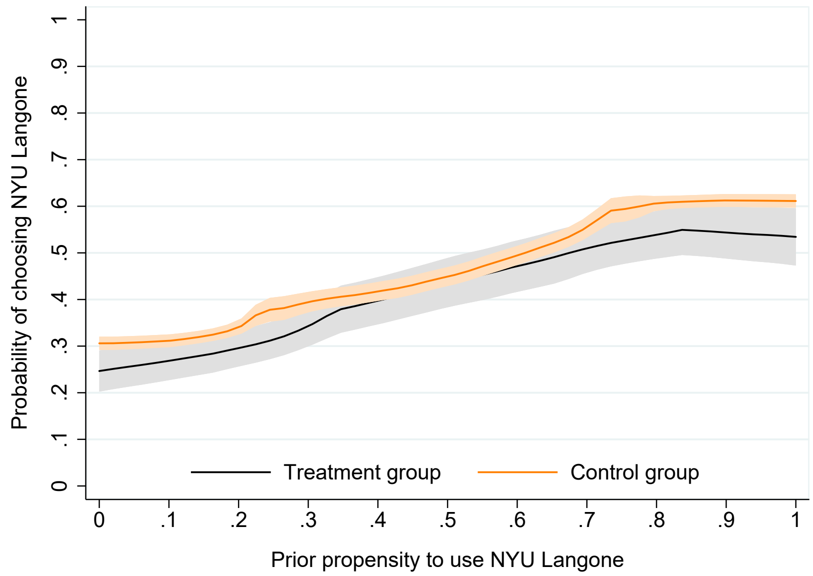


Notes: This figure shows the probability of choosing Bellevue in the first episode of period 2 as a function of the prior propensity to use the facility. This propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to Bellevue. Included: patients who were loyal to Bellevue at the moment of the storm. Kernel-weighted local regression estimates. The shaded areas represent 95% confidence intervals.

Thus far, I have shown that for patients in the treatment group the probability of choosing the affected hospital after its reopening decreased relative to patients in the control group. I now explore to what extent non-returning patients substitute the affected hospital in period 2 with the same facility they used in period 1. Even if patients are less likely to return to the affected hospital once it reopens, they might gravitate towards other hospitals in a proportional way (conditional on heterogeneity); however, if they continue using the hospital they visited in period 1 once the affected facility reopens, then this points to state dependence as the source of the difference in patient behavior. In order to capture the effect of interest, I estimate a conditional logit model of hospital choice using the first episode of period 2. Explanatory variables include hospital-type of visit fixed effects, the distance from the centroid of the patient's zip code of residence to the facility, interactions between the type of insurance of the patient and an indicator for public hospital, and an

contacting them).

Figure 1.6: Probability of choosing NYU Langone in period 2



Notes: This figure shows the probability of choosing NYU Langone in the first episode of period 2 as a function of the prior propensity to use the facility. This propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to NYU Langone. Included: patients who were loyal to NYU Langone at the moment of the storm. Kernel-weighted local regression estimates. The shaded areas represent 95% confidence intervals.

indicator for whether the hospital was used in the previous episode of care; in addition, I allow the utility from choosing the affected hospital to vary with the prior propensity to use that facility. The estimates from the model are reported in Table 1.5. Previous use of a facility is a strong predictor of the hospital choice of a patient: having used a hospital in period 1 increases the probability of choosing that same facility in period 2 more than four times above the baseline probability (the magnitude of this effect is similar across the three case studies).

In summary, I have provided evidence that patients who experienced a forced shift in their loyalty state are less likely to return to their usual hospital than similar unaffected patients. Moreover, patients who do not return to the affected hospital favor the new facility they used in period 1. These patterns indicate the presence of state dependence in hospital choices of patients. I now discuss in more detail the assumptions underlying my empirical strategy in order to identify potential problems.

Table 1.4: Probit estimates

Days since reopening	0	30	60	90	120	150
Panel A: Coney Island Hospital						
Treatment status	-0.091	-0.084	-0.080	-0.080	-0.075	-0.072
(Std Error)	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)	(0.012)
Sample size	41,153	38,640	36,318	34,174	32,084	30,095
Panel B: Bellevue						
Treatment status	-0.105	-0.100	-0.104	-0.108	-0.104	-0.099
(Std Error)	(0.012)	(0.013)	(0.014)	(0.015)	(0.016)	(0.017)
Sample size	25,550	23,684	21,891	20,287	18,856	17,560
Panel C: NYU Langone						
Treatment status	-0.065	-0.066	-0.066	-0.064	-0.062	-0.058
(Std Error)	(0.014)	(0.015)	(0.015)	(0.016)	(0.017)	(0.017)
Sample size	21,295	20,155	19,056	17,967	16,914	15,944

Notes: This table shows the results from the estimation of probit models where the dependent variable is one if the patient visited the affected hospital in period 2. The explanatory variables are an indicator for treatment status, propensity-bin fixed effects, zip code fixed effects, diagnosis group fixed effects, and a spline of the number of days between the reopening of the affected hospital and the episode. The indicator for treatment status is one if the patient needed hospital care while the affected hospital was closed for repairs. The estimates reported correspond to the average marginal effect of treatment status. Days since reopening refers to the episodes included in the estimation sample: if days since reopening is x, then only episodes that took place at least x days after the reopening of the affected hospital are included in the estimation sample.

First, I assume that the timing of hospital visits is independent of patients' preferences for hospitals. In other words, I assume that whether a patient seeks hospital care in period 1 or not is independent of her preferences for different hospitals. If patients with very strong preferences for the affected hospital postponed elective medical care until this facility reopened, then there is a selection problem. In particular, the control group would include a higher proportion of patients with strong preferences for the affected facility, which might explain the gap in choice probabilities between the two groups.

Table 1.5: Logit estimates

	Case study		
	Coney Island Hospital	Bellevue	NYU Langone
Used in previous episode	1.968	2.109	1.890
	(0.068)	(0.101)	(0.091)
Distance	-0.130	-0.476	-0.321
	(0.003)	(0.008)	(0.008)
Affected hospital x Prior propensity			
Prior propensity: 0 - 0.25	0.011	-0.058	-0.036
	(0.104)	(0.092)	(0.111)
Prior propensity: 0.25-0.50	0.338	0.419	0.381
	(0.067)	(0.072)	(0.077)
Prior propensity: 0.50-0.75	0.774	0.871	0.795
	(0.054)	(0.061)	(0.064)
Prior propensity: 0.75-1	1.217	1.442	0.999
	(0.040)	(0.044)	(0.043)
Public hospital x Insurance			
Insurance: Medicare	0.027	0.877	1.022
	(0.045)	(0.058)	(0.099)
Insurance: Medicaid	0.350	1.052	1.724
	(0.037)	(0.050)	(0.122)
Insurance: Other	1.315	1.969	2.290
	(0.044)	(0.058)	(0.115)

Notes: This table shows estimates of the conditional logit model of hospital choice in the first episode of period 2. The previous use variable is an indicator for whether the hospital considered was chosen by the patient in her previous episode of care. Prior propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to the affected hospital. In addition to the variables shown, the model includes hospital - type of visit fixed effects. The effect of previous hospital use is economically significant: having used a hospital in period 1 increases the probability of choosing that same facility in the first episode after the reopening of the affected hospital 4.4, 4.6, and 3.9 times above the baseline probability, respectively.

Second, the affected hospitals restored services gradually and sometimes at reduced capacity. If capacity constraints were important, I might be artificially inflating persistence due to state dependence: some patients might have wanted to receive care at CIH in period 2 but were unable to do so due to capacity constraints. For this to bias the results, it must be the case that patients in the treatment group were more affected by capacity constraints than patients in the control group.

To address these concerns, I perform the same analysis excluding episodes that took place close to the start of period 2. Capacity constraints likely were more of an issue during the first days after affected hospitals resumed operations, so inference using episodes that occurred later in period 2 should be less affected by bias from this channel. In addition, using episodes further away from the reopening date should lower the impact of selection into treatment status by patients who delay care. I estimate a probit of choosing the affected hospital in period 2 as described above, but excluding cases that took place less than 30, 60, 90, 120, and 150 days after the reopening date. The results are shown in columns 3-7 of Table 1.4. The decrease in the probability of choosing CIH associated with treatment group status is between 7.2 and 8.4 percentage points, depending on the time window considered. Although the effect is smaller than in the main specification, it remains large and significant. For Bellevue and NYU Langone, the effect of treatment status also decreases as I focus on episodes further away from the date of reopening, although the magnitude of the change is smaller than for CIH.

Third, I assume that the closure and unavailability of an affected hospital did not have any impact on patients' underlying preferences for the facility. In particular, this means that the quality of the hospital did not change and that patients did not update their beliefs about the quality of the hospital. If patients in both the control and treatment groups were equally affected along this dimension, this would not be a problem because identification is based on differences in choice probabilities. However, if patients in the treatment group are more likely to adjust their beliefs about the quality of the affected hospital (or if they feel that some aspects about the facility that they liked changed after its reopening), then this would be a confounding factor in the analysis. Although I cannot control for this effect, I believe that I can largely discount it: the closure was directly related to the damage produced by the storm, which was particularly destructive and affected many areas of the city.

In conclusion, while deviations from the identifying assumptions are certainly possible, the robustness of my results suggests that their potential effects are of second-order importance. However, there are several assumptions implicit in the analysis whose impact on the results is not easily ascertained. First, I assumed a specific distribution for the unobservable portion of utility. The selection of probit or logit to estimate the effects of interest responds more to convention than to economic rationale. One could wonder whether I would still find that there is state dependence in hospital choices of patients under other (possibly "non-standard") distributional assumptions. Second, I have assumed that switching costs are symmetric: the utility premium from sticking to the previous choice is independent of the hospital visited in period 1. However, it is possible that some hospitals induce more loyalty from patients than others. Third, the results previously discussed are informative about the mean effect of state dependence, but they provide limited information about heterogeneity in the impact of previous choices on current behavior. However, it is not clear which modeling assumptions are appropriate in this setting. Fourth, I have assumed that utility is linear in its different components. I may need to explore the effect of non-separability of preferences on measured state dependence. Fifth, I have assumed that patients are myopic when choosing hospitals: they do not consider the implications of their current choices for future episodes of care. The question remains whether conclusions change once I allow for forward looking behavior. In summary, although the main source of identification in my setting is evident, there are several modeling choices that could affect the results. In the next section, I address this concern and show evidence of state dependence using a nonparametric model of dynamic discrete choice.

1.5.4. Nonparametric analysis

In this section, I quantify the extent of state dependence in hospital choices of patients using the nonparametric approach proposed by Torgovitsky (2016), adapted to the particulars of my case study. Here, I outline the basic framework and present the main results; additional details can be found in Appendix B.

There are N patients and three periods ($t = 0,1,2$). All patients choose a hospital in periods 0 and 2, and some patients also have to choose a hospital in period 1. There are J hospitals; however, hospital 1 is not available in period 1 and therefore cannot be chosen by patients who need hospital care in that period. If $\mathcal{J}_t$ denotes the choice set in period t, then $\mathcal{J}_0 = \mathcal{J}_2 = \{1,...,J\}$ and $\mathcal{J}_1 = \{2,...,J\}$. This is the setting described in the previous subsection: hospital 1 is the hospital shuttered by Hurricane Sandy and the same definitions of time periods apply. Each patient has a type determined by:

1. Her choice in period $t = 0$, denoted h_0.

2. Her potential choices in each period $t \geq 1$ in which she has to choose a hospital: $h_t(k)$ denotes the alternative the patient would choose in period t had her previous choice counterfactually been k. For a patient who has to choose a hospital in period 1, $h_1(k)$ is the alternative the patient would choose in period 1 had her period 0 choice counterfactually been k, while $h_2(k)$ is the alternative the patient would choose in period 2 had her period 1 choice counterfactually been k. For a patient who does not have to choose a hospital in period 1, $h_2(k)$ is the alternative the patient would choose in period 2 had her period 0 choice counterfactually been k. There are J counterfactual choices $\{h_t(k)\}_{k=1}^{J}$ in period t. It is worth noting that the model places no restrictions on the temporal dependence of potential choices across periods.

3. Whether she needed hospital care in $t = 1$, indicated by $z \in \{0,1\}$.

4. Other variables that indicate time-varying or time-invariant characteristics of the patient, captured by vector $x \in \mathcal{X}$. I restrict attention to discrete covariates, so the set $\mathcal{X}$ is discrete.

In what follows, $\mathcal{H}$ denotes the set of all possible types, h denotes a generic type in $\mathcal{H}$ and $h_t = [h_t(1),...,h_t(J)]$ denotes the corresponding set of potential choices in period t.

Let y_t denote the hospital chosen by a given patient in period t. A patient in my data is associated with a vector of observables $y = \{y_0, y_1, y_2, z, x\}$. I observe the actual hospital choices of patients and their characteristics, but not their types. However, types and observables are related. In particular, a patient with type h makes the following choices:

$$y_0 = h_0 \tag{1.2a}$$

$$\text{For } z = 1 : y_1 = h_1(y_0) \text{ and } y_2 = h_2(y_1) \tag{1.2b}$$

$$\text{For } z = 0 : y_2 = h_2(y_0) \tag{1.2c}$$

Going forward, $\mathcal{Y}$ is the set of all possible vectors of observables, y denotes a generic element of $\mathcal{Y}$, and $y(h) \in \mathcal{Y}$ is the vector of observables generated by a patient with type h through relationships (1.2a)-(1.2c). Note that while type h generates a unique vector of observables $y(h) = \bar{y}$, the latter could have been generated by more than one type.

The data is given by $P : \mathcal{Y} \to [0,1]$, where P is a probability mass function (p.m.f.) with support in $\mathcal{Y}$. In other words, for each possible vector $y \in \mathcal{Y}$, I know the proportion of patients with those observables. For now, I assume that the distribution of observables is known by the researcher.

For any patient, only one element of the vector of potential choices in a given period is known to the researcher. The remaining potential choices are counterfactual. This framework is essentially the potential outcomes framework used in policy evaluation applied to a dynamic discrete choice setting. It postulates a set of potential choices that could be observed in alternative states of the world, with the state variable defined by the alternative chosen in the previous choice occasion.

The notion of state dependence arises naturally in this setting, as a patient might make different choices in different counterfactual scenarios. A patient with type h exhibits state dependence in period $t > 0$ if there are two previous counterfactual choices $k, j \in \{1, ..., J\}$ for which the corresponding period t potential choices are different: $h_t(k) \neq h_t(j)$. Let $sd_t(h)$ be an indicator of whether a patient with type h exhibits state dependence in period t. Identifying state dependence at the individual level requires information on at least two potential choices in a period. As explained above, the data reveals only one component of the vector of potential choices in a period. Therefore, whether there is state dependence at the individual level is untestable. However, we might be able to learn something about the extent of state dependence in the population.

The primitive of the model is the distribution of types in the population. With this distribution, I can define various measures of state dependence. In particular, I want to learn about the proportion of patients who exhibit state dependence in period 2 (which I refer to generically as θ). Then, I view this feature of the distribution of types in the population as the object of interest. If the distribution of types is given by the probability mass function $\mathbb{P}$, then the corresponding parameter is:

$$\theta(\mathbb{P}) = \sum_{h \in \mathcal{H}} \mathbb{P}(h) sd_2(h) \tag{1.3}$$

which is linear in the probabilities of types.

The problem is that the distribution of types is not known by the researcher. Let W be the set of admissible distributions. For the moment, we take W as given. This set contains all the distributions that the researcher considers could have generated the data, based on the assumptions about the data generating process that s/he is willing to maintain. We only consider distributions in W and any such distributions could in principle be the true one. In general, there are many admissible distributions and therefore the parameter of interest is partially identified. The identified set for the parameter of interest is:

$$\Theta(W) = \{\theta(\mathbb{P}) : \mathbb{P} \in W\} = [\theta_l(W), \theta_u(W)] \tag{1.4}$$

In other words, to recover the identified set we need to compute the proportion of patients with

Table 1.6: Period t latent choices, J=2

		$h_t(2)$		
		1	2	
$h_t(1)$	1	$\mathbb{P}_{11}$	$\mathbb{P}_{12}$	$\mathbb{P}_{1*}$
	2	$\mathbb{P}_{21}$	$\mathbb{P}_{22}$	$\mathbb{P}_{2*}$
		$\mathbb{P}_{*1}$	$\mathbb{P}_{*2}$	

Notes: This table is a 2x2 contingency table associated with potential choices in period t. The columns refer to the choices in the state in which the patient chose hospital 2 in $t-1$, while the rows refer to the choices in the state in which the patient chose hospital 1 in $t-1$. $\mathbb{P}_{j*}$ denotes the marginal probability $\mathbb{P}[h_t(1)=j]$, $\mathbb{P}_{*j}$ denotes the marginal probability $\mathbb{P}[h_t(2)=j]$, and P_{ij} denotes the probability $\mathbb{P}[h_t(1)=i, h_t(2)=j]$.

state dependence in period 2 under each admissible distribution using Equation (1.3). The last equality in Equation (1.4) states that the identified set is a closed interval with extreme points $\theta_l(W)$ and $\theta_u(W)$. This characterization of the identified set only holds under certain conditions on $\theta(\mathbb{P})$ and W, which are satisfied in my setting. Therefore, to characterize the identified set I only need to find the admissible distributions with the lowest and highest state dependence:

$$\theta_l(W) = \min_{\mathbb{P}\in W} \sum_{h\in\mathcal{H}} \mathbb{P}(h) sd_2(h) \tag{1.5}$$

$$\theta_u(W) = \max_{\mathbb{P}\in W} \sum_{h\in\mathcal{H}} \mathbb{P}(h) sd_2(h) \tag{1.6}$$

Before discussing the restrictions that define the set of admissible distributions and the computational approach to recover the identified set, let me briefly discuss the two main identification issues we need to deal with.

First, from ordinary (non-experimental) data on patients' hospital choices, I can recover $P[y_t = v|y_{t-1} = w] = \mathbb{P}[h_t(w) = v|y_{t-1} = w]$ for $v \in \mathcal{J}_t, w \in \mathcal{J}_{t-1}$. However, this object differs from $\mathbb{P}[h_t(w) = v]$ due to persistent unobserved heterogeneity. Therefore, the marginal distributions of potential choices in a given period are generally not identified from observational data.

Even if I deal with selection bias (for example, by using experimental data), there is a second identification issue: it is generally not possible to recover the joint distribution of potential choices from the marginal distributions (Heckman et al., 1997). The problem is then to bound an unknown joint distribution from known marginal distributions. This can be illustrated for the case of two alternatives (which I denote 1 and 2) using a 2x2 contingency table, following Heckman et al. (1997). For this example, I focus on the potential choices in period t. In Table 1.6, the columns refer to the choices in the state in which the patient chose hospital 2 in $t-1$, while the rows refer to the choices in the state in which the patient chose hospital 1 in $t-1$. If I observed each patient in both states of the world, I could fill in the table and recover the full distribution. From the full distribution, I could determine the proportion of patients with state dependence in period t[19]. With experimental data, I would be able to estimate the totals for each row and column but I cannot recover the values of particular cells without further assumptions. This issue stresses the fact that it is not possible to point identify the extent of state dependence even with experimental data[20]. This notion seems to have been overlooked by many empirical studies of persistence in consumer choice that focus almost exclusively on dealing with the selection bias issue[21].

The set of admissible distributions consists of those p.m.f. over $\mathcal{H}$ that are compatible with the observed data and the institutional setting, and that satisfy the additional identifying assumptions that I impose. Let $\mathbb{P}$ denote a generic distribution of types[22]. Then, $\mathbb{P}$ is an admissible distribution if the following conditions are satisfied:

Assumption 1: $\forall \bar{y} \in \mathcal{Y}$, $\sum_{h:y(h)=\bar{y}} \mathbb{P}(h) = P(\bar{y})$

[19] In the example, patients with state dependence in t are those with potential choices $[h_t(1) = 1, h_t(2) = 2]$ or $[h_t(1) = 2, h_t(2) = 1]$. These patients make different choices in different counterfactual scenarios (which are defined by the choice made in the previous episode).

[20] Although knowledge of the marginals is not enough to point identify state dependence, it might be enough to put non-trivial bounds on it. In the example, using the Frechet-Hoeffding bounds, we have:

$$\max\{\mathbb{P}_{1*} + \mathbb{P}_{*2} - 1, 0\} \leq \mathbb{P}_{12} \leq \min\{\mathbb{P}_{1*}, \mathbb{P}_{*2}\}$$

.

[21] It should be noted that dealing with selection bias is in general quite complicated.

[22] In particular, $\mathbb{P}$ satisfies the following conditions: 1) $\mathbb{P}(h) \in [0,1], \forall h \in \mathcal{H}$; 2) $\sum_{h \in \mathcal{H}} \mathbb{P}(h) = 1$

This restriction states that the probability mass assigned to types that generate certain observables $\bar{y}$ must be equal to the proportion of patients in the data with those observables.

Assumption 2: $\forall j \in \{1,..,J\}, h_1(j) \neq 1$

This restriction states that hospital 1 cannot be chosen in period 1 by patients who require hospital care during that period. Then, $\mathbb{P}$ assigns probability 0 to any type that does not satisfy the condition above.

Assumption 3: $\forall j,k,m \in \{1,..,J\}$, $h_t(k) = j \Rightarrow h_t(j) = j$ and $h_t(k) = j \Rightarrow h_t(m) \in \{m, j\}$

I restrict the set of admissible distributions to those p.m.f. that assign probability 0 to any type that does not satisfy the conditions above. There are two assumptions implicit in Assumption 3: 1) Monotone treatment response: if a patient would choose hospital A had she chosen another hospital B in the previous episode, then she would also choose hospital A had she chosen hospital A in the previous choice occasion; 2) No partial compatibility: purchases made in different episodes are compatible only if the same alternative is chosen in both occasions. Assumption 3 would hold, for example, if choices are determined by utility maximization and the utility a patient receives from choosing alternative j in period t depends on: 1) Characteristics of the patient and the hospital in that period; 2) Whether the patient chose alternative j in the previous choice occasion, in which case she receives a non-negative utility premium (but there is no utility premium from switching hospitals). Note that the standard logit model of demand (see, for example, Equation 1.1 in Section 1.3) satisfies this assumption. A more detailed discussion of Assumption 3 can be found in Appendix B.

Under Assumptions 1-3[23], the identified set for the proportion of patients with state dependence in period 2 is the interval $[0,1]$: I can find two distributions of types that attribute all the observed persistence to unobserved heterogeneity and state dependence, respectively. In order to get more informative results, I need to impose additional structure. In particular, I translate the identifying assumption discussed in the previous subsection into the following restriction on the set of admissible distributions:

Assumption 4: $\forall j,k \in \{1,..,J\}$ $\forall x \in \mathcal{X}$, $\mathbb{P}[h_2(k) = j|z = 0,x] = \mathbb{P}[h_2(k) = j|z = 1,x]$

This is a conditional independence assumption. It states that the potential choices in period 2 are independent of the treatment status indicator z, conditional on the other covariates x. In other words, the distribution of preferences in period 2 is the same for patients who required hospital care in period 1 and for patients who did not, conditional on other covariates.

This was the main identifying assumption for the analysis in Section 1.5.3. Here, the meaning of this assumption is more clear. If patients in the treatment and control groups have different loyalty states in period 2, then their choices reveal different elements of the vector of potential choices in that period. If the distribution of potential choices is the same for both groups, then I can combine the information obtained from each group to learn about the joint distribution of potential choices.

Assumptions 1 through 4 determine the set of distributions W that I need to consider for identification. As discussed above, I could recover the identified set by computing the parameter of interest for each admissible distribution. However, this approach is not very practical. While an analytical characterization seems too complicated, we can exploit the structure of the problem to characterize the identified set in a computationally simple way. As mentioned before, the identified

[23] In principle, the only assumptions that need to be always considered are Assumptions 1 and 2. As I discuss later, I need to impose Assumption 3 ex-ante to accommodate certain practical limitations. However, whenever possible, we might want to treat Assumption 3 as a regular identifying assumption and evaluate its power independently of the other restrictions.

set is a closed interval. Therefore, I only need to recover the lower and upper bounds on the proportion of patients with state dependence in period 2 to fully characterize the identified set. Consider the problem of recovering the lower bound (the same reasoning applies to the problem of recovering the upper bound). As Equation (1.5) shows, the problem is to find the admissible distribution (probabilities of types) that minimizes the parameter of interest. The latter is linear in the probabilities of types. The set of admissible distributions is defined by Assumptions 1 through 4, which impose linear restrictions on the probabilities of types. Therefore, the lower and upper bounds on the parameter of interest can be found by solving a set of linear programming problems. The full description of the optimization problem can be found in Appendix B. Note that I can recover the identified set under different characterizations of W in a straightforward way: I just need to re-solve the optimization problem under different combinations of assumptions.

To implement this approach, I take certain steps to accommodate practical limitations. The main issue is the size of the set of types: the number of possible potential choices in any period $t \geq 1$ when the cardinality of the choice set is J is J^J, so the number of possible types (without any covariates) is J^{JT+1} if $t = 0,..,T$. To keep the dimensions of the problem to manageable proportions, I make the following modeling choices: 1) I only consider three periods ($T = 2$); 2) I only consider patients who visit the affected hospital or any of its main competitors in any period in which they need hospital care (then, I use sample #3 for the analysis); 3) I impose Assumption 3 ex-ante. Without doing this, the dimensions of the optimization problem would be prohibitive in my application. However, I stress that Assumption 3 has economic content beyond its usefulness to reduce the dimensionality of the problem.

For each case study, I compute the identified set under Assumptions 1 through 4. The results are shown in Table 1.7. The data and the restrictions that I impose on the distribution of preferences allow me to conclude that at least 9.1% of patients in the sample for Coney Island Hospital exhibit state dependence in period 2. In other words, there is no distribution of patient types consistent with Assumptions 1 through 4, such that the proportion of patients with state dependence in period 2 is

Table 1.7: Proportion of patients who exhibit state dependence in period 2: lower bound

Coney Island Hospital	Bellevue	NYU Langone
0.091	0.092	0.129

Notes: This table shows, for each case study, the lower bound on the proportion of patients who exhibit state dependence in period 2 under Assumptions 1 through 4.

lower than 9.1%. In the case of Bellevue and NYU Langone, the lower bounds on the proportion of patients with state dependence are 9.2% and 12.9%, respectively. In all three cases, the upper bound is equal to 1: I cannot reject the hypothesis that all patients exhibit state dependence in period 2 under Assumptions 1 through 4.

The interpretation of the results is the following. I consider the case of Coney Island Hospital to fix ideas. If I knew the distribution of types in the population, then I would be able to compute the proportion of patients who exhibit state dependence in period 2. However, I do not know what the true distribution is. Therefore, I need to consider all distributions that are consistent with the data. I cannot get informative results from the data alone, so I need to impose additional structure. One possibility would be to specify a fully parametric model of patient behavior (for example, a random utility model), map the full structure of the model into the potential choices framework, estimate the parameters of the model, and use these estimates to compute the proportion of patients with state dependence in period 2. However, it is difficult to assess the sources of identification with this approach and therefore it is not clear how measured state dependence would change under different modeling assumptions. The results obtained using the nonparametric approach indicate that no model with a richer structure will produce less than 9.1% of patients with state dependence in period 2, as long as it is consistent with the identifying assumptions imposed. Therefore, I can reject the hypothesis that there is no state dependence under transparent assumptions: the finding of state dependence is not an artifact of particular parametric assumptions but is supported by a more fundamental source of identification (Torgovitsky, 2016). I gain transparency in results at the cost of set identification: there are many values of the parameter of interest that are compatible with the data and the identifying assumptions that I impose.

In summary, this approach allows me to provide robust evidence of the presence of state dependence in hospital choices of patients. I can rule out that persistence in patient behavior is only due to unobserved heterogeneity under minimal assumptions about the data generating process. On the other hand, the identified set is quite wide, so we need to impose additional structure if we want to get more informative results. Therefore, I see the approach discussed in this section as complementary to parametric models of consumer choice: we use these richer models to obtain more meaningful conclusions, but we do this having already shown that the finding of state dependence is not an artifact of the parametric assumptions we impose.

In the previous analysis, I recovered the identified set assuming that the distribution of observables was known, without accounting for sampling error. To conduct statistical inference on the identified set, I follow the strategy proposed by Torgovitsky (2016). The idea is to express the potential choices model as a moment inequalities model. To characterize the identified set, we define a criterion function that penalizes deviations from the identifying assumptions. Then, a sample analog of this criterion function is used as the basis for statistical inference. Specifically, I use it as a test statistic for the null hypothesis that the parameter of interest takes a specific value. To recover the distribution of the test statistic, I use the subsampling approach of Romano and Shaikh (2008). This approach requires solving a linearly-constrained quadratic program for many subsamples. Given that I am interested in bounding the extent of state dependence away from zero, I test the hypothesis that the proportion of patients with state dependence in period 2 is equal to zero. Even though the test is conservative, as pointed out by Torgovitsky (2016), the hypothesis is strongly rejected at usual significance levels.

1.6. The impact of state dependence on health outcomes

In previous sections, I presented evidence that past hospital choices of a patient influence her current choice. If health or other shocks made a patient gravitate towards a certain hospital in the past, then the patient is more likely to choose that same facility for her current medical needs than an otherwise similar patient. Therefore, absent state dependence, demand patterns would be different.

In this section, I analyze the impact of state dependence on health outcomes. Lock-in might prevent patients from re-optimizing after a change in the environment. In particular, lock-in might prevent a patient from switching to hospitals that are more suitable than her previous choice at treating her current medical condition. Then, absent state dependence, patients would choose hospitals that produce better health outcomes. However, this is an empirical matter. It is possible that, absent state dependence, patients would switch to lower quality facilities, in which case the impact of state dependence on health outcomes is reversed.

I study this in the context of hospital choice for heart surgery (Coronary artery bypass grafting, CABG). I estimate a model of hospital choice that quantifies patients' preferences for hospital attributes and incorporates state dependence. I use the estimates of the model to compare expected mortality in the actual state of the world and in a counterfactual scenario in which there is no state dependence. The main question I want to answer is: Absent state dependence, how many more CABG patients would have survived?

There are two reasons why I focus on a particular medical condition rather than doing a more general analysis. First, risk adjustment and estimation of hospital quality are better defined when they are disease specific. Second, quality measures are difficult to compare across different procedures; therefore, it is hard to find a single measure of health outcomes that is appropriate for a wide range of medical conditions. I focus on CABG surgery for various reasons (Gaynor et al., 2016). First, it is a common procedure. Second, it is mostly an elective intervention. Third, risk-adjusted mortality is a well accepted quality metric for CABG surgery. Fourth, there are many studies that analyze hospital choices of patients for CABG surgery, which provide me with a benchmark for assessing any effects I find.

In the analysis, I hold hospital quality fixed as observed in the data: I ignore any feedback between patients' demand patterns and quality choices of hospitals. In order to account for feedback, I would need a formal model of hospital behavior, which is beyond the scope of this study. Therefore, in assessing the impact of state dependence on health outcomes, I only consider the direct effect that takes into account how it affects the allocation of patients to hospitals.

Ideally, I would use the natural experiment discussed in Section 1.5 for this exercise. However, this is not feasible due to the size of the resulting sample: the number of patients in the treatment group who are hospitalized for CABG is very small[24]. As a result, I must rely on observational data to identify the parameters of interest, at the cost of less clean identification than in the previous section. However, I showed in previous sections that past hospital choices of patients have a causal impact on current behavior. Moreover, the identification of state dependence in those settings did not rely on arbitrary parametric assumptions. This lends credibility to the conclusions obtained from models with a richer structure like the one used in this section.

1.6.1. Framework

A patient has $T \geq 1$ episodes of care, where episode T corresponds to CABG surgery. Even though patients seek treatment for the same condition in the last episode, they might have received hospital care for different medical conditions in the past. The utility that a patient obtains from choosing a particular hospital for CABG surgery is a function of observed patient and hospital characteristics (which capture the quality of the match between them), the loyalty state of the patient (which summarizes the history of her past hospital choices), and factors unobserved by the researcher. The patient chooses the hospital in the choice set $\mathcal{J} = \{1, .., J\}$ that gives her the highest utility. In what follows, $\mathbb{P}^{SD}_{ij}$ is the ex-ante probability that patient i chooses hospital j in the actual state of the

[24] Sample sizes are also small for other medical conditions studied in the literature.

world (taking the loyalty state of the patient as given), and $\mathbb{P}_{ij}^{NSD}$ is the choice probability in the counterfactual scenario where there is no state dependence.

The ultimate goal of the analysis is to evaluate the impact of state dependence on the ex-ante probability of death following CABG surgery. Let $Y_i = 1$ denote that patient i dies within 30 days of CABG surgery. For each hospital, there is a potential outcome Y_{ij} that denotes the outcome that would have been observed had patient i counterfactually chosen hospital j for CABG surgery. Potential outcomes are related to observed variables as follows:

$$Y_i = \sum_{j=1}^{J} Y_{ij} h_{ij} \tag{1.7}$$

where $h_{ij} = 1$ indicates that patient i chose hospital j.

Following the notation introduced above, the ex-ante expected mortality of patient i in scenario $K \in \{SD, NSD\}$ is given by:

$$EM_i^K = \sum_{j=1}^{J} \mathbb{P}_{ij}^K E(Y_{ij}) \tag{1.8}$$

This is the expected probability of death taking as given the loyalty state of the patient and before the choice and mortality errors are realized. The impact of state dependence on mortality across all patients who undergo CABG surgery is:

$$\Delta EM = \sum_i [EM_i^{SD} - EM_i^{NSD}] = \sum_i \sum_j [\mathbb{P}_{ij}^{SD} - \mathbb{P}_{ij}^{NSD}] E(Y_{ij}) \tag{1.9}$$

As discussed before, hospital-specific mortality is the same in both scenarios and therefore the difference in health outcomes is only due to different sorting of patients across hospitals. Then, the elements that I need to recover are: 1) The choice probabilities that govern the data generating process; 2) The choice probabilities in the scenario where there is no state dependence, and; 3) Hospital-specific mortality. The choice probabilities are computed from the estimates of a hospital demand model, while I compute hospital quality using patient discharge and mortality data. After describing the data used for the analysis, I discuss these parts of the model in turn.

1.6.2. Data

I construct the sample for the analysis as follows. First, I identify all patients who had CABG surgery in a New York hospital during the period from 2013 to 2015. I identify CABG cases based on the procedures listed in the discharge record of the patient. I then exclude patients who had valve surgery or another major cardiac surgical procedure during the same hospital stay, and patients who were transferred to the hospital from another health care facility. Second, I exclude CABG surgeries performed during an emergency episode. This way, I only study hospital choices of patients who are able to evaluate the different alternatives available. Third, I exclude a small number of patients who are more than 85 years old and patients with insurance other than Medicare, Medicaid, or private insurance. Fourth, I assume that the choice set of a patient consists of all hospitals in the market for CABG surgery (see below) that are within 100 miles of her home (25 miles for patients in New York City). Then, I drop patients who went to a hospital outside this area. The final sample consists of 7,509 patients.

For each patient, I construct the history of hospital visits before the CABG episode. To avoid picking-up past hospital use directly related to the latter, I consider all visits for cardiac care to the hospital chosen for CABG during the 60 days prior to the surgery as part of the CABG episode. The main loyalty variable is an indicator for whether the hospital was used in the previous episode of care.

The set of hospitals that offer CABG surgery is limited: 39 hospitals out of a total of 264 facilities in the State of New York offered this service during the period from 2013 to 2015[25]. Table 1.8 shows summary statistics for these hospitals.

[25]Given my sample restrictions, there are two hospitals that were not chosen by any patient in the resulting dataset. Therefore, in the analysis, the set of hospitals in the market for CABG surgery contains 37 hospitals.

Table 1.8: Hospitals that offer CABG surgery

	Number	Proportion
Public	5	0.14
Teaching	23	0.62
Pediatric surgery	11	0.30
Heart transplant, adult	6	0.16
Heart transplant, pediatric	2	0.05
Total	37	1

Notes: This table shows summary statistics of hospitals in the market for CABG surgery during the period from 2013 to 2015. The last three variables refer to other cardiac procedures performed by a hospital.

Table 1.9 reports summary statistics of the CABG episodes in the working sample. Almost 80% of patients are male. The average patient is around 66 years old, and more than half of the patients are covered by Medicare. On average, a patient had 4.6 episodes before the CABG surgery. The observed 30-day mortality rate is 0.99%. The average distance to the chosen hospital is 17.5 miles and 41% of patients go the closest hospital. There are 14 hospitals in the average patient's choice set.

In 2,905 cases (39% of the total), the hospital used in the previous episode is in the choice set of the patient. On average, this hospital was used in 59% of previous episodes and in 82% of previous visits to hospitals in the choice set; in 63% of the cases, it is the only hospital in the choice set that was used before. The average time elapsed since the last episode is 2.2 years. Past hospital use is a strong predictor of the current hospital choice of the patient: conditional on having used a hospital in the previous episode, the repurchase probability is 0.62.

1.6.3. Hospital quality

To compute hospital-specific mortality, I use patient level data of all hospitalizations for CABG surgery during the period from 2013 to 2015. Following Chandra et al. (2016), I regress an indicator for mortality within 30 days of discharge on characteristics of the patient (age/race/sex interactions),

Table 1.9: Summary statistics of CABG patients

Age	66.4
Female	0.22
White	0.70
NYC	0.36
Private insurance	0.32
Medicare	0.55
Medicaid	0.13
Charlson = 0	0.23
Charlson = 1	0.31
Charlson = 2	0.47
Episodes before CABG	4.63
Died (%)	0.99
Distance	17.6
Closest	0.41
Observations	7,509

Notes: This table shows summary statistics of CABG patients in the working sample.

indicators for being hospitalized for selected conditions in the year previous to CABG surgery[26], and hospital fixed effects. The hospital fixed effects are the risk-adjusted mortality rate estimates for hospitals, which are the hospital quality measures used in the analysis. A regression of the estimated fixed effects on hospital attributes (not reported) suggests that mortality is negatively associated with provision of sophisticated services such as pediatric cardiac surgery, but there are no systematic differences between teaching and non-teaching hospitals.

The State of New York computes risk-adjusted mortality rates for CABG surgery by hospital. However, as pointed out by Chandra et al. (2016), these rates are computed as the ratio between observed and expected mortality, so they are not equivalent to the hospital fixed effects that I estimate. However, I find that these rates and my estimates are highly correlated (coefficient of correlation is 0.81).

[26]These conditions are: acute myocardial infarction, diabetes, diabetes with complications, hemiplegia/paraplegia, renal disease, cancer, metastatic cancer, mild liver disease, moderate/severe liver disease, AIDS, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, COPD, rheumatoid disease, and peptic ulcer.

1.6.4. Hospital demand

The key factors that affect hospital choices of patients are distance from home to the hospital, quality of care, and previous experience with the facility. I assume that the utility that patient i obtains from choosing hospital j for CABG surgery is:

$$u_{ij} = -\alpha_i d_{ij} + \beta_i q_j + \gamma_i \mathbb{I}(L_i = j) + \sum_k \delta_i w_{jk} + \xi_j + \varepsilon_{ij} \tag{1.10}$$

where d_{ij} is the travel distance from patient i's home to hospital j, q_j is the risk-adjusted mortality rate of hospital j, L_i is the loyalty state of the patient, w_{jk} is the value of attribute k (for example, an indicator for public hospital) for hospital j, ξ_j captures unobserved features of the facility, and ε_{ij} reflects idiosyncratic preferences of the patient for the hospital. I assume that the latter is iid distributed according to a Type 1 extreme value distribution.

I assume that there is no outside option. I define the alternative specific constants relative to New York-Presbyterian Columbia University Medical Center, which consistently ranks as one of the top cardiology and heart surgery hospitals in the country according to U.S. News. I refer to the model's parameters other than γ as θ. I allow for observed and unobserved heterogeneity in preferences for hospital characteristics. In particular, the coefficients on distance, mortality, and previous use are a function of patient characteristics (demographics and severity) and unobserved factors. More precisely, I assume that:

$$\alpha_i = \bar{\alpha} + \sum_r \alpha_r^o z_{ir} + \alpha^u \mu_i \tag{1.11}$$

$$\beta_i = \bar{\beta} + \sum_r \beta_r^o z_{ir} + \beta^u \nu_i \tag{1.12}$$

$$\gamma_i = \bar{\gamma} + \sum_r \gamma_r^o z_{ir} + \gamma^u \eta_i \tag{1.13}$$

where z_{ir} is the value of (observed) attribute r for patient i, and $(\mu_i, \nu_i, \eta_i) \sim N(0, I)$ represent idiosyncratic preferences of consumer i for distance, quality, and loyalty, respectively.

This specification of patient utility is similar to past work, although it allows for richer heterogeneity. The third term in Equation 1.10 captures the impact of loyalty on current utility. Specifically, the patient gets a utility premium γ from choosing the hospital to which she is loyal. It can also be interpreted as an implied utility cost conditional on switching hospitals. This implies that for patient i to switch hospitals, she must prefer an alternative option by γ_i more than the hospital used in her previous episode (Handel, 2013).

One possibility is to model the hospital choice of a patient for all her episodes using Equation 1.10[27]. Then, the likelihood function at the patient level would be computed for the *sequence* of hospital choices. This way, I can exploit changes in choice set across episodes for a given patient to learn about her preferences. However, I would need to assume that preferences are stable across episodes. This would imply, for example, that the disutility of travel of the patient is the same for CABG and for other types of episodes. That assumption seems too strong. Instead, I model the choice of hospital in episode T_i (CABG) alone.

The main identification concern is the endogeneity of the previous choice variable. The fact that a patient chose a given hospital in the past indicates that she might have strong unobserved preferences for that facility; if these preferences are persistent across episodes and are not properly accounted for, then the previous use variable is positively correlated with the econometric error term. Then, a positive value of γ might be capturing both the effect of state dependence and persistent unobserved heterogeneity.

Given that the set of hospitals that offer CABG surgery is small, there are many consumers

[27]The different covariates would need to be indexed by t, as characteristics of patients and hospitals might change across episodes (for example, the patient might need hospital care for different diagnoses over time).

who have not developed an attachment to any particular hospital in the market for this procedure. For these consumers, state dependence does not affect preferences between hospitals. Then, the parameters in θ are identified from the choices of these "new" patients. This identification strategy has been used in the literature (see, for example, Handel (2013) and Luco (2016)). The state dependence parameter is then identified by the repurchase behavior of patients who are loyal to a particular facility.

I estimate the model by Maximum Simulated Likelihood (Train, 2009). From the estimates of the model, I recover two sets of choice probabilities: (1) $P_{ij}(\theta,\gamma)$ is the ex-ante probability that patient i chooses hospital j in the actual state of the world: these are the predicted probabilities based on imposing the estimated coefficients of the utility function on the actual data; (2) $P_{ij}(\theta,0)$ is the choice probability when the state dependence parameter is set equal to zero.

1.6.5. Results

The estimates from the demand model (see Table 1.10) indicate that patients dislike traveling for hospital care and value hospital quality. On average, a patient is willing to travel 12 additional miles for getting access to a hospital with one percentage point lower mortality. The impact of state dependence on patients' choices is large: if a hospital was used by the patient in her previous episode, then the probability of choosing that facility is about 3.4 times higher than what would be expected based on other covariates.

The impact of state dependence on health outcomes is a function of the spatial and quality configuration of local markets (the geographic distribution and composition of the patient population in the local area, and the locations and qualities of hospitals). Therefore, although the effects of state dependence on choices are substantial, the effects on health outcomes might not be as large. Moreover, it is not clear that eliminating the frictions driving state dependence will make patients switch to better hospitals in terms of quality.

I consider the impact on health outcomes of a number of thought experiments. I compare the

Table 1.10: Estimates of the hospital demand model

		Estimate	Std Error
Used in previous episode	Constant	1.850	0.219
	Female	0.124	0.199
	No white	-0.012	0.183
	Charlson Index = 1	-0.002	0.233
	Charlson Index = 2	0.533	0.215
	Medicare	-0.067	0.184
	Medicaid	0.432	0.283
	Std dev	2.656	0.224
Distance	Constant	0.140	0.007
	NYC	0.095	0.008
	Female	0.000	0.005
	No white	-0.008	0.006
	Charlson Index = 1	0.004	0.006
	Charlson Index = 2	-0.005	0.005
	Medicare	0.000	0.005
	Medicaid	0.014	0.008
	Std dev	0.045	0.004
Mortality	Constant	-1.867	0.172
	Female	0.053	0.043
	No white	0.008	0.040
	Charlson Index = 1	0.041	0.051
	Charlson Index = 2	-0.029	0.049
	Medicare	0.087	0.046
	Medicaid	0.187	0.058
	Std dev	0.130	0.094

Notes: This table shows the estimates from the hospital choice model. In addition to the coefficients shown in the table, the model includes hospital fixed effects and interactions between indicators for the type of insurance of the patient and an indicator for public hospitals.

Table 1.11: Ex-ante expected mortality relative to the actual state of the world

Panel A: Impact on mortality from reducing state dependence

γ/γ^0	0.8	0.6	0.4	0.2	0
ΔEM	-0.129	-0.291	-0.490	-0.736	-1.007

Panel B: Impact on mortality from reducing disutility from travel

α/α^0	0.8	0.6	0.4	0.2	0
ΔEM	-0.442	-0.950	-1.543	-2.204	-2.776

Notes: This table shows the change in expected mortality (in terms of number of deaths) under a range of counterfactual scenarios, relative to the actual state of the world in which preferences are given by the estimates reported in Table 1.10. The effects reported correspond to patients for whom the hospital used in the previous episode is in the choice set. In Panel A, different columns correspond to different counterfactual scenarios where the state dependence coefficient has been set equal to a given fraction (0.8, 0.6, 0.4, 0.2, 0) of the baseline estimate. In Panel B, different columns correspond to different counterfactual scenarios where the disutility of distance has been set equal to a given fraction (0.8, 0.6, 0.4, 0.2, 0) of the baseline estimate. γ^0 and α^0 are the baseline coefficients, while γ and α are the counterfactual coefficients.

ex-ante expected mortality under different counterfactual scenarios, relative to the actual state of the world. In each scenario, the state dependence parameter is set to a fraction of the baseline estimate: the lower the fraction, the less contaminated by state dependence the hospital choices of patients. We can think of these counterfactuals as illustrating the impact of policies that are partially effective in dealing with the sources of state dependence, or as capturing the possibility that some residual unobserved heterogeneity is picked up by the state dependence parameter. The results are reported in Panel A of Table 1.11. To provide a reference point to analyze the magnitude of these effects, I consider changes in mortality relative to the baseline scenario from reducing patients' disutility from travel. The results are reported in Panel B of Table 1.11.

There are two points to be noted. First, I find that reducing the magnitude of the state dependence coefficient leads to reductions in ex-ante expected mortality. In a counterfactual world where there is no state dependence, one extra person is expected to survive, which implies a 3% reduction in mortality relative to the baseline scenario (there were 33 actual deaths). In other words, state dependence prevents a stronger allocation of patients to higher-quality hospitals. To provide a benchmark for assessing the magnitude of this reduction in mortality, consider the findings of Gaynor et al. (2016). They study the impact of a reform in the English National Health Service that removed constraints on patient choice. They find that CABG patients became more responsive to clinical quality: the reallocation to higher quality hospitals led to a 3% reduction in expected mortality. Second, the impact of eliminating state dependence is similar to the effect of reducing patients' preferences for proximity by 40%. Therefore, the allocative role of state dependence is significant, considering that distance is one of the main determinants of patients' hospital choices.

Following Handel (2013), it should be noted that the impact of state dependence on health outcomes that I find is specific to the setting considered: the direction or magnitude of the impact could be reversed for other patient populations (with different medical conditions) or for the same population in different markets.

1.7. Conclusion

When choosing a hospital, patients favor facilities they have used in the past. While the idea of patient loyalty seems widely accepted, there is no strong prior on whether unobserved heterogeneity or state dependence drive this behavior. These channels have different implications for hospital behavior, the long-run welfare effect of excluding a hospital from an insurer's network, and the design of policies to influence patient demand. To identify the sources of choice persistence in a credible way, I exploit quasi-exogenous shocks that induce a patient to try a new hospital: emergency hospitalizations and temporary hospital closures.

In the first case, I find that patients who visit a new hospital during an emergency hospitalization are more likely to continue using that same facility in subsequent episodes than observationally similar patients. This provides evidence of the presence of state dependence in hospital choices of patients. In cases where the emergency did not shift the loyalty state of the patient, repurchase rates are higher, which indicates that unobserved heterogeneity is also empirically relevant.

In the second case, I exploit the unexpected closures of three hospitals in New York City following Hurricane Sandy. I find that patients who needed hospital care during the time an affected hospital was closed for repairs were less likely to use the facility after its reopening than similar patients who did not have hospital visits during the unavailability window. Moreover, patients continued using the same hospital they visited during the time their usual facility was unavailable. The difference in behavior of these patients points to the presence of state dependence. To provide more credible evidence, I analyze patients' hospital choices using the nonparametric framework of Torgovitsky (2016). I am able to bound the proportion of patients who exhibit state dependence away from zero under a conditional independence restriction between the preferences of the patient and the timing of hospital visits, but without making parametric assumptions about the nature of preference heterogeneity. The higher transparency of this approach (relative to more traditional frameworks to analyze dynamic discrete choice) comes at the cost of set identification.

After showing that state dependence affects patients' choices, I look at its implications for health outcomes. In the context of hospital choice for heart surgery, I find that, absent state dependence, patients would switch to hospitals that are better at producing health outcomes. In particular, in this counterfactual scenario, ex-ante expected mortality would be 3% lower than in the actual state of the world.

There are interesting avenues for future research that originate from the results in this chapter. The first is to investigate the microeconomic fundamentals that drive state dependence in this setting. Quantifying switching costs created by hospital policies regarding the sharing of patient data would be particularly interesting given the diffusion of Electronic Health Records (EHR) during the last years and policy efforts to achieve interoperability of different systems.

It would also be interesting to quantify the contribution of state dependence in hospital choices of patients to inertia in consumers' choices of health insurance plan. This is particularly relevant given the diffusion of narrow networks after implementation of the Affordable Care Act.

Another interesting question is how state dependence distortions hospitals' incentives to invest in quality. For example, hospitals might display bargain-then-ripoff behavior using quality as the adjustment variable. In addition, the possibility of developing long-term relationships with patients might lead hospitals to overinvest in services that patients use early in their lives.

Chapter 2

Persistence in hospital choices of patients for childbirth

2.1. Introduction

In this chapter, I study persistence in hospital choices of patients for childbirth in the State of New York. Specifically, I analyze to what extent mothers use the same hospital for different pregnancies and what the sources of the observed persistence in behavior are. The advantage of analyzing childbirth is that the different episodes of a patient correspond to the same medical condition. However, there are changes in circumstances across pregnancies that might change patients' preferences for hospitals. In particular, I focus on mothers with pregnancies of different clinical complexity and mothers who change residence between pregnancies. On the downside, focusing on childbirth adds the complication that any observed persistence in hospital choices of patients might just reflect loyalty to doctors.

I start by presenting evidence of substantial persistence in hospital choices of patients for childbirth. The raw repurchase probability in the data is very high (close to 75%). Moreover, I find that, on average, patients are more likely to choose the hospital used in the previous pregnancy than observationally similar new mothers. To explore the possibility that loyalty to doctors drives the observed loyalty to hospitals, I compare cases based on whether the previous doctor continues working at the previous hospital. I find that there is substantial persistence in patients' hospital choices that is not associated to loyalty to doctors. In order to gauge the quantitative importance of the persistence found in the data in a more precise way, I then estimate a model of patients' hospital choices that incorporates inertia. Consistent with the descriptive part of the analysis, I find that previous use of a hospital is a strong predictor of the current hospital choice of a patient: having used a hospital in the past increases the probability of choosing that same hospital for the current childbirth 1.7 times above the baseline level. Moreover, previous use of a hospital is an important driver of heterogeneity in willingness-to-pay for a facility.

While in Chapter 1 I used emergencies and temporary hospital closures for identification, here I follow a more traditional identification strategy that relies on the existence of consumers who are new to the market (first-time mothers), whose decisions cannot be contaminated by inertia, and experienced consumers (mothers with previous pregnancies). This is similar to the approach followed by Handel (2013), Polyakova (2016), and Luco (2016), who compare cohorts of new consumers (who make active choices) and established consumers (who have a particular default[1] option).

This work is related to papers that study hospital choices of patients for childbirth. Examples include Phibbs et al. (1993), Ho and Pakes (2014), and Raval and Rosenbaum (2018). The paper most closely related is Raval and Rosenbaum (2017), which was developed independently from the current study. They analyze the determinants of hospital choices for childbirth of mothers in Florida and provide evidence that switching costs are important in that setting. They use a panel data fixed effects estimator to separate persistence in choice due to switching costs and persistence in choice due to unobserved preference heterogeneity. More specifically, they consider women who have three children and switch hospitals between their first and second births. For identification, they compare the hospital choices (for the third birth) of women who attended the same two hospitals for the first two births but in different order. They find that approximately 40% of the observed choice persistence reflects switching costs. The current work follows a different identification strategy to separate the sources of choice persistence and shows that mothers continue using the previous hospital even when there is a change in the choice environment.

The rest of the chapter is organized as follows. In Section 2.2, I describe the data used for the analysis. In Section 2.3, I present descriptive evidence of persistence in hospital choices of patients for childbirth. In Section 2.4, I quantify persistence in hospital choices of patients using a model of hospital demand. In Section 2.5, I provide concluding remarks.

[1] In my analysis, the default option does not refer to the alternative to which the consumer is assigned in case she does not make an active choice.

2.2. Data

I use hospital discharge data for the State of New York from 1995 to 2015 (see Section 1.2 of Chapter 1 for more details about this data). Specifically, I construct a panel dataset of all childbirths during this period. A record is identified as a childbirth if it contains a V27 diagnosis code (ICD-9). For each record, I observe a patient identifier that links births by the same mother over time, admission and discharge dates, hospital and physician identifiers, diagnosis and treatment information, and patient demographics (age, gender, race, zip code of residence). I assign to each diagnosis listed on the discharge record of the patient a severity rank (1 to 3, in increasing order of severity) based on the methodology of Ho and Pakes (2014). I then classify episodes into severity groups based on the ranking of the most severe listed diagnosis. For each episode, I consider the operating physician in the patient's record as the relevant doctor for the decision process: this assumption is sensible as women mostly care about the physician they expect will handle the delivery. The full sample contains 4,935,066 deliveries during the period from 1995 to 2015.

I define the choice set of a patient as the set of all hospitals that: 1) offered regular maternity services at the time of the birth, and; 2) are within 50 miles of the residence of the patient (25 miles for patients living in New York City). I infer provision of maternity services by aggregating deliveries to the hospital level and observing the number of deliveries that take place in each hospital. I consider that a hospital provides maternity services in any particular year if it performed more than 20 deliveries during that period. There are hospitals with a positive but low volume of deliveries in the discharge data that are not feasible options according to this definition. These cases seem to correspond to patients who live in remote areas or to emergency situations.

For the main analysis, I use a subsample that results from the following steps. First, I drop records with a missing patient identifier. This restriction eliminates 0.34% of the observations in the full sample. Second, I ignore cases where the hospital visited was not in the choice set of the patient. This restriction would eliminate, for example, a mother living in Buffalo and visiting a hospital in New York City. Third, I drop patients with missing or invalid values for key variables in at least one record and patients for which the patient identifier does not seem accurate based on inconsistent information about date of birth or time between deliveries. Finally, I exclude patients who had a pregnancy before age 15 and all deliveries after age 45, as I expect these patients to have idiosyncratic characteristics that make them gravitate towards particular hospitals. Taken together, the last three selection steps eliminate 8.5% (6.7%) of observations (patients) in the non-missing patient identifier sample. The resulting sample (which I refer to as the "working" sample) contains 4,498,651 deliveries for 3,280,132 different patients during the period from 1995 to 2015.

In the data, I observe all childbirths of a patient during the period from 1995 to 2015. Some patients might have had births prior to 1995 and therefore I do not observe their complete history of hospital choices. The "main" sample contains childbirths of women in the working sample who were no more than 18 years old in 1995, for whom it is likely that I observe the complete history of hospital choices for childbirth. The main sample contains 1,959,784 deliveries for 1,343,946 patients during the period from 1995 to 2015. In what follows, I refer to patients with no prior record of delivery (within the sample period) as first time mothers and to patients with at least one prior pregnancy as "experienced" mothers.

For each episode, I define the marginal probability of choosing a certain hospital as follows. I first group mothers into bins defined by unique combinations of zip code of residence, severity and year. Then, I aggregate episodes to the bin-hospital level and define the marginal probability that a patient from a certain bin chooses a hospital as the share of new mothers in that bin who choose that facility (see Section 1.4 in Chapter 1 for more details). In the preferred specification, I use first time mothers in the main sample to compute marginal probabilities. Therefore, I do not use hospital choices of mothers who were older than 18 in 1995 but did not have any childbirth experience in a New York hospital in the sample period to compute marginal probabilities. As I will focus on the period from 2005 to 2015, I expect the impact of this problem to be small. To further check that this restriction does not contaminate the results, I performed the same analysis using all first time mothers to compute marginal probabilities and the results (not reported) are very similar.

For the analysis of repurchase behavior, I focus on hospital choices of experienced mothers during the period from 2005 to 2015. I restrict attention to cases in which the hospital visited in the previous pregnancy (which I refer to as the previous hospital) is a feasible option for the patient. First, I drop cases in which the previous hospital is not in the market for maternity services at the time of admission. Second, I ignore cases where the previous hospital is not a practical alternative for the patient. Specifically, I ignore cases where the marginal probability of choosing the previous hospital is lower than 0.01.

I use data on hospital characteristics from the New York Department of Health, SPARCS and AHA. The State of New York has a system of regionalized perinatal services: with perinatal regionalization, hospitals receive a perinatal designation on the basis of the risk profile of patients they are capable to care for. There are four levels of perinatal centers (in increasing order of sophistication): Level 1, Level 2, Level 3, and Regional Perinatal Center; current designations became effective during the early 2000's. Hospitals with a given designation provide perinatal services corresponding to lower levels and in addition are capable to care for patients of higher complexity: for example, Level 1 Perinatal Centers provide care for low-risk women and newborns, while Level 2 Perinatal Centers provide Level 1 perinatal care services and in addition provide care to moderately high-risk mothers and newborns. Gestational age at delivery and birth weight serve as clinical guidelines that help risk stratify pregnancies and define what the adequate level of perinatal care services for each patient is. The severity groups I use in the analysis are not constructed using gestational age and birth weight; however, I expect the correlation between both measures to be high.

Table 2.1 reports summary statistics of childbirths in the working and main samples. The first two columns contain summary statistics of all childbirths during the period from 1995 to 2015, while the last two columns report summary statistics of all childbirths of experienced mothers during the period from 2005 to 2015 where the previous hospital is a feasible option. The childbirths in the

Table 2.1: Summary statistics of childbirths

Births	All		Second and later, 2005-15	
Sample	Working	Main	Working	Main
Age	28.7	24.8	30.3	27.2
Distance (miles)	8.0	7.6	8.4	8.0
Severity 1	0.45	0.40	0.40	0.39
Severity 2	0.50	0.55	0.55	0.56
Severity 3	0.05	0.05	0.05	0.05
Medicaid	0.25	0.29	0.20	0.26
Private	0.71	0.67	0.77	0.71
Other insurance	0.03	0.04	0.03	0.03
NYC	0.49	0.50	0.46	0.47
Repurchase			0.74	0.73
Observations	4,498,651	1,959,784	917,728	553,357

Notes: This table shows descriptive statistics of childbirths. Different columns correspond to different samples and/or time periods.

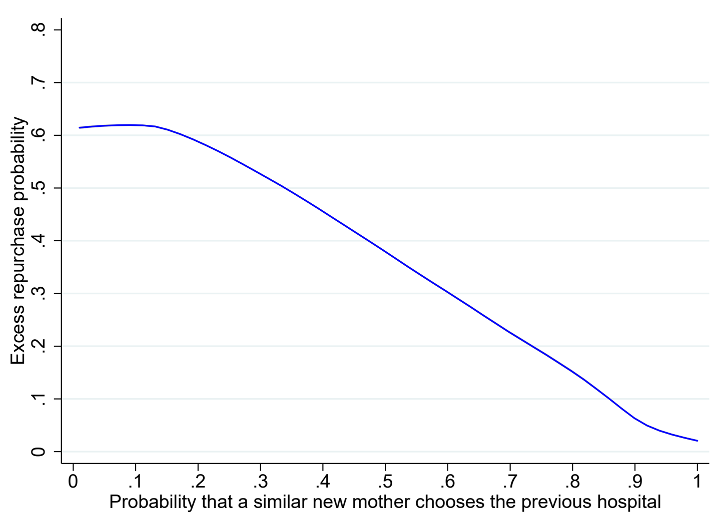
main and working samples are quite similar, although average age (by construction) and traveled distance are higher in the working sample than in the main sample.

2.3. Descriptive evidence

Here, I consider the repurchase behavior of mothers. Unless otherwise noted, I consider births of experienced mothers in the working sample during the period from 2005 to 2015 and for which the previous hospital is a feasible option. By repurchase, I mean that the mother visits the same hospital that she used in the previous pregnancy.

Patients' hospital choices for childbirth are quite persistent: 74% of the deliveries considered took place in the same hospital used by the patient in the previous episode. The fraction is very similar if we only consider second births (73%). The repurchase rate is also high for births of mothers who changed their zip code of residence (59%) and births of higher clinical complexity than the previous one (70%). These results suggest that persistence in patients' hospital choices is high: switching is rare, even when the choice environment changes across episodes.

Figure 2.1: Excess repurchase probability, second and later births



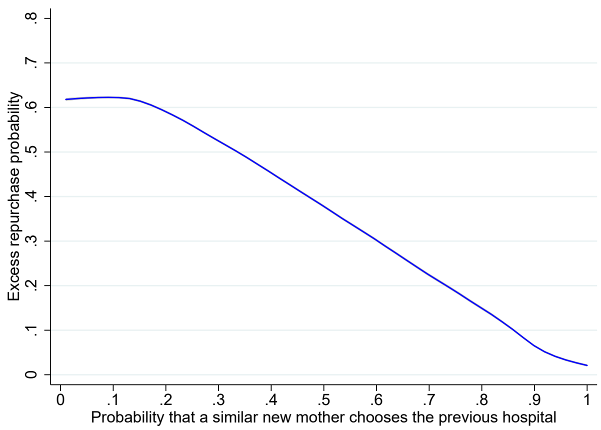
Notes: The figure shows the excess repurchase probability as a function of the marginal probability of choosing the previous hospital. Included: second and later births in the working sample during the period from 2005 to 2015. The shaded areas represent 95% confidence intervals.

To analyze persistence in behavior more precisely, I examine to what extent patients who went to a certain hospital A in the previous episode are more likely to visit hospital A for the current birth than observationally similar first time mothers (based on zip code, severity and year). I want to control, for example, for the possibility that the previous hospital chosen by the patent is used by the majority of observationally similar first time mothers: in this case, a high repurchase probability might only reflect the quality of the match between the patient and the hospital. I pool episodes according to the marginal probability of choosing the previous hospital, define the repurchase rate of a group as the fraction of patients in that group who choose the previous hospital for the current pregnancy, and obtain the excess repurchase probability as the difference between the repurchase rate and the corresponding marginal probability of choosing the previous hospital.

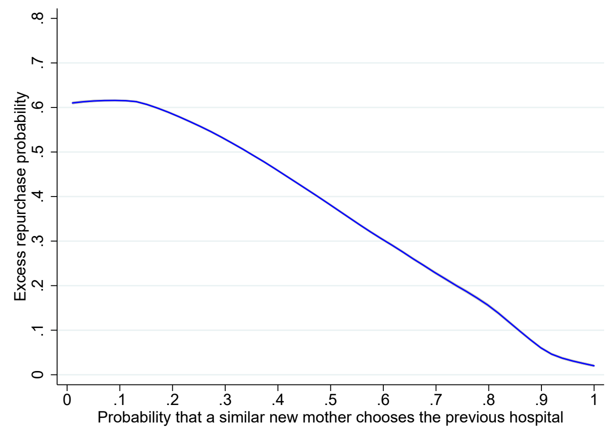
In Figure 2.1, I show the excess repurchase probability schedule constructed pooling all second and later births and all previous hospitals. The excess repurchase probability is positive and large for all values of the marginal choice probability. This means that a patient who visited a certain

Figure 2.2: Excess repurchase probability, by period

(a) 2005-2010



(b) 2011-2015



Notes: The figures show the excess repurchase probability as a function of the marginal probability of choosing the previous hospital. Included: second and later births in the working sample. Different figures correspond to different periods. The shaded areas represent 95% confidence intervals.

hospital A in the past is more likely to choose that same hospital for her current birth than an observationally similar new patient. For the average patient, the marginal probability of choosing the previous hospital is 0.25. The figure shows that this patient is more than three times as likely to return to the previous hospital than a similar first time mother (the excess repurchase probability is slightly higher than 0.5). As mentioned before, this repurchase frequency is computed considering all hospitals and childbirths, so it should be interpreted as the average persistence observed in the data.

In Figure 2.2, I report the excess repurchase probability schedule separately for different periods. The figure shows that the observed persistence in hospital choices of mothers has not changed much during the last years. To analyze whether patient loyalty varies across different types of hospitals, I construct the excess repurchase probability schedule separately for previous hospitals with different perinatal designations. The results are displayed in Figure 2.3. The patterns of repurchase behavior are quite similar across different types of facilities: on average, patients who visited a basic hospital in the previous episode show similar persistence in behavior as patients who went to a more sophisticated hospital in the past.

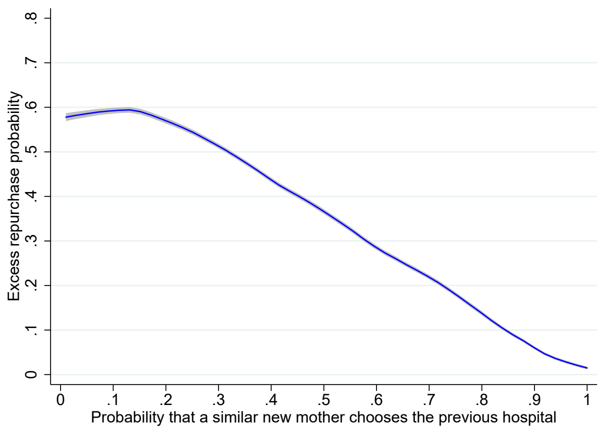
As mentioned before, I am analyzing sequences of choices for the same medical condition and therefore the different choice situations faced by a patient might be quite similar. Given this, I want to know what the extent of choice persistence is when there is a change in the choice environment of the patient. For the analysis below, I only consider second births.

First, I consider patients who change zip code of residence between births: movers experience a change in the relative attractiveness of hospitals based on convenience, which produces the type of variation across choice situations I am interested in. I focus on cases where the patient moves further away from her previous hospital, so there is no concern about sorting of patients into neighborhoods based on preferences for facilities. If the patient reallocates further away from her previous hospital, the potential for unobserved heterogeneity to explain a repeat visit after moving is diminished. To control for the possibility that the relocation of the patient renders the previous hospital too inconvenient for the current episode, I classify episodes according to the extra distance to the previous hospital relative to the distance traveled in the previous episode. In Figure 2.4, I present the excess repurchase probability schedule for different groups. In this case, the marginal probability of choosing the previous hospital is the market share of the previous hospital among similar first time mothers in the destination zip code. The results are very similar across different groups. Independently of the extra distance a patient needs to travel to return to the previous hospital (compared to the previous birth), she is more likely to choose the previous hospital than a similar first time mother in the destination zip code.

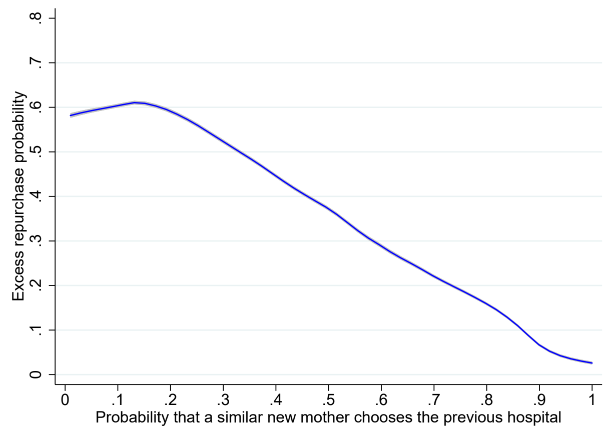
Second, I analyze the choice persistence of non-movers exploiting variation in choice sets induced by transition across severity categories. The risk profile of the first pregnancy made the patient gravitate towards a certain type of hospital in that episode. However, if the risk profile of the second pregnancy changed with respect to the first birth, then the adequacy of facilities for providing care in the current episode might be different than before (as I only consider non-movers,

Figure 2.3: Excess repurchase probability, by perinatal designation of the previous hospital

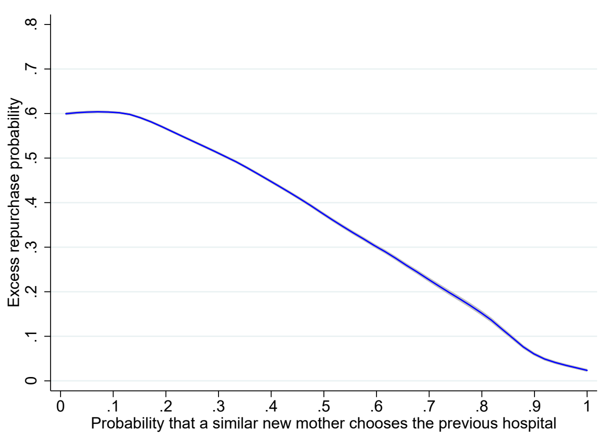
(a) Level 1



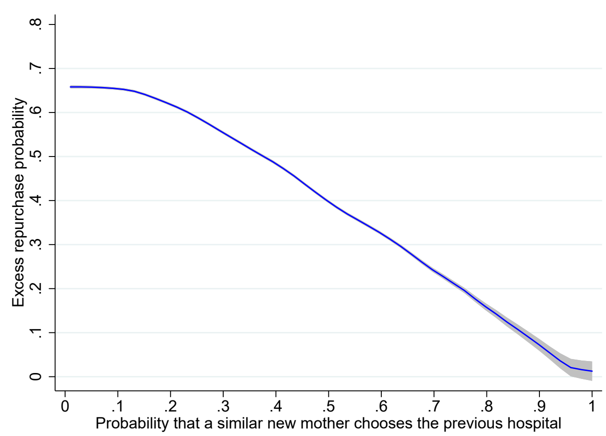
(b) Level 2



(c) Level 3



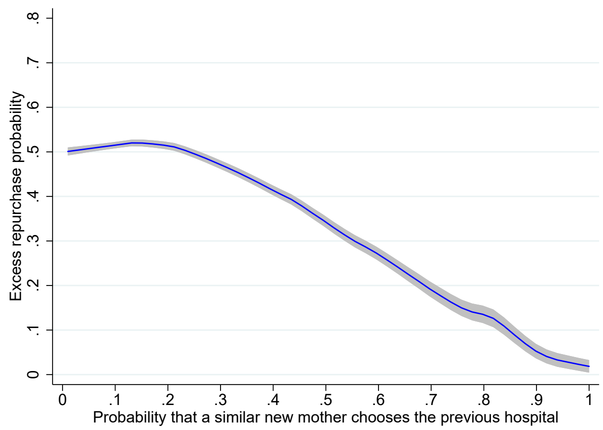
(d) Regional Perinatal Center



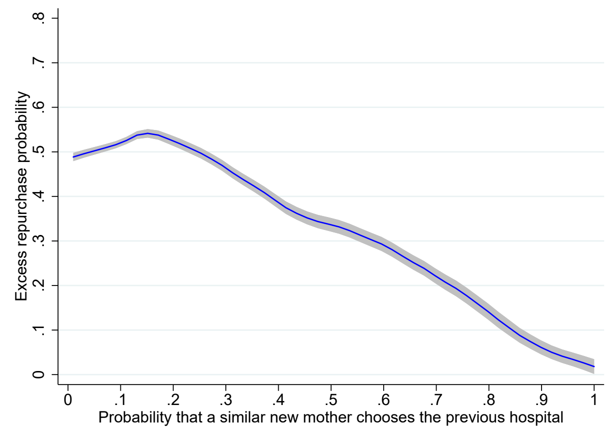
Notes: The figures show the excess repurchase probability as a function of the marginal probability of choosing the previous hospital. Included: second and later births in the working sample during the period from 2005 to 2015. Different figures correspond to different perinatal designations of the previous hospital. The shaded areas represent 95% confidence intervals.

Figure 2.4: Excess repurchase probability, movers, by differential distance

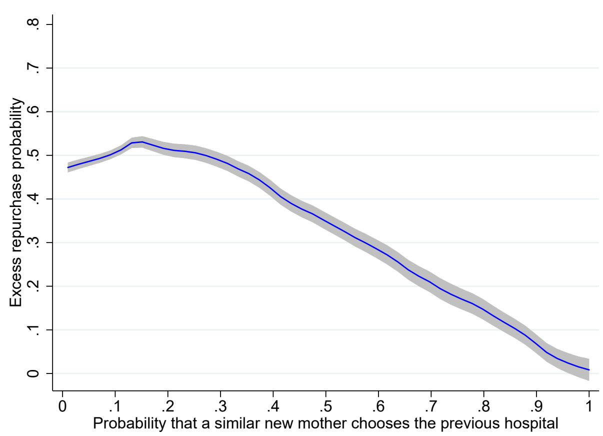
(a) 0 to 2 miles



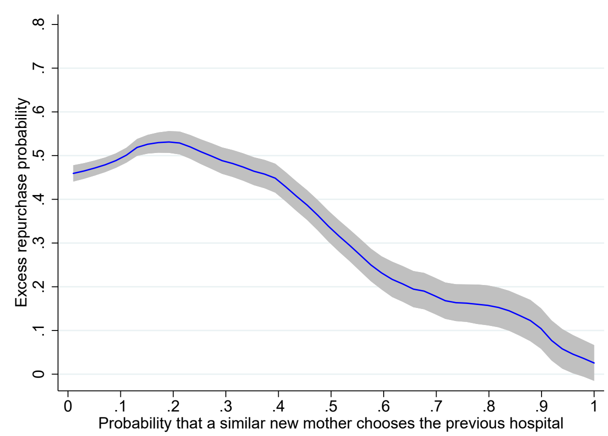
(b) 2 to 5 miles



(c) 5 to 10 miles



(d) 10 to 15 miles



Notes: The figures show the excess repurchase probability as a function of the marginal probability of choosing the previous hospital. Included: second births of movers in the working sample during the period from 2005 to 2015. Different figures correspond to different cases depending on the current distance to the previous hospital compared to the previous episode. The shaded areas represent 95% confidence intervals.

Table 2.2: Raw repurchase probability, by severity and perinatal designation

Designation previous	Severity previous	Severity current 1	2	3
1	1	0.89	0.84	0.72
1	2	0.88	0.85	0.68
1	3	0.87	0.84	0.69
2	1	0.87	0.80	0.70
2	2	0.86	0.83	0.71
2	3	0.85	0.78	0.80
3	1	0.83	0.78	0.71
3	2	0.82	0.80	0.73
3	3	0.81	0.77	0.78
4	1	0.84	0.83	0.80
4	2	0.81	0.84	0.82
4	3	0.76	0.82	0.82

Notes: This table shows the fraction of mothers who visited the same hospital used in the previous pregnancy, according to the perinatal designation of the previous hospital, the severity of the previous pregnancy, and the severity of the current pregnancy. Included: second births of non-mover mothers in the working sample during the period from 2005 to 2015.

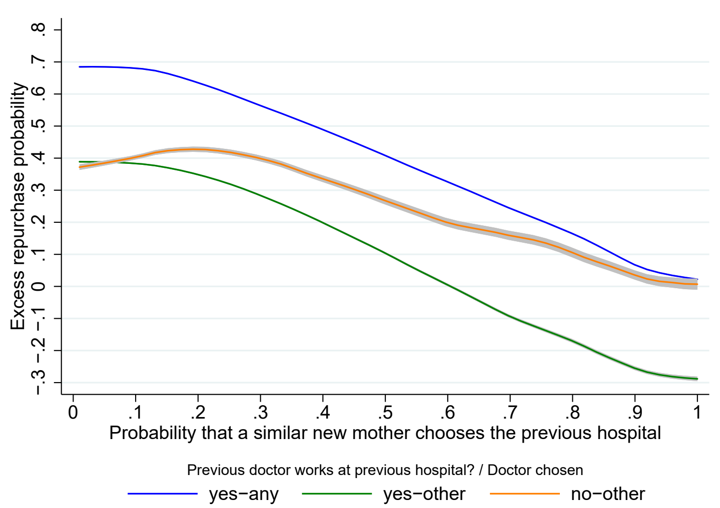
the convenience of the different hospitals is the same as in the first episode). Consider, for example, a mother whose first pregnancy was very complex and therefore was willing to travel several miles to give birth at a Regional Perinatal Center. If her second pregnancy is low-risk, then she might be less willing than before to travel far to give birth at the same Regional Perinatal Center if there is a hospital with a lower perinatal designation closer to her home. Therefore, the extent to which patients remain loyal to a hospital when transitioning across severity groups is informative about the nature of persistence. Of particular interest are patients who initially had a low-risk pregnancy but whose current pregnancy is of higher complexity.

In Table 2.2, I report raw repurchase probabilities for non-movers based on the perinatal designation of the previous hospital and the severity of the previous and current pregnancy. The table shows that, for example, for a patient whose previous birth was of severity 1 and used a Level 1 facility, the probability of using that same hospital for her second birth of severity 2 is 0.84. The main takeaway from Table 2.2 is that repurchase probabilities are quite high across the board. For Level 1 to Level 3 facilities, the repurchase probability tends to decrease as the severity of the current pregnancy increases. Repurchase frequencies for Regional Perinatal Centers are more stable. These results suggest that inertia is lower the higher the mismatch risk-quality created by transitions across severity groups (especially for cases of high severity and low quality), although repurchase frequencies remain quite high in absolute terms.

One of the main concerns in the analysis above is that the persistence in hospital choices of patients that I observe in the data might just reflect loyalty to doctors. As the type of medical condition is similar across episodes, the doctor who treated the patient in the previous episode might be a feasible option in the current one. I consider that a doctor works in a given hospital during a certain semester if s/he is listed as operating physician for at least five deliveries in that facility during that period. Most physicians in the sample practice at one hospital according to my definition. Therefore, if a patient goes to the hospital her doctor practices in and her previous doctor continues working at the same hospital as before, this will produce persistence in hospital choice (although the patient is only loyal to the doctor). To analyze this, I look at repurchase probabilities for different types of cases, depending on where the doctor that handled the previous delivery works at the time of the current birth.

Consider those cases where the previous doctor continues working at the previous hospital. The overall repurchase frequency captures both patients who use the same doctor as in the previous episode and patients who switch doctors. Therefore, it is an upper bound on the extent of persistence in hospital choice. The fraction of cases where the patient chose the same hospital as in the previous childbirth but the delivery was handled by a different doctor gives us a conservative measure of persistence in hospital choice: if there was only loyalty towards doctors, the patient would have chosen the same physician as in the previous birth. Therefore, this is a lower bound for persistence in hospital choice: some of the patients who chose both the same hospital and same doctor as before might only care about using the previous hospital but are not included in this measure. In order to

Figure 2.5: Excess repurchase probability



Notes: The figure shows the excess repurchase probability as a function of the marginal probability of choosing the previous hospital. Included: all second and later births in the working sample during the period from 2005 to 2015. The green and blue lines are constructed using those cases where the previous doctor continues working at the previous hospital. The green line refers to the probability of choosing the previous hospital but not the previous doctor, while the blue line refers to the probability of choosing the previous hospital and any doctor. The red line is constructed using those cases where the previous doctor does not continue working at the previous hospital. The shaded areas represent 95% confidence intervals.

obtain a clean measure of loyalty to hospitals, I analyze cases in which the previous physician does not work at the previous hospital at the time of the current birth. In this case, if a patient chooses the previous hospital, it automatically means that a different physician will be handling her delivery.

Figure 2.5 shows the excess repurchase probability schedule for different cases, depending on whether the previous doctor continues working at the previous hospital and whether the patient chose the previous doctor, the previous hospital, or both. The blue and green lines are constructed considering those cases where the previous physician continues working at the previous hospital. For constructing the blue line, I define a repurchase as a patient who chooses the previous hospital, independently of the doctor chosen. Some of the returning patients use the same doctor as before, so I cannot rule out the possibility that all these patients care about is the doctor who handles the delivery. Therefore, the blue line captures both loyalty to hospitals and doctors. For constructing the green line, I define a repurchase as a patient who chooses the previous hospital but not the previous doctor. Therefore, the green line provides a lower bound for loyalty to hospitals: it only captures persistence in hospital choice but not necessarily all such persistence, as some of the patients who chose both the previous hospital and previous doctor might only care about the hospital visited. The red line is constructed considering those cases where the previous doctor does not continue working at the previous hospital. For these cases, choosing both the previous hospital and the previous doctor is not feasible. I define a repurchase as a patient who chooses the previous hospital. Therefore, the red line provides a clean measure of persistence in hospital choice. As observed in Figure 2.5, the excess repurchase probability is positive and lies between the excess repurchase probability schedule for the first two cases (which was expected as the latter represent upper and lower bounds for persistence in hospital choice). These results suggest that not all the observed persistence in hospital choices of patients is due to loyalty to doctors.

There are a couple of issues to consider regarding this lower bound. First, even if the previous doctor continues working in the previous hospital, it might be the case that s/he does not perform the surgical procedure needed in the current situation (for example, if the first birth was a vaginal delivery and for the second birth a C-section is required). To control for this possibility, I perform the same analysis considering only those cases where the method of delivery was the same as in the previous birth. Second, the doctor might have been unavailable at the time of delivery (because of sudden onset of delivery or because of capacity constraints). To control for this possibility, I also perform the analysis using only deliveries that take place during weekdays, for which I expect that the incidence of unavailability of doctors is relatively low. Repurchase frequencies (not reported) are similar to the general case, which suggests these issues are not biasing results towards finding persistence in hospital choice.

Although I explore the possibility that patients are loyal to doctors, it might be the case that patients are loyal to a physician group and not to individual doctors. Even if the doctor that handled the previous delivery is not available at the previous hospital, a patient might continue using the facility just because other doctors from the same practice work there. Data limitations prevent me from exploring this issue in more detail.

The main descriptive findings of this section are the following. First, there is substantial persistence in hospital choices of patients for childbirth, even after scaling repurchase probabilities by marginal choice probabilities that take into account location, severity and year of admission. Second, the degree of choice persistence is similar across different types of hospitals and patients. In particular, patients who experience a change in the choice environment across births (change of location or severity) tend to choose the previous hospital at a higher rate than similar first time mothers. Third, not all the observed persistence in patients' hospital choices can be explained by loyalty to doctors. The previous analysis illustrates the sources of variation in choice sets that can be exploited for estimating the parameters of a hospital demand model that accounts for state dependence.

2.4. Discrete choice model

In order to quantify switching costs and preferences for hospital attributes, I estimate a conditional logit model of hospital choice for childbirth. This analysis is only intended to illustrate the magnitude of persistence in hospital choices of mothers for childbirth using more traditional tools than the ones used in the descriptive part of the chapter and therefore should be viewed as a complement of the latter.

In general, hospital demand models do not account for the temporal dependence of choices. Exceptions are Shepard (2016) and Raval and Rosenbaum (2017), which include a lagged dependent variable in the demand model (note that, in principle, the coefficient on this variable captures both state dependence and unobserved heterogeneity). I follow this approach and assume that the utility

Table 2.3: Estimation sample, discrete choice model

Age	25.49
Distance (miles)	7.82
Severity 1	0.22
Severity 2	0.71
Severity 3	0.07
Medicaid	0.20
Private insurance	0.76
Other insurance	0.04
Observations	62,009

Notes: This table shows summary statistics of childbirths in the estimation sample.

that patient i gets from choosing hospital j for birth t is given by:

$$u_{ijt} = \alpha_i d_{ijt} + \beta_i Z_j + \gamma_i h_{ij(t-1)} + \varepsilon_{ijt} \tag{2.1}$$

where d_{ijt} is the distance from the centroid of patient's i zip code of residence in episode t to hospital j, Z_j includes hospital characteristics, and h_{ijt} is an indicator for whether the patient went to hospital j for birth t. Therefore, the patient receives a utility premium γ for choosing the same hospital as in the previous episode. As explained below, the model allows for observed heterogeneity in preferences by allowing the coefficients on travel distance, hospital characteristics, and previous use to depend on patient characteristics.

For this application, I consider pregnancies of mothers that live in Erie county. Moreover, I restrict attention to mothers who were at most 18 years old in 1995 (so I am confident that I observe all their childbirths in the data) and who had their first childbirth on or after 2005. The choice set is defined as the set of hospitals in the market for delivery services in the corresponding admission year that are located in Erie county[2]. The resulting sample includes a total of 44,255 mothers and 62,009 pregnancies during the period from 2005 to 2015. Table 2.3 presents summary statistics on this sample.

[2] 99% of childbirths of patients who live in this county take place in hospitals located in the same county.

Table 2.4: Estimates of the conditional logit model

	(1)	(2)
Distance	0.203	0.189
	(0.008)	(0.009)
Previous		2.332
		(0.140)
Repurchase probability	0.38	0.82
Pregnancies	62,009	62,009

Notes: This table shows the results from the conditional logit model of hospital choice for childbirth. The first column contains the estimates from a model that does not include the previous choice of the patient as explanatory variable, while the second column contains the estimates from a model that includes a lagged dependent variable. Both specifications include hospital-severity fixed effects and interactions between patient characteristics (age, severity, type of insurance) and distance and the lagged dependent variable.

I now discuss the estimates from the patient demand model, following the insights of Raval and Rosenbaum (2017). I have tried various specifications that differ in terms of the covariates included (such as the specific interactions between patient and facility characteristics) and the qualitative results are consistent. Therefore, I will focus on the main specifications of the demand model. In all these cases, I include hospital-severity fixed effects and allow the coefficients on travel distance and the indicator of previous use of the hospital to depend on the severity, age, and type of insurance of the patient. Table 2.4 displays the main estimates of interest. In specification 1, hospital choices are a function of travel time and hospital characteristics only. The model predicts that approximately 38% of experienced patients in the estimation sample would return to the previous hospital. In specification 2, I include a lagged dependent variable. The effect of the previous use variable is significant: on average, having visited a hospital in the previous episode increases the probability of choosing that same facility for the current pregnancy 1.7 times above the baseline. The inclusion of this variable therefore produces different predictions of patients' hospital choices. Specifically, more than 80% of experienced mothers in the estimation sample would return to the previous hospital with this specification. I also estimated the model using only first-time mothers, for which state

Table 2.5: Distribution of WTP, by specification

	p10	p25	p50	p75	p90	mean
Specification 1	0.56	0.97	1.71	2.71	3.65	1.92
Specification 2	0.10	0.22	0.56	6.57	11.75	3.39

Notes: This table shows the distribution of WTP for Buffalo General Medical Center according to the estimates from specifications 1 and 2 of the conditional logit model.

dependence is not relevant, and the parameters on distance and other explanatory variables (results not reported) are similar to the estimates from specification 2.

For the following analysis, I only consider women with previous births, whose choices can be affected by state dependence. I use the estimated parameters from specifications 1 and 2 to compute patient-specific choice probabilities. Then, I estimate the willingness-to-pay (WTP) of patients for different hospitals in this market, defined as the utility of adding a facility to the choice set of the patient divided by the marginal utility of distance (based on the estimated distance coefficients from the demand model). Table 2.5 shows the distribution of WTP for a particular facility, Buffalo General Medical Center, according to the estimates from specifications 1 and 2 of the conditional logit model. Table 2.6 shows the proportion of patients who used Buffalo General Medical Center in the previous pregnancy within each quintile of the distribution of WTP for this hospital. For example, this table shows that 23% of patients in the third quintile of the distribution of WTP according to specification 1 used the hospital in the previous pregnancy. There are two main points to be noted. First, adding the previous use variable increases heterogeneity in WTP across patients substantially. Second, the group of patients with high values of WTP for Buffalo General Medical Center contains a relatively high proportion of patients who used the hospital in the previous episode. However, this feature is striking when a lagged dependent variable is included in the demand model.

The results of this section are consistent with the descriptive part of this chapter and with the findings of Shepard (2016), Raval et al. (2017b), and Raval and Rosenbaum (2017). The conclusions are relevant for policy analysis, particularly considering the diffusion of narrow networks during

Table 2.6: Proportion of patients who used the hospital in the previous episode, by quintile of the WTP distribution

Quintile	Specification 1	Specification 2
1	0.10	0.00
2	0.17	0.00
3	0.23	0.02
4	0.38	0.32
5	0.52	0.99

Notes: This table shows the proportion of patients who used Buffalo General Medical Center in the previous pregnancy in each quintile of the distribution of WTP for that facility.

the last years. The results suggest that the impact of network exclusions differs greatly across different consumer segments and that previous hospital choices of patients are an important factor in explaining this heterogeneity. Raval and Rosenbaum (2017) also make the point that incorporating state dependence in models of hospital choice affects the evaluation of a network exclusion, although they focus on how the long and short run effects of the network exclusion depend on the relative magnitude of switching costs and unobserved preference heterogeneity.

2.5. Conclusion

In this chapter, I study persistence in hospital choices of mothers for childbirth. Specifically, I analyze to what extent mothers use the same hospital for different pregnancies and what the sources of the observed persistence in behavior are. I find that mothers continue using the previous hospital more often than what we would expect given the characteristics of the current episode. Moreover, the persistence in choice remains even when I consider mothers who experience large changes in the choice environment across episodes. The identification strategy used here is more traditional than the one employed in Chapter 1 and is based on the existence of new and experienced patients in the market. The medical condition of the patient is the same across episodes, which facilitates the analysis. However, this feature opens up the possibility that the loyalty to hospitals observed in the data in fact captures loyalty to doctors. To explore this possibility, I compare cases that differ on whether the previous doctor continues working at the previous hospital. I find that there is substantial persistence in hospital choice that is not associated to loyalty to doctors. Finally, I estimate a model of hospital choice to quantify preferences for different hospital attributes in a more precise way. This analysis is only intended to illustrate the magnitude of persistence in hospital choices of mothers for childbirth using more traditional tools than the ones used in the descriptive part of the chapter. A full analysis that explores choice persistence in different local markets is beyond the main objective of this study, but I consider this is a relevant starting point.

Chapter 3

The impact of payment reform on the behavior of Long-Term Care Hospitals

3.1. Introduction

Medicare, the federal health insurance program for the elderly and the disabled, sometimes pays medical providers differently for the same service depending on the site of care. The typical example is that Medicare pays more to hospital-owned outpatient centers than to independent physician offices for the same procedures. As a result, the cost for Medicare of treating a patient with a given clinical profile might depend on the care setting used (Pruitt, 2013).

Site-neutral payments try to reduce or eliminate the gap between the different rates by relating payments to patient characteristics, rather than to the setting of care. Medicare has begun incorporating site-neutral payments into payment policies to minimize the reimbursement discrepancy across sites of care. The Medicare Payment Advisory Commission (MedPAC) has recommended site-neutral payments between hospital outpatient departments and physicians' offices for certain services, consistent payment between acute care hospitals and long-term care hospitals for certain classes of patients, and site-neutral payments for select conditions treated in both skilled nursing facilities and inpatient rehabilitation facilities (MedPAC, 2012; MedPAC, 2015; MedPAC, 2016).

In this chapter, I study the effects of a major regulatory reform to the Long-Term Care Hospital (LTCH) industry that introduced site-neutral payments for certain patients treated in this setting. Before the reform, LTCHs were paid an LTCH rate - higher than the rate received by general hospitals for treating a similar patient - for all Medicare patients. Under the new system, LTCHs are paid the LTCH rate for patients with high severity of illness, and a much lower site-neutral rate for treating patients with lower medical acuity. The reform was expected to reduce the incentives of LTCHs to admit patients who have borderline needs for long-term acute care and therefore could be treated less expensively in other settings.

I study the strategies that LTCHs have implemented to mitigate the impact of the new regulatory framework and the effects of these actions on access to care and treatment intensity. First, I analyze whether LTCHs have taken actions to change the composition of their patient pool in order to improve their profitability. More specifically, I assess how the incentives of hospitals for admitting site-neutral patients changed after the implementation of the new system. Second, I explore how LTCHs have modified clinical practices and treatment decisions in response to the change in financial incentives. In particular, I analyze how the distortions in the timing of patient discharge documented in the literature have changed after the introduction of the new system.

The main findings of this study are the following. First, hospitals screen patients who are referred for long-term acute care more actively than before the reform and are more likely to reject those who will not (or are unlikely to) meet the criteria for exclusion from the site-neutral rate. However, there is heterogeneity in behavior across hospitals. Second, many for-profit hospitals alter treatment patterns in response to the change in reimbursement. In particular, these hospitals are more likely to discharge site-neutral patients before the short-stay outlier threshold and right after crossing it. While the first response was expected given the change in the payment schedule, the latter result is hard to explain given that marginal incentives did not change for those cases. I analyze alternative explanations for this "anomaly" in hospital behavior, but I do not find conclusive empirical support for any of them. Therefore, the reasons for some of the changes in discharge patterns remain largely unexplained.

This study is related to different strands of literature. First, it is related to studies about the LTCH industry. This industry (and more generally the post-acute care sector) has been under-studied in comparison to the general hospital industry. As a result, there are few papers that study this setting. The papers most closely related to my study are Kim et al. (2015), Eliason et al. (2016) and Einav et al. (2017). I postpone a detailed discussion of these studies until the basic elements necessary for the analysis are introduced. My analysis of discharge patterns follows this prior work. However, while previous studies investigate the impact of the Medicare LTCH payment schedule on the discharge behavior of hospitals during the pre-reform period, I analyze how LTCHs changed their behavior after the introduction of the new payment system.

More generally, this chapter is related to studies of the impact of financial incentives on provider behavior. Particularly relevant are previous papers analyzing how providers respond to changes in payment systems. See, for example, McGuire and Pauly (1991), Hodgkin and McGuire (1994), Gruber and Owings (1996), Gruber et al. (1999), Dafny (2005), and Ho and Pakes (2014).

The rest of the chapter is organized as follows. Section 3.2 provides an overview of the LTCH industry and discusses the payment reform that is the subject of this study. Section 3.3 describes the data used in the analysis. Section 3.4 presents descriptive evidence of the impact of the new payment system on the admission and discharge decisions of hospitals. Section 3.5 analyzes the change in hospital behavior brought about by the reform using a dynamic model of hospital behavior. Section 3.6 contains concluding remarks.

3.2. Institutional background

In this section, I provide a brief overview of the Long-Term Care Hospital (LTCH) industry and its regulatory framework. Then, I describe the payment system for LTCHs, putting emphasis on the regulatory reform that is the subject of this study. This material is based on IPPS / LTCH PPS Proposed and Final Rules issued by CMS, MedPAC reports (MedPAC, 2012; MedPAC, 2015; MedPAC, 2016), and CMS (2016).

3.2.1. Long-Term Care Hospitals

To qualify as an LTCH, a hospital must be certified by Medicare and have an average Medicare inpatient length of stay of 25 days or longer. There are no additional facility or patient criteria: LTCH status does not depend on the type of services provided by the hospital nor on the specific types of patients admitted. However, as a result of the regulatory requirement mentioned above, the patient population and scope of services found in LTCHs differ from those found in general acute care hospitals (GACHs).

The distinction between GACHs and LTCHs is administrative rather than medical. In particular, LTCH status refers to how the hospital is paid by Medicare. Both GACHs and LTCHs provide inpatient services and are paid for each Medicare patient based on the Prospective Payment System (PPS)[1]. This is a per discharge payment system with a fixed reimbursement based on the Diagnostic Related Group (DRG) to which the patient's stay is grouped. LTCHs are paid according to the LTCH PPS, which uses MS-LTC-DRGs as a patient classification system, while GACHs are paid according to the IPPS, which uses MS-DRGs for classifying patients. The DRGs in the LTCH PPS and IPPS are the same, but in most cases the payment that a hospital receives for treating a given Medicare patient is higher if the hospital has LTCH status.

LTCHs provide inpatient services for patients who require extended recovery time after an acute health episode. Therefore, they are considered part of the post-acute care (PAC) setting. In addition to LTCHs, PAC providers include inpatient rehabilitation facilities (IRFs), skilled nursing facilities (SNFs), and home health agencies (HHAs). LTCHs are the least used type of PAC provider, but they care for patients with the highest level of severity of illness: general hospital stays discharged to LTCHs have the highest average cost and average length of stay among all discharges to PAC providers (Tian, 2016).

Patients are typically admitted to LTCHs from GACHs once they are medically stable. LTCH patients can be discharged home, to a less sophisticated PAC facility, to an acute care hospital, or to a hospice. Following Einav et al. (2017), I refer to the first two types of discharges as "downstream" discharges and to the other two types of discharges as "upstream" discharges. Discharges can also be due to the death of the patient in the LTCH.

Table 3.1 shows the evolution of the Medicare LTCH industry during the period from 2012 to 2017. The volume of LTCH services has been declining during the last years. However, Medicare spending on this segment is significant: in Fiscal Year (FY) 2015, Medicare paid $5.1 billion for about 152,000 LTCH stays nationwide. In terms of clinical characteristics, the Medicare LTCH population tends to be sicker over time.

[1]LTCHs have been paid based on the PPS since 2003. See Subsection 3.2.2 for details.

Table 3.1: Evolution of the Medicare LTCH industry, 2012-2017

FY	Discharges	Total Spending ($B)	Mean Spending ($)	DRG weight	Diagnoses	Procedures
2012	163,229	5.34	32,707	1.10	15.76	2.25
2013	160,299	5.30	33,074	1.12	16.31	2.32
2014	155,826	5.14	32,973	1.12	16.83	2.33
2015	152,348	5.09	33,421	1.13	17.40	2.19
2016	147,847	4.84	32,753	1.14	18.83	2.02
2017	136,476	4.19	30,716	1.15	19.23	2.13

Notes: This table shows the evolution of Medicare LTCH discharges, spending, and patient characteristics during the period from FY 2012 to FY 2017. DRG weight refers to the average FY 2015 DRG weight of discharges in the corresponding fiscal year.

There are approximately 420 LTCHs certified by Medicare. The geographic distribution of LTCHs across the country is uneven: although there are LTCHs in almost every state, 25% of LTCHs are concentrated in Texas and Louisiana. An LTCH can be freestanding or co-located within a GACH, as a hospital-within-hospital (HwH), provided it is independently owned and has its own medical staff. In the latter case, the LTCH operates as a separately licensed hospital within the host hospital. LTCHs are predominantly for-profit hospitals. There are two main for-profit chains, Kindred and Select, which jointly operate about 40% of facilities in the country (see Section 3.3 for descriptive statistics of LTCHs). Medicare accounts for about two thirds of LTCH patients. According to company reports by Kindred, patients covered by non-government payors are generally more profitable than those covered by Medicare and Medicaid.

3.2.2. Payment reform

Here, I describe how the payment an LTCH receives from Medicare is determined and the main elements of the regulatory reform that is the main subject of this study. I discuss two payment systems: the "old" (or "pre-reform") system applied to Medicare discharges occurring in cost reporting periods beginning before October 1, 2015, while the "new" (or "post-reform") system applies to Medicare discharges occurring in cost reporting periods beginning on or after October

1, 2015. In both systems, each LTCH stay is assigned to a DRG based on the principal diagnosis, secondary diagnoses, surgical procedures, sex, age, and discharge status of the patient. Unless otherwise noted, the average length of stay (ALOS) of a DRG refers to the ALOS of discharges in that DRG during the previous fiscal year. The payment schedule refers to the relationship between the payment received by the LTCH and the length of stay of the patient.

Under the old system, the payment an LTCH receives from Medicare depends on the LOS of the patient relative to the short-stay outlier (SSO) threshold, defined as 5/6 of the ALOS of the corresponding DRG. If the LOS of the patient was greater than the SSO threshold, then the hospital received the full MS-LTC-DRG payment independently of the LOS of the patient. If the LOS of the patient was lower than or equal to the SSO threshold, then the stay was paid as a SSO and the payment from Medicare is a non-decreasing function of the LOS of the patient. See Appendix C for details on how the payment for SSO cases is determined. I refer to the payment determined this way as the LTCH PPS rate[2].

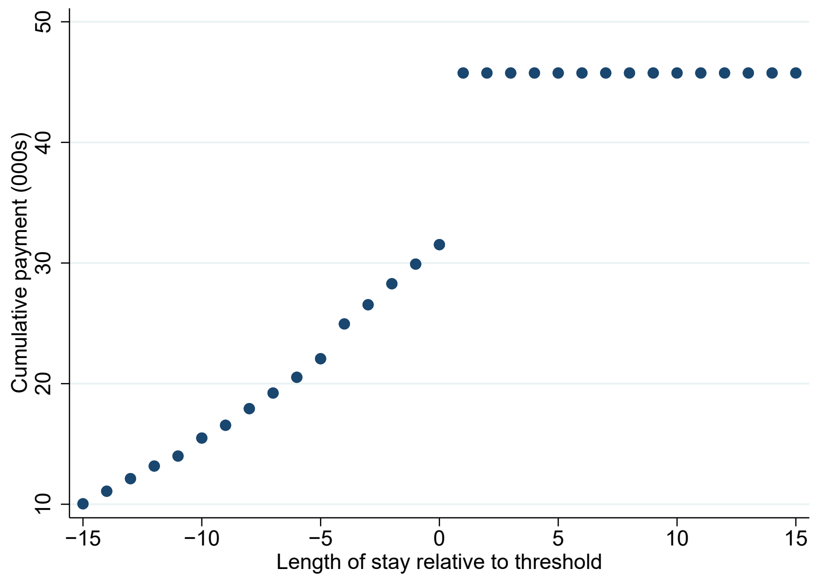
Figure 3.1 shows the average payment schedule faced by LTCHs during FY 2015 (I discuss the data used for constructing the figure in Section 3.3). For each patient, I determine the payment schedule faced by the hospital visited. I define the normalized LOS as the difference between the LOS of the patient and the SSO threshold corresponding to the DRG to which the LTCH stay was grouped. If the normalized LOS is negative or zero, then the stay would be paid as a SSO under the LTCH PPS; if it is positive, the hospital would receive the full DRG payment under such system[3]. Then, I pool the payment schedules of all patients[4] based on the normalized LOS and compute the mean payment for each day. Before the SSO threshold, the per discharge payment increases approximately linearly with the LOS of the patient, at a rate of $1,400 per day. At the

[2]LTCHs were exempt from Medicare prospective payment through FY 2002. The LTCH PPS was phased in over three years beginning in FY 2003. Before that, Medicare paid LTCHs based on estimated costs. I do not discuss the payment system that was in place before 2003, as the focus of the analysis is the reform implemented in 2015.

[3]Day 1 is the first day for which the LOS of the patient is higher than the SSO threshold, while day 0 is the last day for which the case is paid as a SSO.

[4]Note that I am pooling payment schedules across hospitals and DRGs.

Figure 3.1: Payment schedule, FY 2015 discharges



Notes: This figure shows the average payment schedule for LTCH discharges in FY 2015. Average payment across hospitals and DRGs.

SSO threshold, there is a discontinuity in the payment schedule: the extra revenue from keeping a patient beyond the SSO threshold is approximately 48% of the payment that the hospital would have received from discharging the patient the day just before crossing it. Once the SSO threshold is crossed, the payment is fixed and does not change with the LOS of the patient.

The new system is a two-tiered payment system mandated by the Bipartisan Budget Act (December, 2013). Under this system, discharges that meet certain statutorily defined criteria continue receiving the LTCH PPS rate as before the reform: the hospital receives the full MS-LTC-DRG payment or the SSO payment depending on the patient's LOS and the ALOS of the corresponding DRG. On the other hand, discharges that do not meet the criteria are paid at a site-neutral payment rate, which is the lower of the IPPS comparable per diem amount and the estimated cost of the case; the SSO policy does not apply to these cases. I refer to discharges that meet the exclusion criteria as "standard" cases (and I refer to the LTCH PPS rate as the standard rate), and to discharges that do not meet the exclusion criteria as "site-neutral" cases. The criteria for exclusion from the

Table 3.2: Clinical profile of patients in FY 2015, by type

	Site-neutral	Standard
DRG weight	0.96	1.29
Diagnoses	16.80	18.12
Procedures	1.90	2.48
Discharged death	0.08	0.17
Discharged upstream	0.11	0.14
Discharged downstream	0.79	0.67
Discharged other	0.03	0.02
30-day mortality	0.10	0.16
90-day mortality	0.19	0.27
Less than 45 y/o	0.05	0.03
45 to 64 y/o	0.24	0.21
65 to 74 y/o	0.32	0.36
75 to 84 y/o	0.23	0.27
More than 85 y/o	0.15	0.13

Notes: This table shows information about the health profile of site-neutral and standard cases during FY 2015. 30 and 90-day mortality rates are conditional on a live discharge from the LTCH.

site-neutral payment rate are the following (CMS, 2016) (see Appendix C for additional details):

1. The LTCH discharge does not have a principal diagnosis relating to a psychiatric diagnosis nor to rehabilitation.
2. The LTCH stay was immediately preceded by a discharge from a subsection (d) hospital (general hospital).
3. The immediately preceding stay in a subsection (d) hospital included at least 3 days in an intensive care unit (ICU) (ICU criterion), or the patient received at least 96 hours of ventilator services during the LTCH stay (ventilator criterion).

Table 3.2 shows that standard patients are more severely ill than site-neutral patients: they are assigned to DRGs with higher weights, have more reported diagnoses, have a higher number of procedures performed, and are more likely to die in the hospital or to be discharged upstream.

Therefore, under the new system LTCHs are paid the LTCH PPS rate for patients with high severity of illness, and a lower site-neutral rate - comparable to what general hospitals receive - for treating patients with lower medical acuity.

The new payment system started with a four-year phase-in period tied to each LTCH's cost reporting period[5]. During the phase-in period, payment for discharges that do not meet the exclusion criteria are based 50% on the LTCH PPS rate and 50% on the new site-neutral rate. The phase-in period for a facility corresponds to the first four cost reporting periods beginning on or after October 1, 2015, while full implementation of the new system starts with the first cost reporting period starting on or after October 1, 2019[6]. For example, if the start of the cost reporting period of a hospital is July 1, the phase-in period began on July 1, 2016, and full implementation of the new system will begin on July 1, 2020. Site-neutral discharges before July 1, 2016, are paid according to the old system; site-neutral discharges between July 1, 2016, and June 30, 2020, are paid the 50/50 blend rate; for discharges on or after July 1, 2020, payment Medicare patients not meeting the patient criteria will be determined by the site-neutral payment rate only. In what follows, the payment under the new system for site-neutral cases refers to the 50/50 blend, as all hospitals are currently in the transition period.

Due to the way the new payment system was rolled out, LTCHs had incentives to modify their cost reporting periods to delay its implementation. For example, if the cost reporting period of a hospital begins November 1, then the new payment system applies to all discharges on or after November 1, 2015. If the cost reporting period begins September 1, then the new payment system applies to all discharges on or after September 1, 2016. In the latter case, the hospital receives the LTCH PPS rate for site-neutral cases for 10 more months than in the former case. In Table 3.3, I show the distribution of hospitals according to their cost reporting periods for the period from FY

[5]The cost reporting period of a hospital refers to the 12-month period of operations covered by reports submitted to Medicare.

[6]Originally, the transition period included the first two cost reporting periods of a hospital beginning on or after October 1, 2015. The Bipartisan Budget Act of 2018 extended the transition period by an additional two years to include cost reporting periods beginning before September 30, 2019.

Table 3.3: Distribution of hospitals, by start of cost reporting period, FY 2012-2016

	2012	2013	2014	2015	2016
January	87	78	73	62	66
February	16	15	15	13	13
March	18	17	13	13	13
April	15	13	14	12	12
May	14	19	9	9	9
June	33	29	28	25	26
July	35	33	33	33	33
August	16	14	16	16	17
September	84	102	125	144	139
October	18	17	16	14	15
November	13	13	12	11	11
December	10	9	5	7	5
Hospitals	359	359	359	359	359

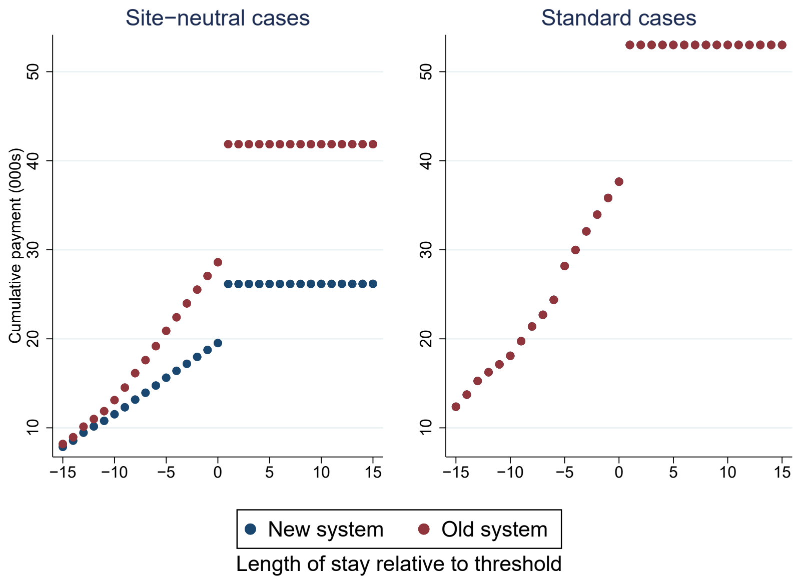
Notes: This table shows the distribution of hospitals according to the month in which the cost reporting period begins. Included: working hospitals that were active during the entire period from FY 2012 to FY 2017.

2012 to FY 2016: many hospitals moved the start date of their cost reporting periods to September, delaying as much as possible the start of the new payment system.

Figure 3.2 shows the average payment schedule for discharges in FY 2017, separately for site-neutral and standard cases. For each type of case, I compute the average payment under the old and new systems. There are several things to note. First, the payment for standard cases is the same under the new and old systems. Second, the average payment under the old system for standard cases is higher than for site-neutral cases, which reflects that the former are higher severity cases. Third, there is a significant reduction in payment for site-neutral cases under the new system. Before the SSO threshold, payment increases approximately linearly with the LOS of the patient at a daily rate of $780 under the new system, while the per diem rate under the old system was approximately $1,350. At the SSO threshold, under the new system hospitals only get 50% of the pre-reform jump in payments. Beyond the SSO threshold, the payment to the hospital is independent of the LOS of the patient under both systems[7]. Table 3.4 shows the distribution of the effective payment cut (at

[7] Under the new system, payments for site-neutral cases beyond the SSO threshold might increase slightly with LOS.

Figure 3.2: Payment schedule, FY 2017 discharges



Notes: This figure shows the average payment schedule for LTCH discharges in FY 2017. The "new system" payment refers to the payment actually received by the hospital and the "old system" payment is the payment that the hospital would have received had the reform not been implemented. Average payment across hospitals and DRGs.

the observed LOS) for site-neutral discharges in FY 2017. On average, site-neutral cases suffered a 30% payment reduction compared to the pre-reform rate; for the median discharge, the payment cut was 36%. Therefore, the adverse financial impact of the new payment system on LTCHs when caring for site-neutral patients is significant.

There are other provisions under the new regulatory system; I mention two that might be relevant for the analysis. First, for cost reporting periods beginning on or after October 1, 2019, at least 50% of Medicare discharges must meet the exclusion criteria in order for a facility to maintain Medicare certification as LTCH. Second, site-neutral cases are excluded from the ALOS computation used to determine LTCH status.

However, the incidence of such cases seems to be very small.

Table 3.4: Variation in payment for site-neutral discharges in FY 2017

	p10	p25	p50	p75	p90	Mean
SSO case: No	-41%	-40%	-39%	-36%	-33%	-38%
SSO case: Yes	-34%	-28%	-13%	-3%	-3%	-15%
Total	-41%	-40%	-36%	-25%	-3%	-30%

Notes: This table shows summary statistics of the distribution of the realized variation in payment for site-neutral cases in FY 2017. The variation in payment refers to the difference between the payment under the new system and the payment under the old system at the observed length of stay of the patient.

3.3. Data

The primary source of data for this study is the LTCH PPS Expanded Modified MEDPAR[8] Limited Data Set files for FY 2014 to 2017[9]. Each file is an administrative dataset containing information on all Medicare LTCH discharges during that fiscal year. Each record is a discharge and it includes characteristics of the patient (age, gender, race, state of residence), quarters of admission and discharge, source of admission, discharge status, the Medicare provider number of the hospital, the DRG to which the inpatient stay was grouped, length of stay, primary and secondary diagnoses, primary and secondary procedures, charges, payments to the hospital, and days from admission to death. In addition, the files indicate whether the discharge meets the criteria for exclusion from the site-neutral rate and the number of days the patient spent in an ICU during her previous hospital stay.

The full dataset contains 592,497 LTCH discharges from 452 hospitals during the period from FY 2014 to 2017. The working dataset includes 536,774 LTCH discharges from 361 hospitals (which I refer to as working hospitals) that: i) were active during the entire period from FY 2015 to 2017; ii) are not government-owned facilities, and; iii) for which there is complete information to construct the payment schedule (see Appendix C). The first restriction implies that I exclude

[8]Medicare Provider Analysis and Review.

[9]For a detailed description of the data, see: https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/LTCHPPSMEDPAR.html

Table 3.5: Summary statistics of LTCH discharges

	Pre-reform	Post-reform
Female	0.48	0.48
White	0.72	0.71
Black	0.20	0.21
Other race	0.08	0.08
Less than 45 y/o	0.04	0.04
45 to 64 y/o	0.23	0.22
65 to 74 y/o	0.34	0.37
75 to 84 y/o	0.25	0.25
More than 85 y/o	0.14	0.12
Diagnoses	17.54	19.32
Procedures	2.23	2.13
Discharged death	0.13	0.13
Discharged upstream	0.13	0.13
Discharged downstream	0.72	0.72
Discharged other	0.03	0.02
Length of stay	28.13	26.68
Payment	36,025	33,170
Other for-profit	0.39	0.43
Select	0.21	0.19
Kindred	0.27	0.26
Nonprofit	0.12	0.11
Co-located	0.29	0.29
Standard case	0.56	0.65
Discharges	137,401	126,025

Notes: This table reports summary statistics of LTCH discharges in the working sample. The pre-reform period corresponds to FY 2015, while the post-reform period is FY 2017.

discharges from a few hospitals that closed before the end of my sample period or that entered the market after October 1, 2015. The second restriction is due to the idiosyncratic nature of public facilities.

For the main analysis, I define four different groups of discharges, depending on whether they: 1) meet the criteria for exclusion from the site-neutral rate, and; 2) occurred before or after the rollout of the new system. In the pre-reform period, there was no distinction between site-neutral and standard cases. When I refer to site-neutral cases in the pre-reform period, I mean cases that would have been classified as such had the patient criteria been in effect in that period. Because site-neutral payments were implemented on a rolling basis depending on each hospital's cost reporting period start date, the pre and post-reform periods are hospital-specific. In the main analysis, I use FY 2015 (2017) as the pre (post) reform period for all hospitals. With this definition, all hospitals were subject to the old payment system in the pre-reform period, while in the post-reform period all hospitals were subject to the new payment system. However, hospitals differ in terms of the amount of time that they have been exposed to the new payment system during the post-reform period. I also performed the analysis using alternative definitions of pre and post-reform periods (see Appendix C); the results (some of them are reported in Appendix D) are very similar to those for the baseline case discussed in the main text. Table 3.5 contains summary statistics of the working sample.

The main limitations of the discharge data are the following. First, I do not observe the zip code nor the county of residence of the patient. Second, I only observe quarters of admission and discharge, but not exact dates. Third, there is very limited information about the previous inpatient stay of an LTCH patient. While I observe the number of days that the patient spent in an ICU during the preceding stay, I do not have diagnosis or treatment information about that episode, nor I observe the identity of the hospital in which the patient was treated. Fourth, there are no patient identifiers, so I cannot link LTCH readmissions to the corresponding index episodes. Fifth, I observe the discharge destination of the patient but I do not have information about post-discharge Medicare spending or utilization.

I link the discharge data to data on hospital characteristics from the Centers for Medicare & Medicaid Services (CMS) Healthcare Cost Report Information System (HCRIS), the CMS LTCH PPS Impact Files, the American Hospital Association (AHA) Annual Survey Database, and company annual reports. In Table 3.6, I report descriptive statistics of my sample of LTCHs. LTCHs are small compared to general hospitals: the average hospital had 61 beds and close to 500 patients during 2016. In terms of type of ownership, 42% of LTCHs are for-profit facilities affiliated with Select or Kindred and 40% are for-profit facilities not affiliated with any of the latter; the rest are nonprofit

Table 3.6: Summary statistics of LTCHs

	Mean	Std. Dev.
Total discharges	497.3	332.9
Beds	61.7	53.2
% Medicare	0.64	0.16
Utilization	0.64	0.17
Other for-profit	0.40	-
Select	0.23	-
Kindred	0.19	-
Nonprofit	0.17	-
Co-located	0.33	-
Hospitals	361	

Notes: This table reports summary statistics of the LTCHs included in the working sample. Hospital characteristics correspond to those indicated in the Medicare cost report of the hospital for FY 2016 (FY 2015 if the latter is not available). Summary statistics are reported across hospitals.

hospitals (I do not include government-owned facilities in the analysis). About one third of LTCHs in my sample are co-located within a general hospital. For the average hospital, Medicare accounts for almost two thirds of all discharges. Capacity utilization is calculated from cost report data as the total number of patient days (Medicare and non-Medicare) divided by the number of bed days available. The average utilization rate across hospitals in my sample was 64% during 2016.

3.4. Descriptive analysis

In this section, I present descriptive evidence of the impact of the payment reform on the admission and discharge decisions of LTCHs. In the analysis, I will be comparing the behavior of hospitals in the pre and post-reform periods. Unless otherwise noted, the pre and post-reform periods correspond to FY 2015 and 2017, respectively. I classify hospitals into four groups (Select, Kindred, other for-profit, and nonprofit) based on the type of ownership indicated in cost reports for FY 2016[10]. Given the way that groups are defined and that my sample of hospitals includes facilities that were

[10]This means, for example, that a hospital that was affiliated with Select in 2015 and was acquired by Kindred in 2017 will be considered as a Select facility in the analysis. The number of movements of hospitals across groups between periods is small.

active during the entire period from FY 2015 to 2017, the hospital composition of the different groups is the same in the pre and post-reform periods. Therefore, post-reform changes at the group level are not due to entry or exit of hospitals nor to changes in ownership. Although the set of hospitals included in a given group is the same pre and post-reform, it is possible that the relative weights of hospitals within each group might have changed with the reform; for example, hospitals that experience the largest increase in the number of discharges will have a higher weight in the post-reform period. I will comment on the impact of this composition effect in the analysis below.

3.4.1. Framework

To guide the analysis, I consider the framework proposed by Einav et al. (2017) and Eliason et al. (2016), adapted to the particulars of my setting. The objective is to provide an interpretive framework for reviewing the empirical results discussed below, which will serve as the basis for a more formal analysis later. A patient is referred to an LTCH. The referral can occur immediately after an inpatient stay in a general hospital, but this is not necessary. Patients have different types defined by: 1) Health status at the time of the referral; 2) The clinical condition for which they seek treatment (DRG), and; 3) Whether the discharge satisfies the criteria from exclusion from the site-neutral payment rate. I assume that the hospital observes all these variables at the time of the referral[11]. The researcher does not observe the health status of the patient.

The LTCH observes the type of the patient and decides whether to admit her or not. The hospital admits the patient if the expected payoff is higher than the expected cost of treatment. If the patient is admitted, the LTCH decides on a course of treatment. Although a course of treatment involves several variables, I assume that it can be summarized by the length of stay[12]. Therefore, the LTCH first faces a decision of whether to admit the patient or not. Once the patient is admitted, the health

[11] Although in practice whether the patient meets the criteria for exclusion from the site-neutral rate might be determined after the patient is admitted, I assume that this is known at the time of admission. In particular, the hospital knows to which DRG the patient is assigned and whether she will receive more than 96 hours of mechanical ventilation.

[12] One dimension in which LTCHs could alter treatment plans is the time the patient spends on mechanical ventilation (see discussion below).

of the patient evolves stochastically over time and the LTCH faces a daily decision of whether to retain the patient, discharge her upstream, or discharge her downstream. The basic trade-off is the following. By discharging the patient, the hospital gets a non-monetary payoff that depends on the health of the patient (it captures the portion of patient utility that is internalized by the hospital). By keeping the patient, the hospital gets a monetary payoff (marginal profit) that depends on the Medicare payment schedule and the marginal cost of treatment, and a non-monetary payoff that captures the evolution of the health of the patient due to the additional treatment. The payment that the LTCH receives depends on the type of the patient and the treatment chosen. The payment schedule faced by the LTCH (given the type of the patient) depends on the regulatory regime to which the LTCH is subject, which in turn depends on the calendar date and the cost reporting period of the LTCH. I assume that the marginal cost of treatment is the same in the pre and post-reform periods, conditional on the type of the patient.

Therefore, LTCHs can react to the decrease in reimbursement on an extensive and an intensive margin. On the extensive margin, LTCHs can refuse to admit site-neutral patients. On the intensive margin, they can modify their discharge behavior to increase reimbursement for these cases.

In the analysis, I assume that the regulatory regime does not affect the pool of patients discharged from GACHs and referred to LTCHs. Then, admission, treatment and discharge (referral) decisions of GACHs are treated as exogenous. In particular, this implies the following. First, there are no changes in the underlying conditions that are being referred for LTCH care. Second, the distribution of patient health at the time of discharge from the GACH is exogenous. Third, GACHs do not alter the treatment plans for their patients in order to manipulate their types. In particular, GACHs do not hold patients longer in the ICU so they qualify for the standard LTCH rate later. Last, it is possible that GACHs decide not to request admission of site-neutral patients to an LTCH anticipating that the latter will turn down these unprofitable patients. In what follows, I consider this as a request followed by a rejection. Note that although the pool of patients who are referred to LTCHs is not affected by the regulatory reform, the pool of patients effectively treated in LTCHs could be affected by the change in the payment policy.

In principle, hospitals can take actions so that patients who otherwise would qualify as site-neutral cases can be categorized as high-acuity patients. Once a patient is admitted to an LTCH, there are three possibilities:

1. The patient did not have a previous inpatient stay, in which case the hospital will be paid the site-neutral rate;

2. The patient had a previous inpatient stay and spent 3 or more days in an ICU, in which case the hospital will be paid the standard rate;

3. The patient had a previous inpatient stay but did not spend at least 3 days in an ICU. If the patient subsequently receives more than 96 hours of mechanical ventilation, the hospital will be paid the standard rate; otherwise, it will receive the site-neutral rate.

Note that exclusion from the lower site-neutral payment rate is mostly dependent upon events that are outside of the LTCH's direct control. In order to modify the type of a patient, the LTCH can do two things. First, it can try to influence treatment decisions (specifically, the number of days spent in the ICU) during the previous hospital stay of the patient. It seems general hospitals are unlikely to artificially extend the number of days that a patient spends in the ICU so the patient qualifies for the higher standard rate in the posterior LTCH stay. The main reason for this is the high opportunity cost of an occupied ICU bed. This opportunity cost likely varies with capacity utilization, as suggested by Kim et al. (2014) and Freedman (2016), so there might situations in which manipulation of types is more feasible. Although I cannot test this assumption directly (given that I do not have data on inpatient stays at general hospitals), I view it as a reasonable approximation. Second, the LTCH could provide additional ventilator time to an admitted patient so she qualifies for the higher standard rate. However, the scope for such behavior is limited for various reasons. First, mechanical ventilation is used to support a single failing organ system (the

Table 3.7: Criteria for exclusion from the site-neutral rate

	Site-neutral cases		Standard cases	
	Pre-reform	Post-reform	Pre-reform	Post-reform
Excluded	0.30	0.35	-	-
None	0.70	0.65	-	-
ICU Only	-	-	0.70	0.73
Ven Only	-	-	0.02	0.01
Both	-	-	0.28	0.26

Notes: This table shows information about the criteria that discharges met or failed to meet, by period and type. "Excluded" refers to psychiatric or rehabilitation cases and to patients without a hospital stay previous to the LTCH admission. "None" refers to "non-excluded" patients who do not meet the ICU nor the ventilator criterion.

lungs), so not all patients are candidates for this procedure. Second, the risk for a patient from unnecessary time on a ventilator is quite high. Third, the procedure codes for mechanical ventilation are not sensitive to more aggressive coding practices. Finally, as discussed below, manipulating the time patients spend on mechanical ventilation would affect a very narrow segment of the LTCH population.

Table 3.7 shows the classification of site-neutral and standard cases before and after the reform according to the exclusion criteria met. Most of the standard cases (in both the pre and post-reform periods) satisfy the ICU criterion, so they already qualify for the standard rate at the time of admission to the LTCH. The patients whose type could have been manipulated by providing additional time on a ventilator are those standard patients who had a previous inpatient stay, do not meet the ICU criterion, but meet the ventilator criterion. These patients accounted for a very small proportion (less than 2%) of standard discharges during the post-reform period.

Therefore, in the analysis I will consider that the type of a patient referred to an LTCH is given from the point of view of the latter. Although deviations from this assumption are certainly possible, the incidence of those cases seems relatively low and therefore I view the assumption as a reasonable approximation.

3.4.2. Incidence of site-neutral cases

I start by analyzing the exposure of LTCHs to the payment reductions brought about by the reform. The incidence of site-neutral cases is defined as the proportion of Medicare discharges that are site-neutral cases. On a pure static sense, the hospitals most affected by the reform are those with a higher incidence of site-neutral cases in the pre-reform period. Next, I analyze how the incidence of site-neutral cases has changed with the implementation of the new payment system in order to assess how payment reform has affected LTCHs' financial incentives to treat Medicare patients.

The overall incidence of site-neutral cases is significant. Looking at the full dataset, approximately 45% of LTCH Medicare discharges in FY 2014 and 2015 would have been classified as site-neutral cases had the new payment system been in effect in those years. After the new payment system was introduced, the incidence of site-neutral cases declined: it was 36% during FY 2017[13].

In the main analysis, I study heterogeneity in the incidence of site-neutral cases across hospitals with different type of ownership. The results are displayed in Table 3.8. The first and second rows of each panel show the distribution of incidence across hospitals in the pre and post-reform periods, respectively. Figures 3.3 and 3.4 provide similar information graphically for the totality of working hospitals. After the rollout of the new payment system, we observe a shift to the left in the distribution of incidence of site-neutral cases. The first (second) quartile shifted from 33% (42%) in the pre-reform period to 13% (33%) in the post-reform period (see Panel 5), while the upper part of the distribution did not experience large changes. However, there is heterogeneity along this dimension across different types of hospitals. In the case of for-profit hospitals other than Kindred and Select (Panel 1) and Kindred hospitals (Panel 3), there is a shift to the left in the lower part of the distribution, while the upper part did not change much. In the case of Select hospitals (Panel 2), there was a large reduction in the incidence of site-neutral cases: while before the reform 90% of Select facilities had an incidence above 20%, after the reform 90% of Select hospitals had an

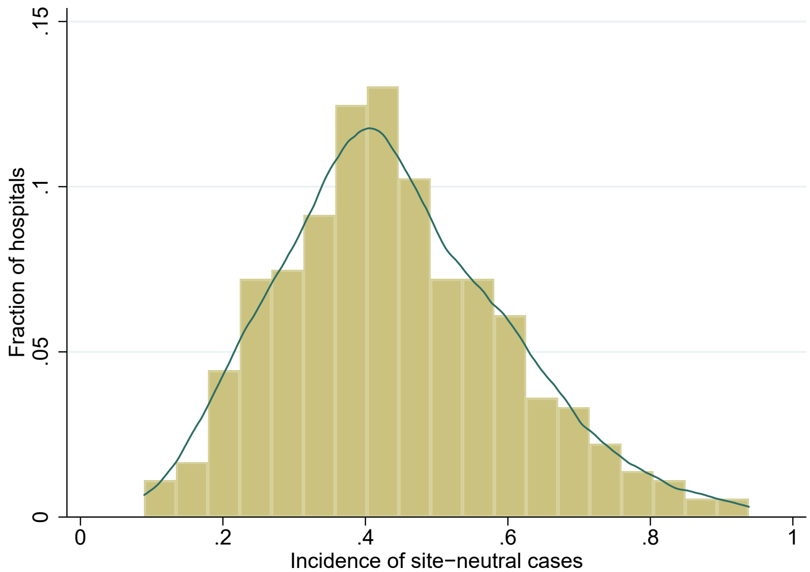
[13] During FY 2016, the incidence of site-neutral cases was 42%. However, the old and new payment systems coexisted during this time period.

Table 3.8: Incidence of site-neutral cases

	p10	p25	p50	p75	p90	Hospitals
1-Other for-profit						
Incidence pre-reform	0.31	0.39	0.51	0.61	0.70	146
Incidence post-reform	0.21	0.32	0.46	0.60	0.71	146
Change incidence (perc)	-0.42	-0.24	-0.07	0.03	0.16	146
Change incidence (points)	-0.18	-0.11	-0.03	0.01	0.08	146
Change neutral cases (perc)	-0.54	-0.37	-0.12	0.12	0.46	146
Change standard cases (perc)	-0.20	-0.09	0.10	0.27	0.54	146
2-Select						
Incidence pre-reform	0.20	0.27	0.36	0.44	0.49	83
Incidence post-reform	0.04	0.05	0.07	0.09	0.11	83
Change incidence (perc)	-0.90	-0.86	-0.81	-0.67	-0.55	83
Change incidence (points)	-0.42	-0.38	-0.28	-0.19	-0.09	83
Change neutral cases (perc)	-0.92	-0.89	-0.84	-0.71	-0.60	83
Change standard cases (perc)	-0.08	0.05	0.21	0.34	0.55	83
3-Kindred						
Incidence pre-reform	0.25	0.32	0.38	0.48	0.59	69
Incidence post-reform	0.18	0.25	0.38	0.48	0.56	69
Change incidence (perc)	-0.35	-0.25	-0.06	0.05	0.18	69
Change incidence (points)	-0.12	-0.09	-0.02	0.02	0.06	69
Change neutral cases (perc)	-0.54	-0.40	-0.16	0.00	0.15	69
Change standard cases (perc)	-0.45	-0.22	-0.09	0.09	0.22	69
4-Nonprofit						
Incidence pre-reform	0.20	0.30	0.42	0.56	0.73	63
Incidence post-reform	0.10	0.19	0.29	0.39	0.53	63
Change incidence (perc)	-0.62	-0.43	-0.30	-0.08	0.03	63
Change incidence (points)	-0.26	-0.21	-0.13	-0.03	0.01	63
Change neutral cases (perc)	-0.71	-0.61	-0.39	-0.18	-0.08	63
Change standard cases (perc)	-0.29	-0.15	0.05	0.21	0.40	63
Total						
Incidence pre-reform	0.24	0.33	0.42	0.54	0.65	361
Incidence post-reform	0.06	0.13	0.33	0.48	0.61	361
Change incidence (perc)	-0.83	-0.55	-0.21	-0.02	0.12	361
Change incidence (points)	-0.33	-0.19	-0.09	-0.01	0.05	361
Change neutral cases (perc)	-0.86	-0.65	-0.32	-0.04	0.19	361
Change standard cases (perc)	-0.26	-0.11	0.08	0.26	0.46	361

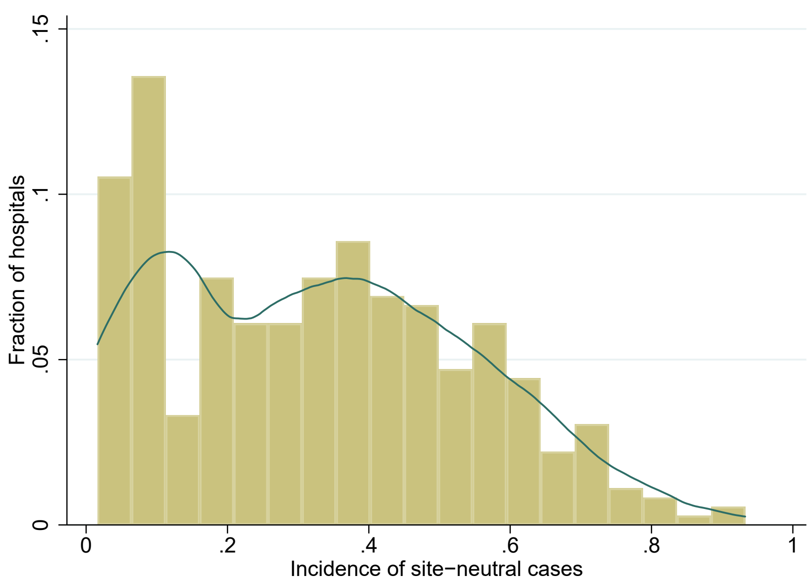
Notes: This table shows summary statistics of the distribution of the incidence of site-neutral cases across hospitals.

Figure 3.3: Incidence of site-neutral cases, pre-reform



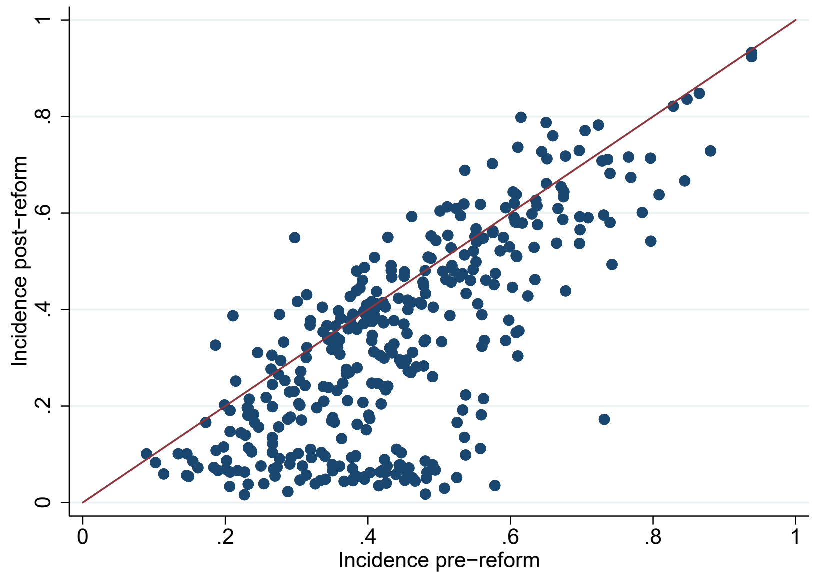
Notes: This figure shows the distribution of the incidence of site-neutral cases across hospitals during the pre-reform period.

Figure 3.4: Incidence of site-neutral cases, post-reform



Notes: This figure shows the distribution of the incidence of site-neutral cases across hospitals during the post-reform period.

Figure 3.5: Incidence of site-neutral cases



Notes: This figure shows the incidence of site-neutral cases in the pre and post-reform periods for each hospital. The 45° line is included for reference.

incidence below 11%. Between these cases, nonprofit hospitals (Panel 4) show a general decrease in incidence, but not as striking as for Select facilities[14].

Next, I analyze the change in incidence at the hospital level, measured both as percentage of the pre-reform incidence and in percentage points. The results are shown in the third and fourth rows of each panel of Table 3.8. Figure 3.5 provides similar information graphically for the totality of working hospitals. We observe that the decline in incidence was generalized in the case of Select and nonprofit facilities. In the case of Kindred hospitals and for-profit hospitals other than Select and Kindred, the median hospital experienced a modest reduction in incidence, although incidence increased in some facilities.

Finally, I examine how the number of site-neutral and standard cases changed with the reform. The results are shown in the fifth and sixth rows of each panel of Table 3.8. The main point is that in general the decline in incidence of site-neutral cases at the hospital level is explained by a decrease

[14]Similar conclusions arise if I weight hospitals by the number of Medicare admissions (not reported).

in the number of site-neutral cases while the number of standard cases either grows or declines at a lower rate than the former.

Although type of ownership is an important dimension of hospital heterogeneity, there is substantial variation within groups. To have a better sense of the effects of the payment reform on different types of hospitals, I regress various measures related to the incidence of site-neutral cases at the hospital level on facility characteristics. The results are displayed in Table 3.9. Different columns correspond to different dependent variables. During the pre-reform period, Select, Kindred and nonprofit facilities had a lower incidence of site-neutral cases than other for-profit facilities. In the post-reform period, we observe that the differences in incidence between other for-profit hospitals, on one hand, and nonprofit and specially Select facilities, on the other hand, are much more prominent than before the reform. Similar conclusions arise if we look at the change in incidence or number of site-neutral cases. In terms of other dimensions of hospital heterogeneity, I do not observe differences in incidence of site-neutral cases between freestanding and co-located hospitals, or between hospitals of different size. However, I observe that hospitals with higher capacity utilization in the pre-reform period experience a larger decrease in the incidence and number of site-neutral cases post-reform. If capacity constraints were binding in the pre-reform period, then the hospital was forced to reject some patients (both site-neutral and standard cases, which were similar in terms of Medicare reimbursement during that period). In the post-reform period, the hospital can deal with capacity constraints by admitting standard patients who were rejected before in detriment of site-neutral patients, which leads to a decrease in the incidence of site-neutral cases[15].

Under my assumptions about the behavior of GACHs, the decrease in the incidence and number of site-neutral cases documented above means that LTCHs are rejecting site-neutral patients who were admitted before the reform. Therefore, the policy not only impacts directly into providers'

[15] To control more directly for the impact of capacity constraints on the admitting behavior of LTCHs, I could exploit within-year variation in capacity utilization. Due to data limitations, I cannot implement this approach.

Table 3.9: Regression of incidence of site-neutral cases on hospital characteristics

	(1) Incidence pre-reform	(2) Incidence post-reform	(3) Change incidence (%)	(4) Change incidence (pp)	(5) Change neutral cases	(6) Change std cases
Select	-0.141***	-0.356***	-0.620***	-0.217***	-0.721***	0.079
	[0.019]	[0.017]	[0.032]	[0.018]	[0.046]	[0.047]
Kindred	-0.105***	-0.089***	0.017	0.016	-0.142**	-0.209***
	[0.020]	[0.023]	[0.031]	[0.013]	[0.051]	[0.043]
Nonprofit	-0.065*	-0.147***	-0.178***	-0.081***	-0.351***	-0.080
	[0.027]	[0.026]	[0.040]	[0.018]	[0.053]	[0.046]
Co-located	-0.035*	-0.026	0.003	0.007	0.040	0.025
	[0.017]	[0.017]	[0.029]	[0.014]	[0.042]	[0.035]
Beds	-0.000	-0.000	-0.000	0.000	0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Utilization	-0.055	-0.191**	-0.298***	-0.109**	-0.260*	0.128
	[0.054]	[0.059]	[0.078]	[0.034]	[0.113]	[0.103]
Constant	0.558***	0.598***	0.099	0.022	0.111	0.044
	[0.041]	[0.045]	[0.055]	[0.024]	[0.099]	[0.091]
Observations	361	361	361	361	361	361
Adjusted R^2	0.158	0.506	0.586	0.443	0.389	0.092

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows the results of hospital-level regressions of measures related to the incidence of site-neutral cases on hospital characteristics. Different columns correspond to different dependent variables. The omitted type of control category is for-profit hospitals other than Kindred and Select. Utilization refers to the capacity utilization rate of the hospital in the previous period.

bottom lines, but it also makes LTCHs turn away patients who were previously treated in these facilities and that now must seek treatment in less sophisticated settings[16].

3.4.3. Discharge patterns

Following Kim et al. (2015), Eliason et al. (2016) and Einav et al. (2017), I study the pattern of discharges as a function of length of stay relative to the SSO threshold. The studies mentioned above consider the pre-reform period and provide evidence that the financial incentives created by Medicare's LTCH PPS influence the timing of LTCHs' discharge decisions. The main findings of these studies are that LTCHs: 1) increase the length of stay of patients under the threshold so they are paid the full MS-LTC-DRG rate, and; 2) discharge patients earlier once they cross the threshold in order to minimize costs (given that the marginal payment is zero after the threshold). These results indicate a strong response of hospital behavior to financial incentives.

In order to analyze the pattern of discharges, I construct two measures for each day. First, I compute the proportion of patients who are discharged downstream that day. Second, I compute the discharge hazard rate: the probability that a patient is discharged downstream that day, conditional on the patient not having been discharged so far. For example, consider the case where the normalized LOS is equal to 1. If the first measure is 0.1, then it means that 10 percent of discharges take place the first day after the SSO threshold. If the second measure is 0.1, then it means that 10 percent of the discharges that take place after the threshold occur the first day after the threshold. I focus on patients who are discharged downstream because it is easiest for hospitals to manipulate the timing of these discharges based on financial incentives.

The objective is to analyze whether the pattern of discharges changed with the introduction of the new system. I distinguish between site-neutral and standard cases. For each type of case, I compare discharge behavior in the pre and post-reform periods, using both the proportion of discharges and

[16] A more direct approach to analyze the impact of payment reform on access to care would be to model the probability that a patient is discharged to an LTCH after a stay in a general acute care hospital.

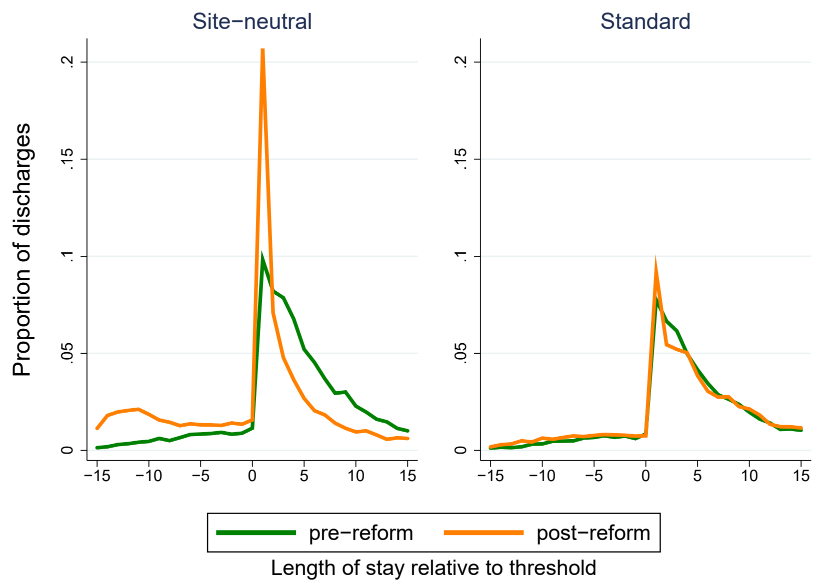
the discharge hazard rate for the analysis. To further link the change in discharge behavior to the payment reform, I show in Appendix D that discharge patterns for site-neutral cases in FY 2014 and 2015 are almost identical for all types of hospitals. The discharge patterns corresponding to the pre-reform period have been previously documented, while the comparison of discharge patterns for different types of patients and regulatory regimes is, to the best of my knowledge, new. In addition, in Appendix E I expand the analysis of pre-reform cases by comparing discharge patterns of Medicare and non-Medicare LTCH patients in Florida and Texas.

Figures 3.6 to 3.13 show the change in discharge patterns for different types of hospitals, based on type of ownership. In the main specification, I use a direct pre/post-reform comparison. To account for the possibility that changes in discharge patterns are due to composition effects, I repeat the analysis reweighting the post-reform cases so they reflect the DRG-hospital composition of the pre-reform period. The results (reported in Appendix D) are very similar to the baseline case considered here.

The main descriptive results of this section are the following. First, the discharge patterns for site-neutral and standard cases in the pre-reform period are very similar. Second, hospitals modified their discharge behavior in the post-reform period differently for site-neutral and standard cases. On one hand, the pattern of discharges for standard cases did not change much with the introduction of the new payment regime. This is not surprising given that the new system did not change the form LTCHs are reimbursed for these cases. On the other hand, there is heterogeneity across different types of hospitals in the way the pattern of discharges for site-neutral cases changed with the introduction of the new regime.

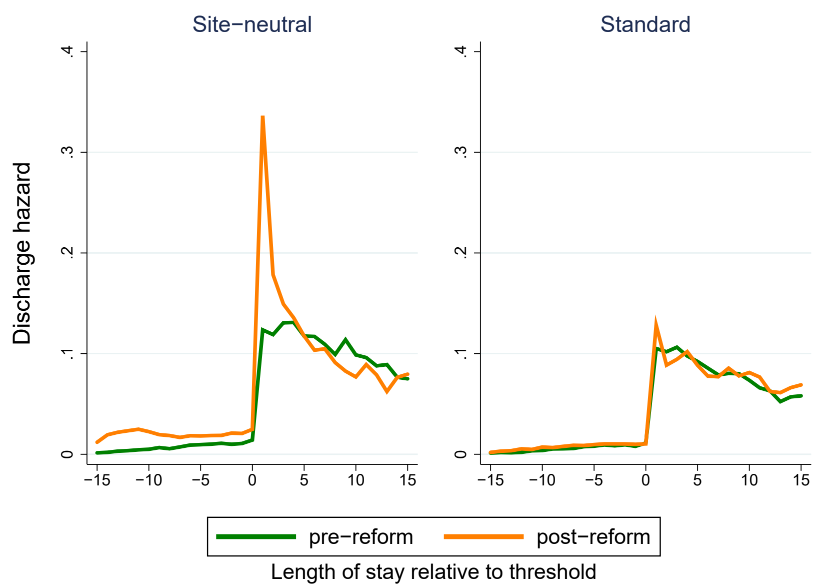
In the case of for-profit hospitals other than Select and Kindred, both the proportion of discharges and the discharge hazard rate before the threshold increased with respect to the pre-reform period. The discharge hazard rate for the day immediately after the threshold more than doubles compared to the pre-reform level: conditional on continuing in the hospital after the threshold, the probability that a site-neutral patient is discharged downstream the first day after it is above 0.30 after the rollout

Figure 3.6: Proportion of discharges, other for-profit hospitals



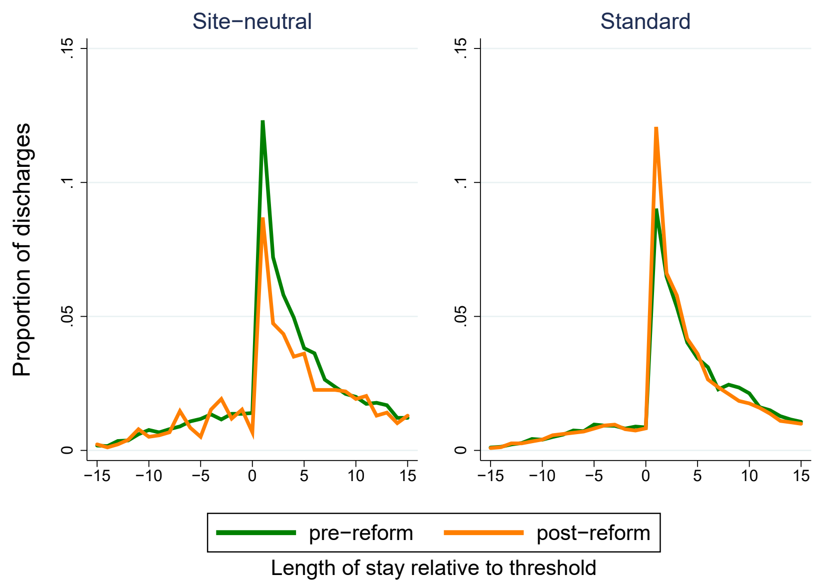
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure 3.7: Discharge hazard rate, other for-profit hospitals



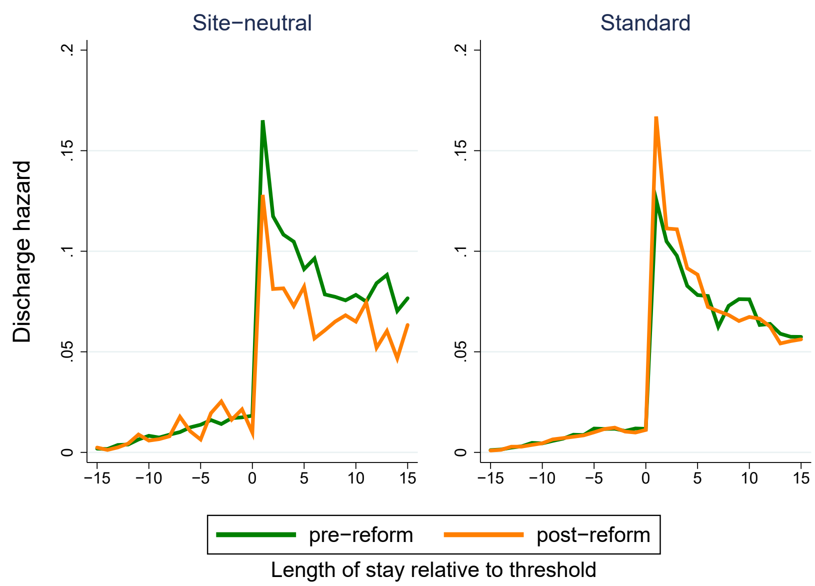
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure 3.8: Proportion of discharges, Select hospitals



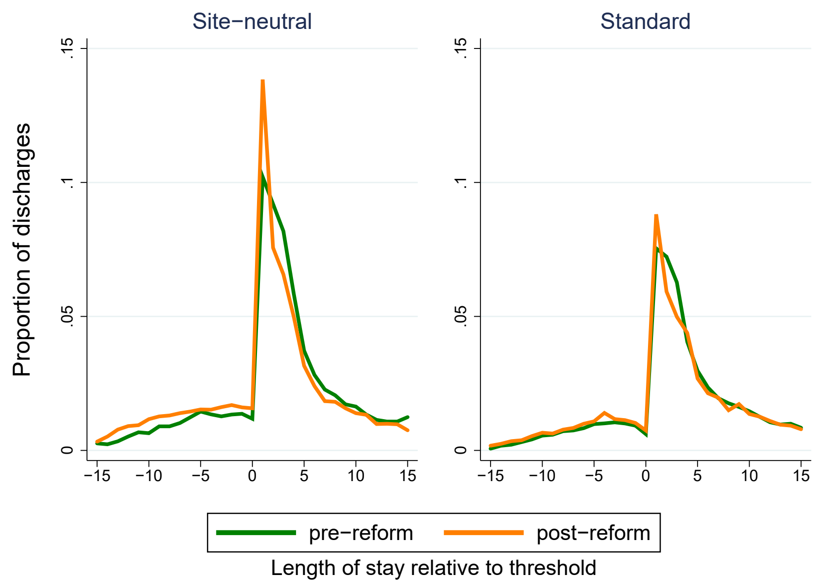
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure 3.9: Discharge hazard rate, Select hospitals



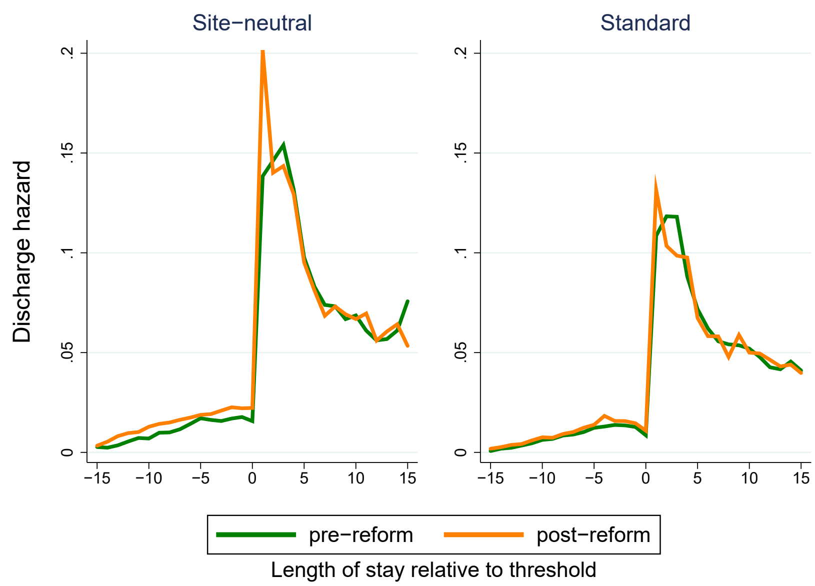
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure 3.10: Proportion of discharges, Kindred hospitals



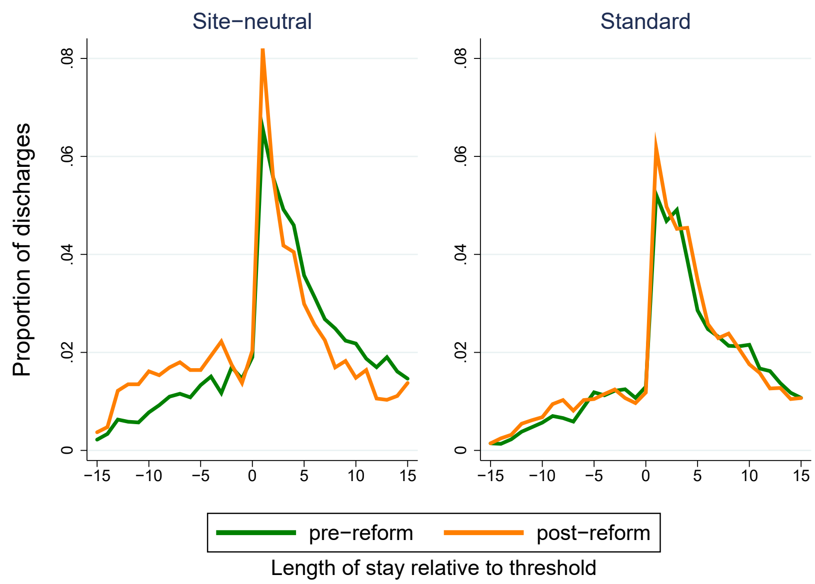
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure 3.11: Discharge hazard rate, Kindred hospitals



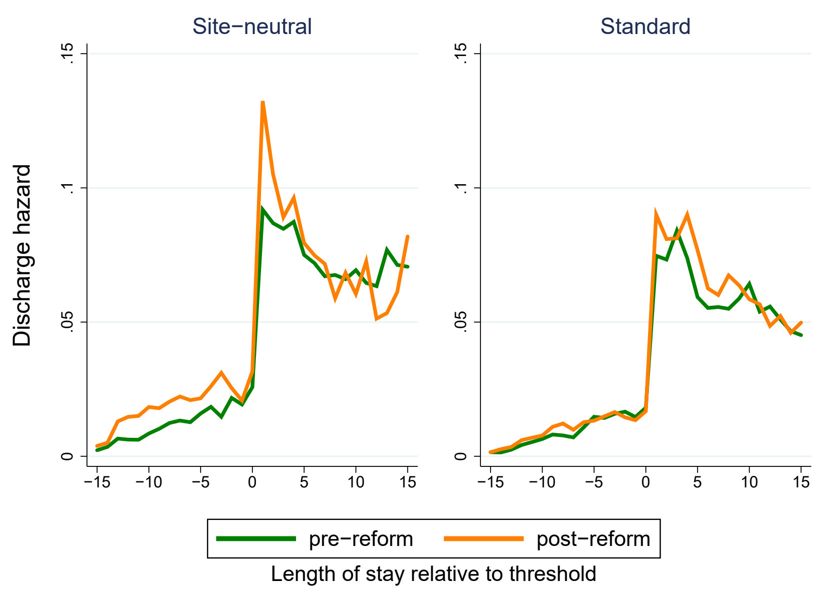
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure 3.12: Proportion of discharges, nonprofit hospitals



Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold) .

Figure 3.13: Discharge hazard rate, nonprofit hospitals



Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

of the new system, compared to 0.12 before the reform. Discharge hazard rates stay relatively high for about five days after the threshold before returning to pre-reform levels.

In the cases of nonprofit and Kindred hospitals, the discharge hazard rate before and immediately after the threshold is higher post-reform. However, the difference in discharge behavior after the threshold compared to the pre-reform period is not as striking as for other for-profit hospitals.

In the case of Select hospitals, the discharge hazard rate before the threshold is similar in both periods but they are less likely to discharge patients downstream after the threshold under the new payment regime.

How did the incentives to retain a patient change with the reform? Consider a patient of given health and assume that the evolution of the patient's health as a function of LOS did not change with the rollout of the new system. Before the threshold, the incentives to retain the patient in the hospital have a static and a dynamic component. The static component is given by the additional payment the hospital receives from Medicare (net of marginal cost, which I assume is the same in both periods), which is lower under the new payment system. The dynamic component reflects the possibility of keeping the patient after the threshold and receive a large jump in payment, which is also lower post-reform. Then, the incentives to retain a patient before the threshold are lower than before, and therefore we should expect an increase in discharges in this region. After the threshold, marginal revenue is zero pre and post-reform. Under the assumption that marginal costs did not change, then the marginal profit is the same in both periods. Consider a patient who remains at the LTCH after the threshold (the hospital already received the large jump in payments). If LTCHs make decisions based on comparing the marginal profit and the impact on the patient's health of an extra day in the hospital, then we should expect that the discharge hazard rate after the threshold does not change post-reform. Given this and the expected increase in SSO cases, we should observe a decline in the proportion of discharges that take place after the threshold.

The analysis above was simplistic in many dimensions. First, I compared the incentives that hospitals face after the threshold under both payment systems for the same patient. However, the pool of patients who remain at the hospital after the threshold might have changed with the reform. I explore the possibility that selection is driving the change in behavior below. Second, there are many factors that are not considered in the analysis. For example, I ignore heterogeneity in the change in the payment schedule across hospitals and DRGs. In Section 3.5, I show that the change in discharge patterns is not in accordance with the predictions from a structural model of hospital behavior that accounts for these and other determinants of discharge decisions. Third, I have assumed that discharge decisions are driven by marginal profits and therefore that the marginal utility of profits is constant. I explore the validity of this assumption below.

In summary, the discharge behavior of LTCHs for standard cases did not change with the reform. In relation to site-neutral cases, Select, Kindred and nonprofit facilities did not adjust their behavior much with respect to the pre-reform period, while other for-profit facilities are more likely to discharge patients before the threshold but they are also more likely to discharge patients immediately after the threshold. The latter result is difficult to explain within the framework discussed in Section 3.4.1 given the way that the payment reform modified the incentives of hospitals.

Next, I explore alternative hypotheses to explain the spike in discharges that take place immediately after the threshold under the new payment system. For the analysis below, I construct different measures of the change in discharge behavior of a hospital at the threshold. For each hospital, I consider: 1) The variation in the proportion of patients discharged downstream immediately after the threshold between the pre and post-reform periods, both in absolute and relative terms (as a proportion of the pre-reform proportion); 2) The variation in the downstream discharged hazard rate immediately after the threshold between the pre and post-reform periods, both in absolute and relative terms (as a proportion of the pre-reform hazard). A positive value for any of these variables should be interpreted as the hospital increasing its strategic behavior in the post-reform period. I construct these measures separately for site-neutral and standard cases. In what follows, unless otherwise noted, the change in the discharge behavior of a hospital refers to the variation in discharge behavior for site-neutral cases.

Hypothesis 1: Change in certification requirements.— In order to keep LTCH certification, the Medicare ALOS of a hospital must be 25 days or longer in the most recent cost reporting period. Before the reform (cost reporting periods starting before October 1, 2015), all Medicare patients were included in the ALOS calculation. After the reform (cost reporting periods starting on or after October 1, 2015), site-neutral cases are excluded from this calculation.

As the previous literature has shown, LTCHs manipulate the timing of discharges for financial reasons. The analysis assumes that LTCHs maximize the payoff from treating patients subject to the payment schedule. However, if the optimal discharge rule induces an ALOS lower than 25 days, the hospital will adjust its discharge behavior so the LTCH certification requirement is met. Therefore, if in the pre-reform period the LOS of patients was extended (beyond the optimal point) to meet the certification requirement, then we should expect that the LOS of site-neutral patients decreases post-reform. This could explain the observed changes in discharge patterns, considering that the extra mass of discharges at the threshold is drawn from the right of the distribution of discharges.

To explore this possibility, I compute the ALOS for each hospital during the last pre-reform cost reporting period. There are various reasons why we should not expect that LTCHs set the ALOS exactly to 25 days if the optimal discharge policy does not meet the certification requirement. However, the higher the excess ALOS with respect to the 25 days mark, the lower the distortions introduced in the discharge decisions of the hospital compared to the first best. I then test whether the change in the discharge behavior of the hospital is related to different measures of how close the ALOS of the hospital was to the 25 days mark in the pre-reform period, after controlling for other hospital characteristics. Table 3.10 shows the results of regressions that include the raw average length of stay of the hospital in the pre-reform period as a covariate. The results do not lend much support to the hypothesis that the change in certification requirements can explain the change in hospital discharge behavior. I have tried several specifications for defining hospitals with

Table 3.10: Impact of average length of stay on the discharge behavior of LTCHs

	(1) Change prop (pp)	(2) Change prop (%)	(3) Change haz (pp)	(4) Change haz (%)
ALOS	-0.003	-0.019	-0.005*	-0.030
	[0.002]	[0.015]	[0.002]	[0.017]
Select	-0.102***	-0.453**	-0.212***	-0.740***
	[0.021]	[0.156]	[0.026]	[0.186]
Kindred	-0.051*	-0.400*	-0.129***	-0.647***
	[0.022]	[0.162]	[0.027]	[0.192]
Nonprofit	-0.027	-0.117	-0.077**	-0.195
	[0.023]	[0.173]	[0.029]	[0.205]
Co-located	-0.029	-0.350**	-0.025	-0.365*
	[0.018]	[0.131]	[0.022]	[0.155]
Beds	-0.000	-0.001	-0.000	-0.002
	[0.000]	[0.001]	[0.000]	[0.001]
Constant	0.130*	1.089**	0.318***	1.789***
	[0.055]	[0.403]	[0.067]	[0.478]
Observations	361	357	361	357
Adjusted R^2	0.078	0.055	0.195	0.085

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows the results of hospital-level regressions of the change in discharges at the threshold for site-neutral patients on the ALOS of the hospital in the pre-reform period and other facility characteristics.

ALOS close to 25 days, such as including indicators for hospitals in different quintiles of the ALOS distribution in the pre-reform period as covariates, and the results are consistent.

Hypothesis 2: Composition effect.— In the analysis above, I compared the discharge behavior of LTCHs in a given day under both payment systems assuming that the patient was the same in both periods. However, it might be the case that the pool of patients who remain at the LTCH after crossing the threshold is different (in terms of health status) in the pre and post-reform periods.

Table 3.11: Health profile of site-neutral patients

	For-profit		Select		Kindred		Nonprofit	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
No previous stay	0.33	0.35	0.22	0.63	0.29	0.30	0.22	0.24
ICU days = 0	0.56	0.54	0.63	0.28	0.58	0.57	0.64	0.61
ICU days = 1	0.05	0.05	0.06	0.03	0.05	0.05	0.06	0.06
ICU days = 2	0.06	0.06	0.08	0.06	0.07	0.07	0.08	0.08
Discharged downstream	0.82	0.84	0.81	0.71	0.80	0.80	0.79	0.79
Discharged upstream	0.11	0.11	0.11	0.17	0.10	0.11	0.13	0.14
In-hospital mortality	0.07	0.05	0.08	0.12	0.09	0.08	0.07	0.07
30-day mortality	0.10	0.09	0.11	0.13	0.11	0.10	0.10	0.10
90-day mortality	0.18	0.18	0.20	0.22	0.20	0.20	0.18	0.17

Notes: This table shows summary statistics of the health profile of all site-neutral patients, by period and type of hospital.

This can be due to two things: 1) The composition of admitted patients changed: LTCHs might screen site-neutral patients in the post-reform period and selectively admit those patients whose expected length of stay coincides with the window of maximum profitability; 2) The evolution of the cross-sectional distribution of health over time changed due to the different discharge behavior of LTCHs before the threshold: even if the pool of admitted patients is the same, the change in discharge behavior implies that by the time the threshold is crossed, the patient population that remains at the LTCH is different.

To explore this possibility, I first compare the health profile of site-neutral patients treated in LTCHs before and after the implementation of the new payment system, by type of hospital. Table 3.11 compares the pool of all admitted patients in each period, while Table 3.12 compares the health profile of patients who remain at the LTCH after the threshold is crossed. The first four variables capture the health of the patient at the time of admission: indicators for patients with no previous inpatient stay and for patients with a previous stay and 0, 1 or 2 days spent in the ICU during that episode. The other five variables capture the health of the patient at the time of discharge: indicators for patients discharged downstream, patients discharged upstream, patients who died in

Table 3.12: Health profile of site-neutral patients who remain at the LTCH after the threshold

	For-profit		Select		Kindred		Nonprofit	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
No previous stay	0.33	0.33	0.21	0.60	0.28	0.29	0.21	0.23
ICU days = 0	0.57	0.56	0.65	0.30	0.59	0.58	0.65	0.61
ICU days = 1	0.04	0.05	0.06	0.03	0.05	0.05	0.06	0.07
ICU days = 2	0.06	0.06	0.08	0.06	0.07	0.07	0.08	0.09
Discharged downstream	0.90	0.90	0.89	0.84	0.89	0.88	0.87	0.86
Discharged upstream	0.07	0.07	0.07	0.10	0.07	0.08	0.09	0.10
In-hospital mortality	0.03	0.03	0.04	0.06	0.05	0.05	0.04	0.04
30-day mortality	0.08	0.08	0.09	0.11	0.09	0.09	0.08	0.09
90-day mortality	0.17	0.16	0.19	0.21	0.19	0.19	0.17	0.16

Notes: This table shows summary statistics of the health profile of site-neutral patients who remain at the LTCH after the SSO threshold, by period and type of hospital.

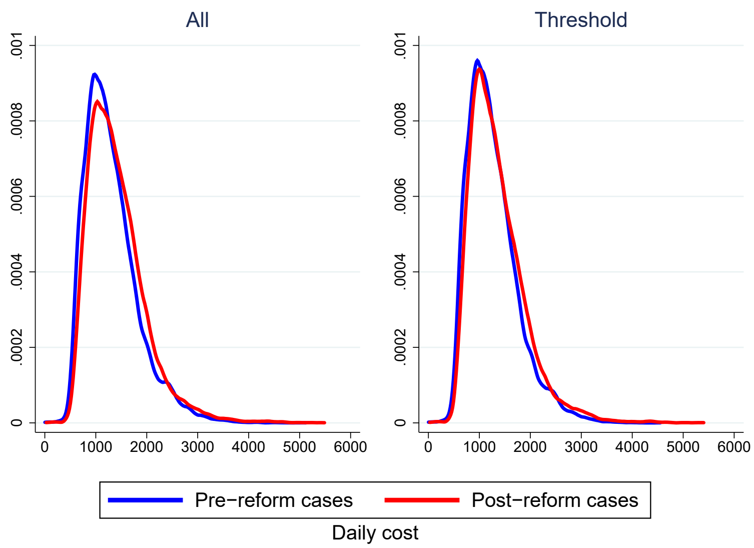
the hospital, and patients who died within 30 and 90 days of a live discharge from the LTCH. Table 3.11 indicates, for example, that before the reform 56% of all site-neutral patients treated in other for-profit hospitals had an inpatient stay with no ICU days immediately before admission to the LTCH, while this percentage was 54% for post-reform cases.

The health profile of patients is very similar in the pre and post-reform periods. The case of Select hospitals seems to be the exception: post-reform site-neutral patients in these hospitals seem to be more severely ill than patients in the pre-reform period. It should be noted, however, that Select hospitals account for a very small fraction (4.2%) of site-neutral cases in the post-reform period.

Second, I compare pre and post-reform site-neutral cases in terms of the estimated daily cost of the case. Figures 3.14 to 3.17 show the pre and post-reform distributions for different types of hospitals and patients. In line with the previous result, pre and post-reform cases are very similar in terms of their estimated daily cost (except for discharges from Select hospitals), suggesting that selection does not seem to account for the observed change in discharge patterns.

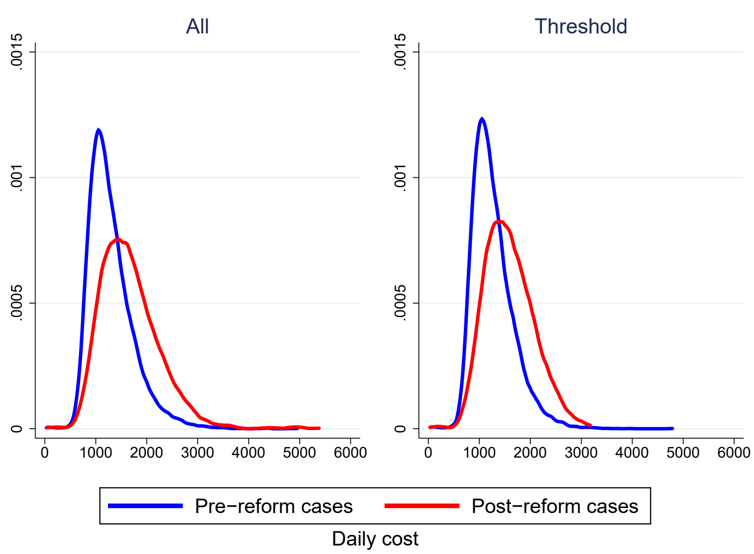
In summary, the data provides limited evidence of changes in the composition of the patient

Figure 3.14: Distribution of estimated daily cost, site-neutral cases, other for-profit hospitals



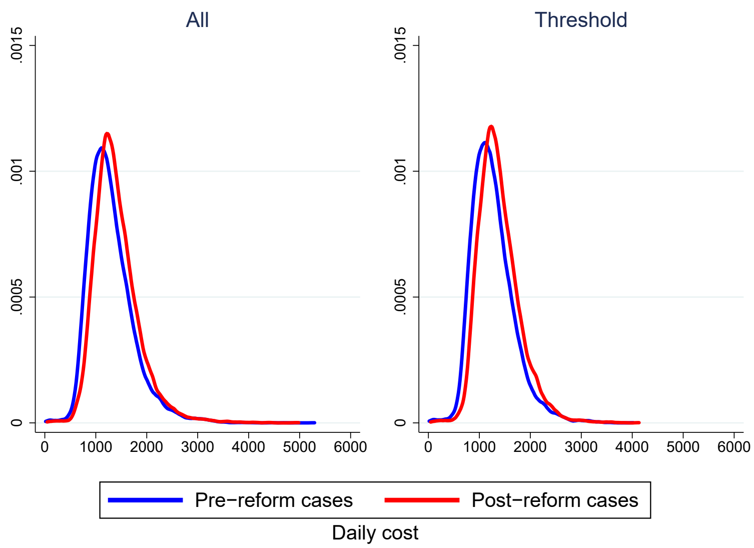
Notes: This figure shows the distribution of estimated daily cost for pre and post-reform site-neutral cases. The left figure shows the distribution for all admitted patients, while the right figure shows the distribution for those patients who were discharged after the SSO threshold.

Figure 3.15: Distribution of estimated daily cost, site-neutral cases, Select hospitals



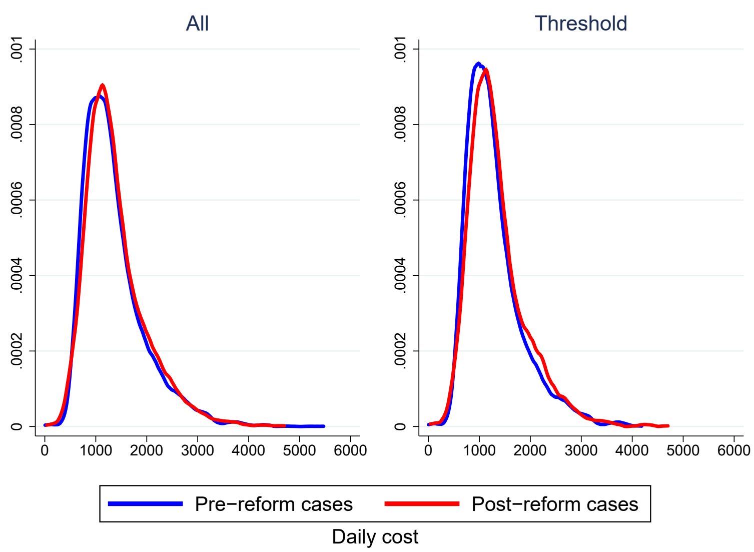
Notes: This figure shows the distribution of estimated daily cost for pre and post-reform site-neutral cases. The left figure shows the distribution for all admitted patients, while the right figure shows the distribution for those patients who were discharged after the SSO threshold.

Figure 3.16: Distribution of estimated daily cost, site-neutral cases, Kindred hospitals



Notes: This figure shows the distribution of estimated daily cost for pre and post-reform site-neutral cases. The left figure shows the distribution for all admitted patients, while the right figure shows the distribution for those patients who were discharged after the SSO threshold.

Figure 3.17: Distribution of estimated daily cost, site-neutral cases, nonprofit hospitals



Notes: This figure shows the distribution of estimated daily cost for pre and post-reform site-neutral cases. The left figure shows the distribution for all admitted patients, while the right figure shows the distribution for those patients who were discharged after the SSO threshold.

pool that could lead to the large change in discharge behavior documented above. However, the previous analysis ignores the possibility of selection based on unobserved health and therefore a more detailed analysis is needed; I comment on this in Section 3.5.

Hypothesis 3: Decreasing marginal utility of profits.— So far, I have focused on the incentive effects of changes in marginal reimbursement, but the level of reimbursement might also affect hospital behavior. The literature on the effects of fee changes on provider behavior (see, for example, McGuire and Pauly (1991), Hodgkin and McGuire (1994), and Gruber and Owings (1996), among others) suggests that the average payment level might have a significant effect on hospital utilization even when marginal incentives remain constant. In particular, many studies have concluded that lowering fees will raise treatment intensity. Some of these "perverse" effects documented by the literature are linked to the existence of income effects[17]. For example, induced demand models suggest that providers may increase inducement to mitigate the effect of negative income shocks. In my context, the negative income shock is produced by the decrease in reimbursement for site-neutral cases and inducement is associated with the strategic discharge of patients during the window of maximum profitability for the hospital.

In the analysis above, I have assumed that hospitals are pure profit maximizing entities. With full profit maximization, there are no income effects and therefore a change in the level of profits (keeping marginal profit constant) does not change behavior. In a more general utility maximizing model, this does not need to be the case given that there are income effects. The strength of the income effect is associated with the behavior of the marginal utility of profits.

Although the marginal profit after the threshold did not change with the reform, the *level* of profits is lower. Consider a patient of given health immediately after crossing the threshold (so the hospital has already received the jump in payments). In both periods, the hospital loses the marginal cost c by keeping the patient an additional day. However, by keeping the patient an extra

[17] For examples outside this literature, see Camerer et al. (1997), Farber (2005), and Fehr and Goette (2007).

day, the health of the patient might change and this could lead to a higher utility when discharging the patient later. Under the old payment system, the level of profits of the hospital is high, so losing the marginal cost has a relatively minor impact: although profits are eroded with the increase in LOS, they remain at a high level for some days. Under the new system, the level of profits is much lower than before, so the relative effect of losing the marginal cost is higher: the boost in profits from the jump in payment at the threshold is exhausted quickly with the increase in LOS. Therefore, if the profit-utility schedule is concave (the marginal utility of profits is decreasing), the utility loss from retaining the patient an extra day in the hospital increased post-reform. In this scenario, the hospital has stronger incentives to discharge the patient soon after the threshold under the new system. This would explain the spike in the discharge hazard rate immediately after the threshold that I documented above.

To explore this possibility, I need to test the assumption that hospitals simply maximize profits. Ultimately, it is a matter of the magnitude of the income effect. Therefore, I look for empirical support for the idea that income effects are important in my setting. Given that income effects occur at the hospital level, I test this hypothesis by relating the change in discharge behavior brought about by the reform to hospital characteristics that proxy for the relevance of income effects.

The first set of variables are related to the financial performance of the hospital. An increase in outside income or an improvement in the financial position of the hospital should affect treatment decisions in a utility maximizing model, while in a profit-maximizing model these variables have no effect. The idea is that in a general utility maximizing model the incentive to manipulate the timing of discharges for financial reasons is higher for hospitals with a greater need for additional profits. As measures of financial distress, I use the debt to asset ratio of the hospital and the fraction of the hospital's discharges that are reimbursed by Medicaid.

Second, income effects will be most pronounced for hospitals with a large exposure to the payment reductions brought about by the reform. Therefore, the second variable of interest is the share of the hospital's patients that are Medicare site-neutral cases. This variable should not be relevant if hospitals only maximize profits, but it would have an impact on discharge behavior in the context of utility maximization.

Another way to test for profit maximization is to consider revenue-side cross-payer and cross-segment effects (Hodgkin and McGuire, 1994). Changes to the payment schedule for Medicare site-neutral cases should have no effect on Medicare standard cases and privately insured patients under profit maximization. This is not the case in a utility maximization model. In the latter world, the hospital will seek to mitigate the impact of the payment reform by increasing revenue from all sources, not only from site-neutral patients. As mentioned in Section 3.3, I do not have data on LTCH privately insured patients, so I cannot examine the impact of payment reform on this segment. However, it seems that the possibility of extracting higher profits from these patients is rather limited. First, from company reports it seems that patients covered by non-government payers are generally more profitable than those covered by Medicare, so competition for those patients is intense. Second, although I do not have information about the form of the payment schedule for privately insured patients, it seems unlikely that it features the sort of non-linearities present in the Medicare payment schedule. In Appendix E, I examine pre-reform discharges from all payers in Texas and Florida and show that distortions in the timing of discharge are much smaller in the case of privately insured patients, which lends support to the second point above.

I run a hospital-level regression of the change in behavior brought about by the reform on the variables mentioned above and other hospital characteristics. The results are displayed in Table 3.13. The coefficients associated with the two measures of financial distress considered are statistically insignificant in all but one specification. The coefficient associated with the fraction of the hospital's patients who are Medicare site-neutral cases is positive and statistically significant in all but one specification. I have tried other specifications for defining hospitals with a high incidence of Medicare site-neutral cases or hospitals in financial distress; the qualitative results are consistent with the ones presented here. To explore whether the payment reform affected the discharge behavior of hospitals for standard cases, I run a similar regression but using the change in discharge behavior for standard cases as the dependent variable. The results are displayed in Table 3.14. The coefficient associated with the fraction of the hospital's patients who are Medicare site-neutral cases is positive in all specifications but statistically significant in only one of them. The coefficient associated with the debt to asset ratio is positive and statistically significant in two specifications.

The results offer mixed evidence about the merits of the hypothesis that income effects explain the observed change in discharge behavior. First, the nature of the results depends on the specific measure of change in behavior considered. Second, an important portion of the heterogeneity in the change in discharge behavior across hospitals remains unexplained. Third, the observation that discharge patterns for standard cases are very similar before and after the rollout of the new payment system suggests that hospitals behave as profit maximizing entities.

3.5. Structural analysis

In this section, I analyze the discharge behavior of LTCHs using the dynamic structural model proposed by Eliason et al. (2016). The basic setting is the following. First, a patient requests admission to an LTCH and the hospital decides whether to admit her or not. If the patient is admitted, the LTCH faces a daily decision of whether to retain the patient or discharge her. The decision is based on the financial incentives provided by the Medicare payment schedule and other non-monetary factors such as the effect of treatment on the health of the patient.

I start by describing the model. I follow the original formulation as close as possible given my data. The main objective of the analysis is to quantify the impact of financial incentives on hospital discharge behavior. First, I use data on pre-reform cases to estimate the parameters of the model and analyze how discharge patterns would change under the new payment schedule. This exercise assumes that the pool of admitted patients is not affected by the reform and therefore is useful to analyze how the payment change would affect patients who were treated at LTCHs in the pre-reform

Table 3.13: Regression of variation in discharge patterns for site-neutral cases on hospital characteristics

	(1) Change prop (pp)	(2) Change prop (%)	(3) Change haz (pp)	(4) Change haz (%)
Select	-0.079***	-0.321*	-0.173***	-0.563**
	[0.023]	[0.160]	[0.030]	[0.186]
Kindred	-0.034	-0.307*	-0.102***	-0.522***
	[0.019]	[0.124]	[0.024]	[0.147]
Nonprofit	-0.022	-0.100	-0.066*	-0.169
	[0.020]	[0.186]	[0.028]	[0.247]
Co-located	-0.024	-0.323**	-0.017	-0.333*
	[0.016]	[0.109]	[0.021]	[0.129]
Beds	-0.000	-0.001	-0.000	-0.002
	[0.000]	[0.001]	[0.000]	[0.001]
% Medicaid	0.221	1.152	0.135	0.865
	[0.120]	[1.043]	[0.172]	[1.224]
Debt	-0.000	-0.000	-0.000*	-0.000
	[0.000]	[0.000]	[0.000]	[0.001]
% site-neutral	0.191*	1.090	0.324***	1.469*
	[0.077]	[0.629]	[0.087]	[0.662]
Constant	-0.019	0.129	0.044	0.391
	[0.031]	[0.249]	[0.038]	[0.293]
Observations	360	356	360	356
Adjusted R^2	0.095	0.060	0.222	0.091

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows the results of hospital-level regressions of the variation in discharge behavior at the threshold for site-neutral patients on hospital characteristics. Different columns correspond to different measures of pre/post-reform variation in discharge patterns.

Table 3.14: Regression of variation in discharge patterns for standard cases on hospital characteristics

	(1) Change prop (pp)	(2) Change prop (%)	(3) Change haz (pp)	(4) Change haz (%)
Select	0.055***	0.198	0.061***	0.120
	[0.013]	[0.113]	[0.016]	[0.116]
Kindred	-0.008	-0.182	-0.013	-0.201
	[0.012]	[0.097]	[0.016]	[0.111]
Nonprofit	0.019	0.189	0.023	0.165
	[0.012]	[0.151]	[0.016]	[0.151]
Co-located	0.001	-0.103	0.004	-0.103
	[0.009]	[0.086]	[0.012]	[0.088]
Beds	0.000*	0.000	0.000*	0.000
	[0.000]	[0.001]	[0.000]	[0.001]
% Medicaid	0.005	0.027	-0.011	-0.105
	[0.066]	[0.607]	[0.086]	[0.633]
Debt	0.000***	-0.002	0.000***	-0.003
	[0.000]	[0.003]	[0.000]	[0.003]
% site-neutral	0.065	0.551	0.094	0.625*
	[0.040]	[0.295]	[0.053]	[0.299]
Constant	-0.041*	-0.011	-0.049*	0.026
	[0.017]	[0.161]	[0.022]	[0.176]
Observations	360	357	360	357
Adjusted R^2	0.067	0.008	0.057	0.002

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows the results of hospital-level regressions of the variation in discharge behavior at the threshold for standard patients on hospital characteristics. Different columns correspond to different measures of pre/post-reform variation in discharge patterns.

period. Second, I use data on both periods to uncover changes in hospitals' response to financial incentives. Importantly, the exercise ignores the admission stage of the process.

3.5.1. Model

Consider a patient who has been admitted to an LTCH. The hospital faces a daily decision of whether to retain the patient or discharge her. During day t, the hospital receives a flow payoff from treating the patient that depends on the payment schedule faced by the hospital, the marginal cost of treating the patient, and other non-monetary considerations. If the hospital discharges the patient at the end of the day (so the length of stay is equal to t), then it frees up a bed (whose value is normalized to zero). If the hospital retains the patient, then it faces a similar decision next day. The value for the LTCH of treating a patient during day t is given by:

$$V(t,\varepsilon_{kt},\varepsilon_{dt}) = u_t + \max\{\varepsilon_{kt} + \delta EV(t+1), \varepsilon_{dt}\} \tag{3.1}$$

where u_t is the flow payoff of treating the patient during day t, $(\varepsilon_{kt},\varepsilon_{dt})$ are idiosyncratic error terms associated with the options of keeping and discharging the patient at the end of day t (observed by the hospital but not by the researcher), and $EV(t+1)$ is the ex-ante value function.

Under the assumption that the error terms have an iid type I extreme value distribution, the ex-ante value function is given by:

$$EV(t) = u_t + \log(1 + e^{\delta EV(t+1)}) \tag{3.2}$$

and the discharge hazard (the probability that the patient is discharged at the end of day t given that she has not been discharged earlier) is:

$$P(t) = \frac{1}{1 + e^{\delta EV(t+1)}} \tag{3.3}$$

The flow payoff depends on the marginal revenue from treating the patient during day t (p_t), the marginal cost (c), the day of the week, and the number of days that the patient has spent in the hospital. More specifically, the flow payoff is parameterized as:

$$u_t = \alpha p_t - \beta c + \gamma_{0,DRG} + \gamma_{1,DRG} t + \gamma_{2,DRG} t^2 + \gamma_{3,DRG} t^3 + \sum_{i=1}^{7} \phi_i \mathbb{I}(day_t = i) \tag{3.4}$$

The flow payoff includes a cubic polynomial in the absolute number of days since admission; I allow the associated parameters to be DRG and period (pre and post-reform) specific. The α and β parameters are interacted with the type of the hospital (for-profit other than Select and Kindred, Select, Kindred, and nonprofit), but are restricted to be the same pre and post-reform.

To estimate the model, I specify a period T and define a parameter Ω that represents the value of treating the patient beyond T. I allow the continuation value to be DRG and period specific. The model can then be solved backwards recursively using Equation 3.2. Then, I can determine the discharge hazard rate for each day t using Equation 3.3. With this, I can compute the probability of the observed length of stay for each patient. I then estimate the model by maximum likelihood.

3.5.2. Estimation and results

The estimation sample consists of site-neutral patients in the working sample who were ultimately discharged downstream. I consider the nine most-common DRGs in the data[18]. I consider all pre and post-reform cases, so I do not restrict attention to FY 2015 and 2017 as in the analysis above. Table 3.15 contains summary statistics of the estimation sample.

The marginal revenue p_t is computed directly from the Medicare payment schedule as explained in Appendix C. The estimate of marginal cost c is computed as the product of the charges indicated in the record of the patient and the cost-to-charge ratio of the hospital.

The model has 105 parameters. The parameters associated with the polynomial in the absolute

[18]These DRGs jointly account for approximately 32% of site-neutral discharges in the working sample.

Table 3.15: Estimation sample, summary statistics

	Pre-reform	Post-reform
DRG 189	0.23	0.27
DRG 207	0.04	0.05
DRG 539	0.12	0.11
DRG 559	0.07	0.06
DRG 560	0.07	0.07
DRG 570	0.07	0.08
DRG 592	0.14	0.12
DRG 638	0.08	0.07
DRG 871	0.18	0.15
Other for-profit	0.43	0.60
Select	0.18	0.07
Kindred	0.29	0.24
Nonprofit	0.10	0.10
Co-located	0.28	0.24
Length of stay	27.41	23.83
Observations	38,075	13,981

Notes: This table shows summary statistics of the sample used for estimating the structural model.

number of days since admission and the continuation values are DRG and period specific. The parameters on marginal revenue and marginal cost depend on the type of the hospital, but are the same across periods. The day of week fixed effects are the same across periods. The main interest lies in the α coefficients, which capture the responsiveness of LTCHs to financial incentives.

I estimate different versions of the model. Specification 1 considers only pre-reform cases. Specification 2 considers only post-reform cases. Specification 3 considers both pre and post-reform cases and allows the γ and Ω coefficients to be period specific. Finally, specification 4 considers both pre and post-reform cases but restricts the γ coefficients to be equal across periods.

In Table 3.16, I report the estimates of the α parameters for the different specifications. The results are consistent with those reported by Eliason et al. (2016). First, α is positive and statistically significant for all types of hospitals. Second, for-profit hospitals other than Select and Kindred are the most responsive to financial incentives.

Table 3.16: Estimated α parameters, by specification

	(1)	(2)	(3)	(4)
For-profit	1.58	2.89	1.77	1.66
	(0.021)	(0.050)	(0.018)	(0.137)
Select	1.36	2.80	1.51	1.35
	(0.024)	(0.094)	(0.023)	(0.122)
Kindred	1.17	2.84	1.37	1.23
	(0.019)	(0.057)	(0.018)	(0.105)
Nonprofit	1.16	2.64	1.35	1.21
	(0.028)	(0.075)	(0.025)	(0.169)

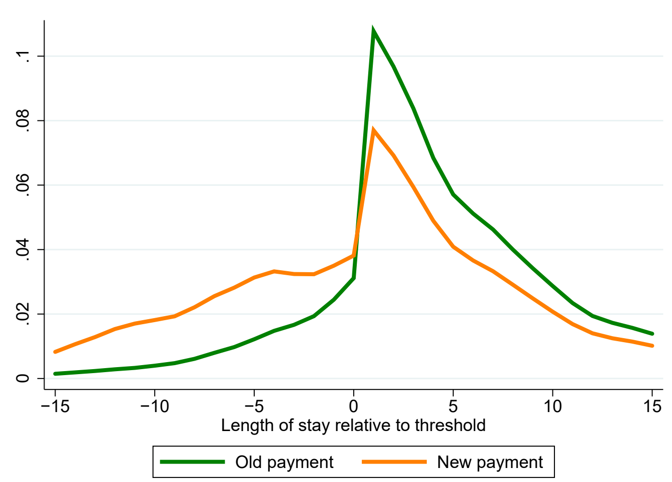
Notes: This table shows the estimated α parameters for the different specifications. Standard errors in parenthesis.

Comparing specifications 1 and 2, we observe that the α coefficients are much higher for the latter. The rationale is the following. The marginal payment before the threshold and the jump in payment at the threshold are lower in the post-reform period. However, the spike in discharges at the threshold does not decrease much or even increases under the new payment schedule. In order for the model to match this response, it must be that hospitals are more responsive than before to financial incentives. When data on both the pre and post-reform periods is used, a single α parameter governs the response to financial incentives in both periods. Therefore, the maximum likelihood estimate will be between the first two cases.

With the estimated parameters, I use simulation to recover the discharge patterns predicted by the model. As in Section 3.4.3, I define the normalized length of stay and compute the proportion of discharges and the discharge hazard rate for each day.

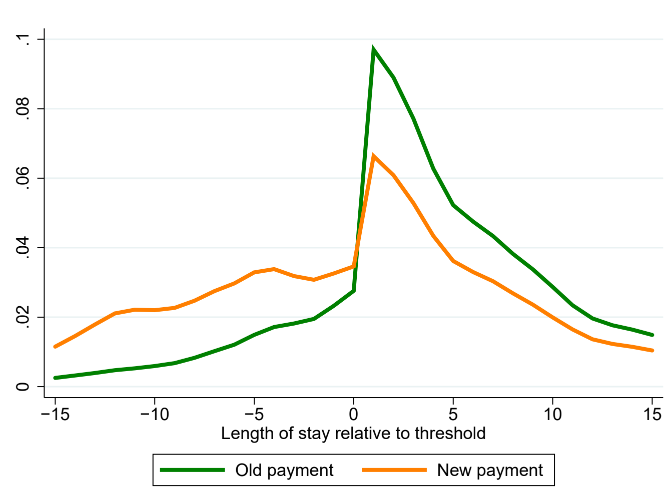
First, I consider specification 1 and recover the discharge patterns predicted by the model under the observed (old) payment schedule and under the new payment schedule. The objective is to analyze how the payment reform would affect the discharge behavior of hospitals. As Eliason et al. (2016) point out, this counterfactual only provides insight on how the new payment system would affect those patients who were treated at LTCHs during the pre-reform period. Figures 3.18 to 3.21 show the results. The predicted patterns are consistent with the argument provided in Section 3.4.3.

Figure 3.18: Predicted discharge patterns, pre-reform cases, other for-profit hospitals



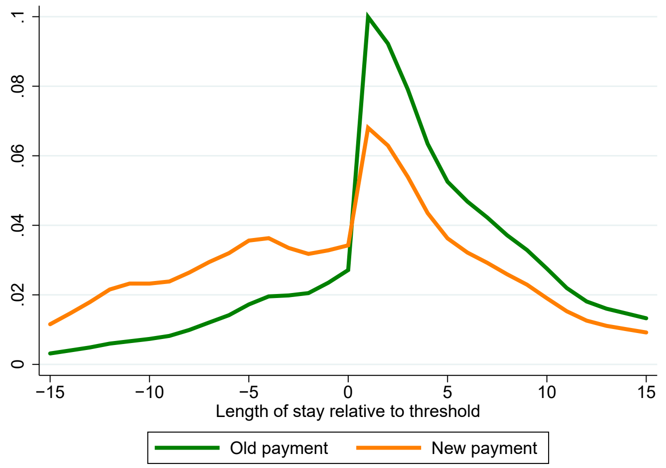
Notes: This figure shows the discharge patterns for pre-reform cases predicted by specification 1 of the model under the observed (old) payment schedule and under the new payment schedule.

Figure 3.19: Predicted discharge patterns, pre-reform cases, Select hospitals



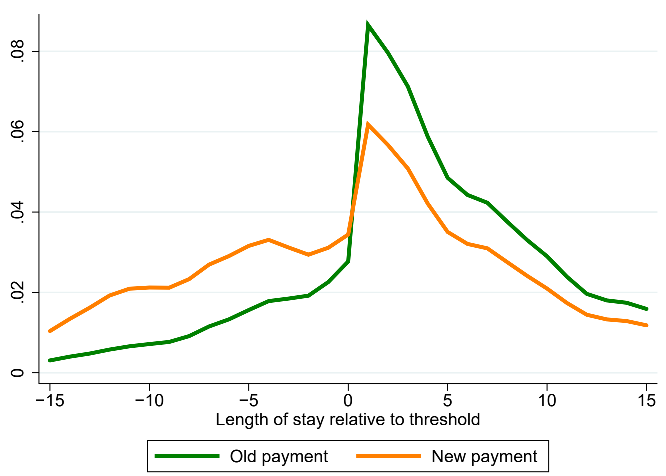
Notes: This figure shows the discharge patterns for pre-reform cases predicted by specification 1 of the model under the observed (old) payment schedule and under the new payment schedule.

Figure 3.20: Predicted discharge patterns, pre-reform cases, Kindred hospitals



Notes: This figure shows the discharge patterns for pre-reform cases predicted by specification 1 of the model under the observed (old) payment schedule and under the new payment schedule.

Figure 3.21: Predicted discharge patterns, pre-reform cases, nonprofit hospitals



Notes: This figure shows the discharge patterns for pre-reform cases predicted by specification 1 of the model under the observed (old) payment schedule and under the new payment schedule.

In particular, the payment reform leads to an increase in the proportion of discharges that take place before the SSO threshold and a decrease in the spike in discharges at the threshold. This is driven by an increase in the discharge hazard before the threshold.

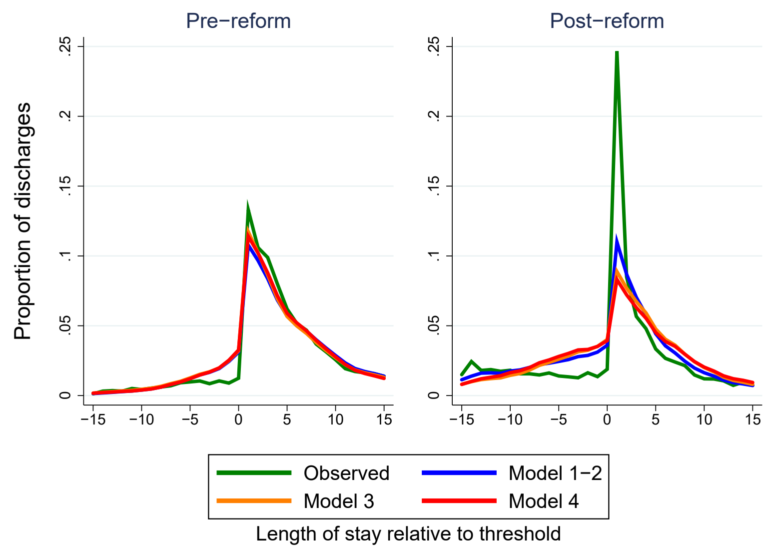
Second, I analyze the fit of the model. Figures 3.22 to 3.24 show the discharge patterns predicted by the different specifications of the model as well as the observed discharge patterns, pooling across all DRGs but separately for pre and post-reform cases.

The fit of the model for pre-reform cases is quite good. The model generates the spike in discharges at the threshold documented before, except for the case of Select hospitals. The performance of the model to match post-reform behavior is not as good. In all cases, specifications 3 and 4 generate similar discharge patterns. Therefore, allowing the polynomial in the absolute number of days since admission to differ across periods does not change the fit of the model substantially. In the case of nonprofit hospitals (and to a smaller extent Kindred and Select facilities), the model estimated using only post-reform cases does a relatively good job matching the spike in discharges at the threshold. In the case of other for-profit hospitals, none of the specifications is able to match the spike in discharges at the threshold. Even when only post-reform cases are considered, the fit is relatively poor. When both pre and post-reform cases are used for estimation and the coefficients on marginal revenue and cost are constrained to be equal across periods, the fit worsens. When all the coefficients are restricted to be equal, the fit also worsens but not significantly.

In summary, the dynamic model considered here supports the findings of previous sections. Given the way payments for site-neutral cases changed with the reform, we expect an increase in the number of SSO discharges and a smaller spike in discharges at the threshold. In the case of other for-profit hospitals, the model does a very poor job replicating the increase in the discharge hazard rate at the threshold post-reform. For the rest of the facilities, the model provides a reasonable fit to post-reform discharge behavior, although it indicates a substantial increase in the responsiveness of hospitals to financial incentives.

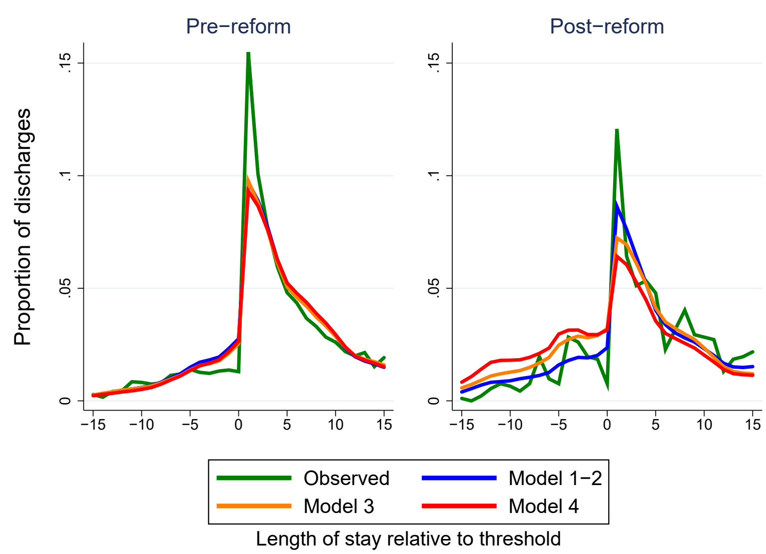
As mentioned in Section 3.4.3, while there is limited evidence that the composition of LTCH site-

Figure 3.22: Discharge patterns, other for-profit hospitals



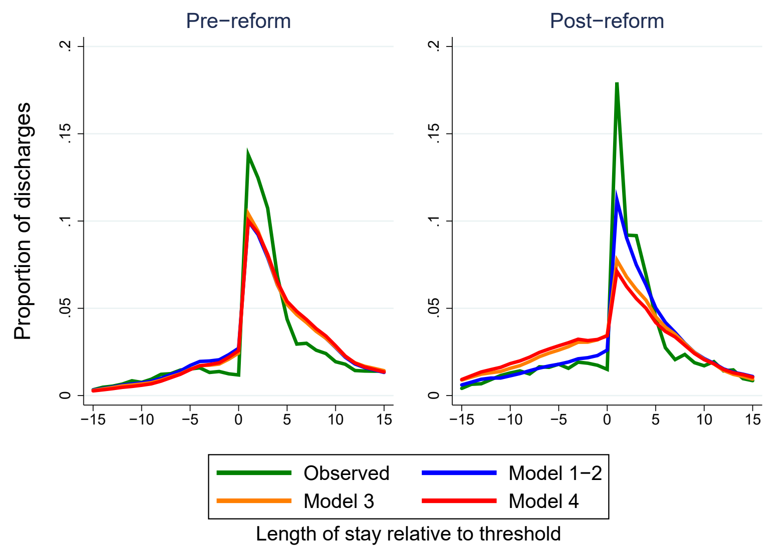
Notes: This figure shows the discharge patterns predicted by the model and the discharge patterns observed in the data.

Figure 3.23: Discharge patterns, Select hospitals



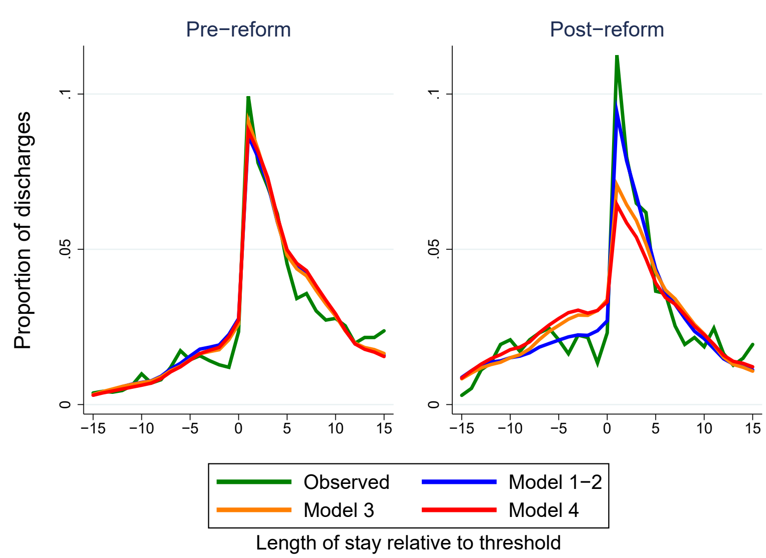
Notes: This figure shows the discharge patterns predicted by the model and the discharge patterns observed in the data.

Figure 3.24: Discharge patterns, Kindred hospitals



Notes: This figure shows the discharge patterns predicted by the model and the discharge patterns observed in the data.

Figure 3.25: Discharge patterns, nonprofit hospitals



Notes: This figure shows the discharge patterns predicted by the model and the discharge patterns observed in the data.

neutral patients changed with the reform in terms of observables, there is the possibility of changes in unobservable health. The model discussed in this section has a limited ability to detect changes in the composition of patients. The model of Einav et al. (2017) is better suited to accommodate potential differences in the LTCH patient mix. I lack the data required to estimate this model for my application; specifically, I do not have information about post-discharge Medicare utilization for LTCH patients. I estimated a simpler version of their model and obtained similar conclusions as with the model of Eliason et al. (2016) discussed in this section. However, as mentioned before, additional data is needed to fully implement their approach.

3.6. Conclusion

In this chapter, I study the impact of a major regulatory reform to the LTCH industry. LTCHs suffer significant payment reductions compared to the pre-reform period when caring for site-neutral (low severity) cases. In terms of admission decisions and discharge practices after the SSO threshold, there are varied responses to the new payment system.

In the case of Select facilities, there was a widespread and dramatic decrease in both the incidence and number of site-neutral cases. In the case of nonprofit facilities, both the incidence and number of site-neutral cases decreased; however, the magnitude of these changes is smaller than for Select hospitals. In the case of Kindred hospitals and for-profit hospitals other than Select and Kindred, the incidence and number of site-neutral cases have decreased on average; however, a non negligible number of hospitals are attracting more site-neutral cases.

In relation to discharge patterns, other for-profit hospitals are much more likely to discharge patients immediately after the SSO threshold compared to the pre-reform period, while the behavior of other types of hospitals has not changed much with respect to the pre-reform period. The change in discharge patterns of other for-profit hospitals at the threshold does not seem consistent with the change in the payment schedule faced by hospitals who evaluate alternatives based on marginal quantities. I use a dynamic structural model of hospital behavior used in previous work on this industry to show that the observed change in discharge behavior differs from the expected change given the way the payment schedule was modified. I explore alternative explanations for the unexpected change in discharge behavior, but I do not find conclusive empirical support for any of them. Therefore, the reasons for some of the changes in discharge patterns remain largely unexplained.

With a framework that explains changes in both admission and treatment patterns, there are interesting policy questions that can be answered. First, what is the predicted impact of the reform on the configuration of the LTCH industry? In particular, will lower reimbursement lead to more closures over the next few years?

Second, what is the impact of this policy on health outcomes? My results suggest that the strategy of decreasing LTCH reimbursement rates for low severity patients has hampered patient access to care. For site-neutral patients, the reform seems to have induced substitution towards less sophisticated settings such as SNFs. If site-neutral patients are treated in lower tier post-acute providers post-reform, then health outcomes could be affected. Therefore, the savings brought about by the reform need to be weighed against potential adverse effects on the health of patients. For site-neutral patients who continue being treated in LTCHs, the payment reform can have effects on health outcomes through changes in the timing of discharge.

Third, has competition for standard patients intensified under the new payment system? If this is the case, has more fierce competition provided hospitals with incentives to improve the quality of their services? What is the contribution of this channel to changes in the clinical outcomes of these patients?

Finally, the analysis would also shed light on the merits of alternative regulatory designs for the LTCH industry and the post-acute care system in general. It would be interesting to study the impact of alternative exclusion criteria to define patient types, particularly a policy that modifies the exclusion criteria so that only patients with 8 or more previous ICU days qualify for the standard rate. This was the suggestion made by the Medicare Payment Advisory Commission (MedPAC, 2016): it considered that a threshold lower than eight days is too low to separate patients who require long-term post-acute care from patients who can be treated appropriately in other (lower cost) settings.

I lack the necessary data to address these questions at the present time. Specifically, I need data on Medicare inpatient and SNF utilization. The discharge data used in this study contains only information on LTCH stays. Due to lags in data availability, obtaining the required dataset (RIF MedPAR file) with post-reform data was not feasible by the date this study was completed. However, I view them as interesting avenues for future research.

Bibliography

Andrews, R. L. and T. Srinivasan (1995). "Studying consideration effects in empirical choice models using scanner panel data". *Journal of Marketing Research*, pp. 30–41.

Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997). "Labor supply of New York City cabdrivers: One day at a time". *Quarterly Journal of Economics* 112 (2), pp. 407–441.

Card, D., C. Dobkin, and N. Maestas (2009). "Does Medicare save lives?" *Quarterly Journal of Economics* 124 (2), pp. 597–636.

Carlson, J. A., L. S. Dafny, B. A. Freeborn, P. M. Ippolito, and B. W. Wendling (2013). "Economics at the FTC: Physician acquisitions, standard essential patents, and accuracy of credit reporting". *Review of Industrial Organization* 43 (4), pp. 303–326.

Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson (2016). "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector". *American Economic Review* 106 (8), pp. 2110–44.

CMS (2016). "Long Term Care Hospital Prospective Payment System Payment System Series".

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

Desai, S. (2016). "The Role of Interoperability in Hospital Digitization: Network Effects and Information Blocking".

Doyle, J., J. Graves, J. Gruber, and S. Kleiner (2015). “Measuring returns to hospital care: Evidence from ambulance referral patterns”. *Journal of Political Economy* 123 (1), pp. 170–214.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010). “State dependence and alternative explanations for consumer inertia”. *RAND Journal of Economics* 41 (3), pp. 417–445.

Einav, L., A. Finkelstein, and N. Mahoney (2017). *Provider Incentives and Healthcare Costs: Evidence from Long-Term Care Hospitals*. Working Paper 23100. National Bureau of Economic Research.

Eliason, P. J., P. L. E. Grieco, R. C. McDevitt, and J. W. Roberts (2016). *Strategic Patient Discharge: The Case of Long-Term Care Hospitals*. Working Paper 22598. National Bureau of Economic Research.

Ericson, K. M. M. (2014). “Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange”. *American Economic Journal: Economic Policy* 6 (1), pp. 38–64.

Farber, H. S. (2005). “Is tomorrow another day? The labor supply of New York City cabdrivers”. *Journal of Political Economy* 113 (1), pp. 46–82.

Farrell, J. and P. Klemperer (2007). “Coordination and lock-in: Competition with switching costs and network effects”. *Handbook of Industrial Organization* 3, pp. 1967–2072.

Fehr, E. and L. Goette (2007). “Do workers work more if wages are high? Evidence from a randomized field experiment”. *American Economic Review* 97 (1), pp. 298–317.

Freedman, S. (2016). “Capacity and utilization in health care: the effect of empty beds on neonatal intensive care admission”. *American Economic Journal: Economic Policy* 8 (2), pp. 154–85.

Gaynor, M., C. Propper, and S. Seiler (2016). “Free to Choose? Reform, Choice, and Consideration Sets in the English National Health Service”. *American Economic Review* 106 (11), pp. 3521–57.

Goldfarb, A. (2006a). “State dependence at internet portals”. *Journal of Economics & Management Strategy* 15 (2), pp. 317–352.

— (2006b). “The medium-term effects of unavailability”. *Quantitative Marketing and Economics* 4 (2), pp. 143–171.

Gruber, J., J. Kim, and D. Mayzlin (1999). “Physician fees and procedure intensity: the case of cesarean delivery”. *Journal of Health Economics* 18 (4), pp. 473–490.

Gruber, J. and M. Owings (1996). “Physician financial incentives and cesarean section delivery”. *RAND Journal of Economics* 27 (1), p. 99.

Handel, B. R. (2013). “Adverse selection and inertia in health insurance markets: When nudging hurts”. *American Economic Review* 103 (7), pp. 2643–2682.

Heckman, J. J. (1981). “Heterogeneity and state dependence”. *Studies in labor markets*. University of Chicago Press, pp. 91–140.

Heckman, J. J., J. Smith, and N. Clements (1997). “Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts”. *Review of Economic Studies* 64 (4), pp. 487–535.

HITPC (2015). “Challenges and Barriers to Interoperability, Report to Congress”.

Ho, K. and A. Pakes (2014). “Hospital choices, hospital prices, and financial incentives to physicians”. *American Economic Review* 104 (12), pp. 3841–3884.

Ho, K., J. Hogan, and F. Scott Morton (2017). “The impact of consumer inattention on insurer pricing in the Medicare Part D program”. *RAND Journal of Economics* 48 (4), pp. 877–905.

Hodgkin, D. and T. G. McGuire (1994). “Payment levels and hospital response to prospective payment”. *Journal of Health Economics* 13 (1), pp. 1–29.

Illanes, G. (2016). “Switching Costs in Pension Plan Choice”.

Israel, M. (2005). “Tenure dependence in consumer-firm relationships: An empirical analysis of consumer departures from automobile insurance firms”. *RAND Journal of Economics*, pp. 165–192.

Jung, K., R. Feldman, and D. Scanlon (2011). “Where would you go for your next hospitalization?” *Journal of Health Economics* 30 (4), pp. 832–841.

Kim, S.-H., C. W. Chan, M. Olivares, and G. Escobar (2014). “ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes”. *Management Science* 61 (1), pp. 19–38.

Kim, Y. S., E. C. Kleerup, P. A. Ganz, N. A. Ponce, K. A. Lorenz, and J. Needleman (2015). “Medicare payment policy creates incentives for long-term care hospitals to time discharges for maximum reimbursement”. *Health Affairs* 34 (6), pp. 907–915.

Luco, F. (2016). “Switching Costs and Competition in Retirement Investment”.

McGuire, T. G. and M. V. Pauly (1991). “Physician response to fee changes with multiple payers”. *Journal of Health Economics* 10 (4), pp. 385–410.

MedPAC (2012). “Report to the Congress, Medicare Payment Policy, March 2012”.

— (2015). “Report to the Congress, Medicare Payment Policy, March 2015”.

— (2016). “Report to the Congress, Medicare Payment Policy, March 2016”.

Miller, A. R. and C. Tucker (2009). “Privacy protection and technology diffusion: The case of electronic medical records”. *Management Science* 55 (7), pp. 1077–1093.

— (2014). “Health information exchange, system size and information silos”. *Journal of Health Economics* 33, pp. 28–42.

Nosal, K. (2012). “Estimating switching costs for medicare advantage plans”.

Phibbs, C. S., D. H. Mark, H. S. Luft, D. J. Peltzman-Rennie, D. W. Garnick, E. Lichtenberg, and S. J. McPhee (1993). “Choice of hospital for delivery: a comparison of high-risk and low-risk women.” *Health Services Research* 28 (2), pp. 201–222.

Polyakova, M. (2016). “Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D”. *American Economic Journal: Applied Economics* 8 (3), pp. 165–95.

Pruitt, N. J. (2013). “Site Neutral Payment Methodology and Long Term Services and Supports Implications”.

Raval, D. and T. Rosenbaum (2017). “Why Do Previous Choices Matter for Hospital Demand? Decomposing Switching Costs from Unobserved Preferences”.

— (2018). “Why is Distance Important for Hospital Choice? Separating Home Bias from Transport Costs”.

Raval, D., T. Rosenbaum, and S. A. Tenn (2017a). “A Semiparametric Discrete Choice Model: An Application to Hospital Mergers”. *Economic Inquiry* 55 (4), pp. 1919–1944.

Raval, D., T. Rosenbaum, and N. Wilson (2017b). “Using Disaster Induced Closures to Evaluate Discrete Choice Models of Hospital Demand”.

Romano, J. P. and A. M. Shaikh (2008). “Inference for identifiable parameters in partially identified econometric models”. *Journal of Statistical Planning and Inference* 138 (9), pp. 2786–2807.

Samuelson, W. and R. Zeckhauser (1988). “Status quo bias in decision making”. *Journal of Risk and Uncertainty* 1 (1), pp. 7–59.

Shepard, M. (2016). *Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange*. Working Paper 22600. National Bureau of Economic Research.

Stigler, G. J. and G. S. Becker (1977). “De gustibus non est disputandum”. *American Economic Review* 67 (2), pp. 76–90.

Sudhir, K and N. Yang (2014). “Exploiting the choice-consumption mismatch: A new approach to disentangle state dependence and heterogeneity”.

Tian, W (2016). “An all-payer view of hospital discharge to postacute care, 2013”. *HCUP statistical brief* 205, pp. 1–17.

Torgovitsky, A. (2016). “Nonparametric Inference on State Dependence with Applications to Employment Dynamics”.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Appendix A

Appendix 1 for Chapter 1

A.1. Hurricane Sandy

In this section, I provide additional details about Hurricane Sandy and the resulting hospital closures that I exploit to analyze patient loyalty.

According to the National Weather Service[1], Sandy formed as a tropical storm on October 22th, 2012, in the Caribbean and intensified into a hurricane as it moved northward across Jamaica, Cuba and the Bahamas. It then moved northeast of the United States until turning west toward the mid Atlantic coast on the 28th. It transitioned into a post-tropical cyclone just prior to moving onshore near Atlantic City, New Jersey, on the 29th, and dissipated over western Pennsylvania by the 31th. The storm affected many states, but was particularly severe in New York and New Jersey. It is regarded as one of the most destructive storms in U.S. history: according to several sources, the storm caused more than 100 deaths in the U.S., with at least 40 deaths in New York City (NYC), while economic damages are calculated to have exceeded $50 billion.

The storm hit the New York Metropolitan area on October 29th. Preparations in NYC began on October 26th and included, among other measures, mandatory evacuations, flights cancellations, suspension of public transportation, and cancellation of classes. The storm affected the entire city, although zones close to the water suffered the most. Storm surge caused record tide levels in many areas, producing flooding across the city. Strong winds caused loss of power and large parts of the city and surrounding areas remained without electricity for some days. Although there were certainly long lasting effects from the storm, most areas returned to normality within days.

Five hospitals in NYC were affected[2]: New York Downtown Hospital, Manhattan VA Medical Center, NYU Langone Medical Center, Bellevue Hospital, and Coney Island Hospital. Several hospitals agreed to take some of the affected hospitals' patients. New York Downtown Hospital

[1] https://www.weather.gov/okX/HurricaneSandy

[2] See https://www.cbsnews.com/pictures/nyc-hospitals-evacuated-for-superstorm/ , https://www.reuters.com/article/storm-sandy-bellevue-idUSL1N0B78OY20130207, and https://www.nytimes.com/2012/10/30/nyregion/patients-evacuated-from-nyu-langone-after-power-failure.html for more details about the impact of the storm

evacuated its patients before the storm; it suffered minor damage and was able to fully reopen approximately one week after the storm. The other hospitals, on the other hand, suffered extensive damage and the necessary repair work kept the facilities closed for several weeks.

The Manhattan VA Medical Center was evacuated on October 27th, before the storm hit. It suffered extensive damage to its electrical and mechanical systems as well as clinical equipment. It fully reopened in May 2013.

NYU Langone Medical Center began evacuating around 300 patients on the evening of October 29th after the backup generators at the hospital failed due to flooding. Some limited outpatient services reopened during the first days of November, but the hospital only partially resumed inpatient services on December 27th. The maternity unit and pediatrics reopened on January 14th, 2013. An urgent care center was opened on January 17th, but the emergency room did not fully open until April 24th, 2014.

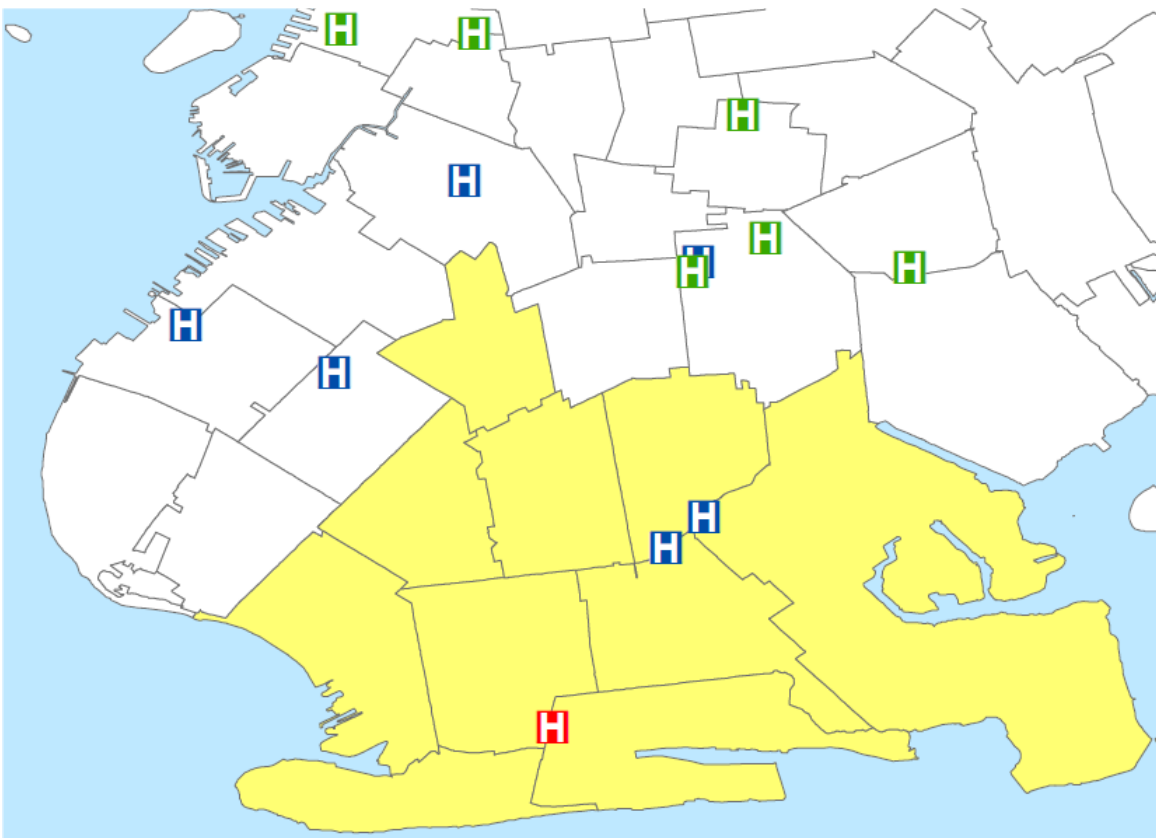
Bellevue Hospital began evacuating about 500 patients on October 31th, when the full extent of the damage became clear: millions of gallons of water poured into its basement, which contained equipment critical to the hospital's operations. The hospital had been operating on backup generators since losing power during the storm. It was the hospital's first evacuation in 276 years. It resumed limited outpatient services on November 19th, but it did not fully reopen inpatient services until February 7th, 2013.

Evacuation of about 180 patients at Coney Island Hospital began on the afternoon of October 30th. The hospital had been using generators since the hurricane caused power outages across Southern Brooklyn. An urgent care center was opened by December 3th but patients were not admitted inpatient. It reopened most of its inpatient beds by February 20th, 2013. The labor and delivery unit did not reopen until June 13th, 2013.

The unavailability of these hospitals meant increased patient volume for other facilities. Anecdotal evidence suggests that the latter had difficulties accommodating extra patients and that capacity constraints were severe, with patients facing abnormally long waiting times for getting care in many cases. Some doctors, residents, and nurses at the affected facilities were able to work at other facilities while their hospital remained closed.

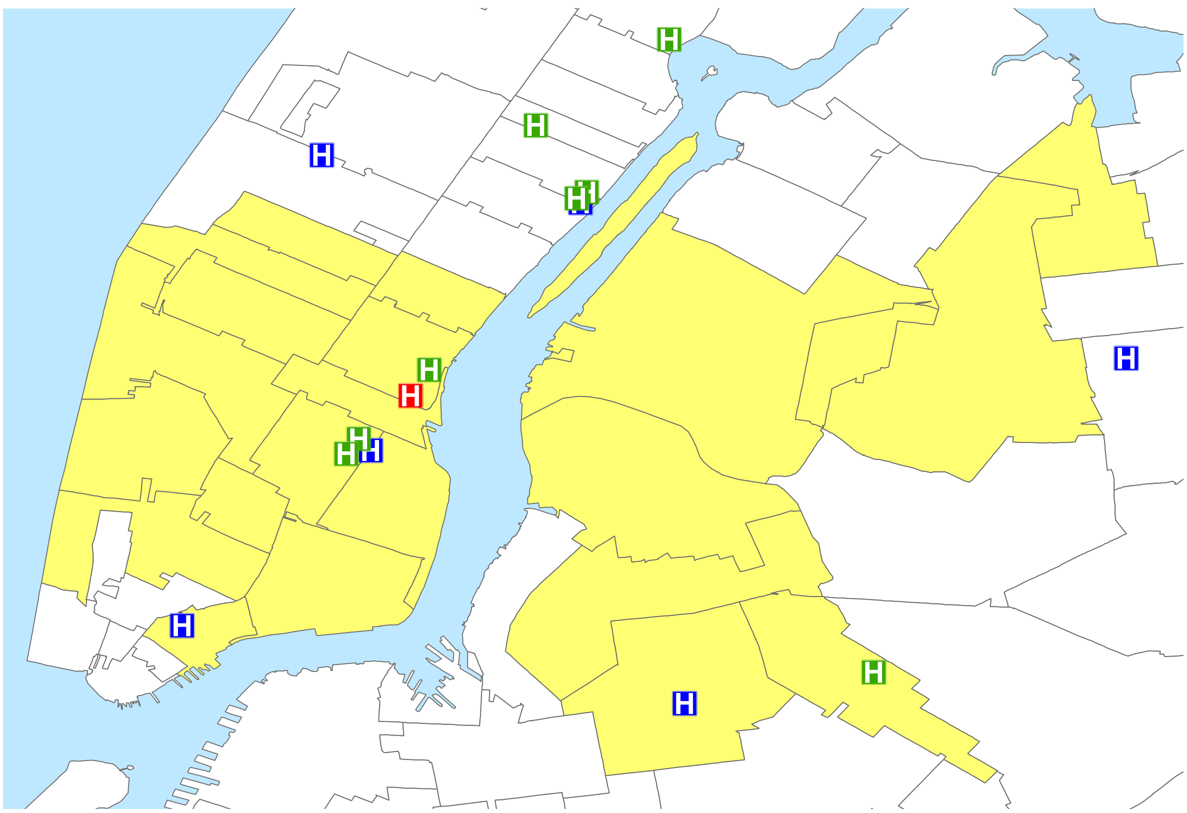
A.2. Additional figures and tables

Figure A.1: Service area of Coney Island Hospital



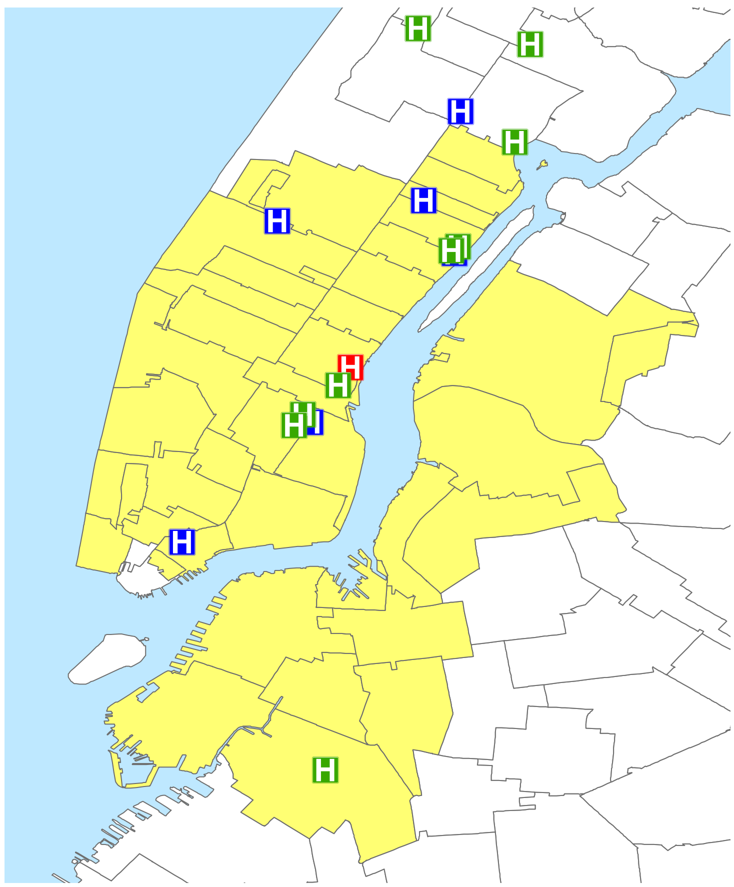
Notes: This figure shows the service area of Coney Island Hospital (zip codes in yellow) constructed as explained in Section 1.5.2. The red hospital is Coney Island Hospital, the blue hospitals are its main competitors, and the green hospitals are other nearby facilities.

Figure A.2: Service area of Bellevue



Notes: This figure shows the service area of Bellevue Hospital (zip codes in yellow) constructed as explained in Section 1.5.2. The red hospital is Bellevue, the blue hospitals are its main competitors, and the green hospitals are other nearby facilities.

Figure A.3: Service area of NYU Langone



Notes: This figure shows the service area of NYU Langone (zip codes in yellow) constructed as explained in Section 1.5.2. The red hospital is NYU Langone, the blue hospitals are its main competitors, and the green hospitals are other nearby facilities.

Table A.1: Summary statistics, Coney Island Hospital

	Sample 1		Sample 2		Sample 3	
	Control	Treatment	Control	Treatment	Control	Treatment
Type of visit:						
Inpatient	0.21	0.26	0.15	0.25	0.22	0.30
Emergency Department	0.58	0.55	0.73	0.66	0.69	0.62
Ambulatory Surgery	0.21	0.19	0.12	0.10	0.09	0.09
Demographics:						
Age	46.1	49.5	39.6	44.8	43.1	47.6
Female	0.56	0.57	0.53	0.56	0.56	0.58
Medicare	0.24	0.33	0.15	0.27	0.23	0.35
Medicaid	0.31	0.33	0.40	0.44	0.37	0.40
Private insurance	0.28	0.23	0.15	0.11	0.22	0.17
Other insurance	0.17	0.11	0.30	0.17	0.17	0.08
Prior propensity:						
0	0.75	0.77	0.10	0.13	0.63	0.71
0-0.25	0.02	0.04	0.01	0.04	0.02	0.04
0.25-0.50	0.03	0.04	0.04	0.09	0.03	0.05
0.50-0.75	0.04	0.05	0.09	0.15	0.05	0.06
0.75-1	0.18	0.10	0.77	0.59	0.28	0.14
Patients	193,222	19,577	38,615	2,541	102,867	8,882

Notes: This table presents summary statistics of the patients used for the case study of Coney Island Hospital. Treatment status refers to whether the patient had a hospital visit in period 1. The different samples are described in detail in Section 1.5.2. Sample #1 contains patients who had at least one hospital visit in both periods 0 and 2. Sample #2 contains those patients in sample #1 who visited Coney Island Hospital in the last episode of period 0. Sample #3 contains those patients in sample #1 who chose either Coney Island Hospital or one of its main competitors in the last episode of period 0, in period 1 (if in the treatment group), and in the first episode of period 2. Characteristics correspond to the first episode of period 2. Prior propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to Coney Island Hospital.

Table A.2: Summary statistics, Bellevue

	Sample 1		Sample 2		Sample 3	
	Control	Treatment	Control	Treatment	Control	Treatment
Type of visit:						
Inpatient	0.16	0.21	0.15	0.20	0.15	0.19
Emergency Department	0.65	0.60	0.75	0.72	0.76	0.72
Ambulatory Surgery	0.19	0.19	0.10	0.09	0.09	0.08
Demographics:						
Age	43.7	48.9	42.4	47.5	41.0	45.4
Female	0.54	0.57	0.46	0.49	0.53	0.57
Medicare	0.21	0.32	0.16	0.27	0.19	0.30
Medicaid	0.34	0.36	0.42	0.47	0.41	0.45
Private insurance	0.28	0.21	0.11	0.09	0.20	0.14
Other insurance	0.16	0.11	0.31	0.18	0.20	0.12
Prior propensity:						
0	0.86	0.87	0.14	0.16	0.74	0.80
0-0.25	0.02	0.03	0.03	0.07	0.03	0.05
0.25-0.50	0.02	0.03	0.05	0.08	0.03	0.04
0.50-0.75	0.02	0.03	0.09	0.15	0.03	0.04
0.75-1	0.08	0.05	0.69	0.54	0.17	0.08
Patients	230,040	20,267	24,047	1,510	88,938	6,940

Notes: This table presents summary statistics of the patients used for the case study of Bellevue Hospital. Treatment status refers to whether the patient had a hospital visit in period 1. The different samples are described in detail in Section 1.5.2. Sample #1 contains patients who had at least one hospital visit in both periods 0 and 2. Sample #2 contains those patients in sample #1 who visited Bellevue in the last episode of period 0. Sample #3 contains those patients in sample #1 who chose either Bellevue or one of its main competitors in the last episode of period 0, in period 1 (if in the treatment group), and in the first episode of period 2. Characteristics correspond to the first episode of period 2. Prior propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to Bellevue.

Table A.3: Summary statistics, NYU Langone

	Sample 1		Sample 2		Sample 3	
	Control	Treatment	Control	Treatment	Control	Treatment
Type of visit:						
Inpatient	0.20	0.22	0.26	0.31	0.28	0.27
Emergency Department	0.53	0.58	0.43	0.44	0.55	0.60
Ambulatory Surgery	0.28	0.21	0.32	0.26	0.18	0.13
Demographics:						
Age	48.2	48.7	50.9	62.5	48.0	50.0
Female	0.55	0.57	0.61	0.59	0.60	0.58
Medicare	0.26	0.31	0.27	0.49	0.27	0.36
Medicaid	0.22	0.32	0.06	0.09	0.17	0.29
Private insurance	0.41	0.27	0.61	0.38	0.50	0.30
Other insurance	0.12	0.11	0.06	0.04	0.06	0.06
Prior propensity:						
0	0.91	0.92	0.19	0.24	0.84	0.89
0-0.25	0.01	0.02	0.02	0.08	0.01	0.03
0.25-0.50	0.01	0.02	0.05	0.12	0.02	0.02
0.50-0.75	0.02	0.02	0.09	0.15	0.02	0.02
0.75-1	0.06	0.03	0.65	0.42	0.11	0.04
Patients	299,009	28,933	20,182	1,118	91,773	8,022

Notes: This table presents summary statistics of the patients used for the case study of NYU Langone. Treatment status refers to whether the patient had a hospital visit in period 1. The different samples are described in detail in Section 1.5.2. Sample #1 contains patients who had at least one hospital visit in both periods 0 and 2. Sample #2 contains those patients in sample #1 who visited NYU Langone in the last episode of period 0. Sample #3 contains those patients in sample #1 who chose either NYU Langone or one of its main competitors in the last episode of period 0, in period 1 (if in the treatment group), and in the first episode of period 2. Characteristics correspond to the first episode of period 2. Prior propensity is defined as the proportion of the patient's hospital visits prior to the storm that were to NYU Langone.

Appendix B

Appendix 2 for Chapter 1

This Appendix provides additional details about the nonparametric framework discussed in Section 1.5.4. First, I explain the optimization problem that is used to recover the identified set. Second, I discuss some of the identifying restrictions in more detail.

B.1. Optimization problem

As in the main text, let $\mathcal{H}$ denote the set of possible types in the population. For each type $h \in \mathcal{H}$, let $sd(h) \in \{0,1\}$ indicate whether a patient with type h exhibits state dependence in period 2. The objective is to recover the identified set for the proportion of patients who exhibit state dependence in period 2. Under certain conditions (which are met in my analysis), the identified set for this parameter is a closed interval. Therefore, I only need to recover the endpoints of the interval. The problem is then to find the admissible distributions under which the extent of state dependence is lowest and highest.

More precisely, I need to find the probabilities $\{p(h)\}_{h\in\mathcal{H}}$ that satisfy the restrictions on the set of admissible distributions and that minimize (maximize) the quantity of interest $\sum_{h\in\mathcal{H}} p(h)sd(h)$. Note that this is a linear function of the probabilities. The restrictions on the set of admissible distributions that I consider are also linear in the probabilities. Then, for example, the lower bound on the extent of state dependence is a solution to the following linear problem:

$$
\begin{aligned}
\underset{\{p(h)\}}{\text{minimize}} \quad & \sum_{h \in \mathcal{H}} p(h) sd(h) \\
\text{subject to} \quad & p(h) \geq 0, \forall h \in \mathcal{H} \\
& \sum_{h \in \mathcal{H}} p(h) = 1 \\
& \sum_{h: y(h) = y} p(h) = P(y), \forall y \in Y \\
& \sum_{h \in \mathcal{H}} w(h) p(h) \leq 0
\end{aligned}
\tag{B.1}
$$

The first two restrictions ensure that I have a proper p.m.f. The third restriction demands that I only consider probability distributions over the set of types that are consistent with the data. For each $h \in \mathcal{H}$, $y(h)$ is the vector of observables associated with a patient of type h. The restriction is that the probability assigned to all types associated with a given vector of observables must be equal to the proportion of patients in the data with these observables. The fourth restriction captures other linear constraints on probabilities such as Assumption 4.

B.2. Ex-ante restrictions

As discussed in Section 1.5.4, even after limiting the size of the choice set and the number of periods considered, the dimension of the set of types is prohibitive from a computational point of view. To deal with this issue, I impose ex-ante restrictions on the set of admissible types. These restrictions are not only useful to reduce the dimensionality of the problem, but they also have economic content. Here, I discuss Assumption 3 in the context of a standard model of consumer choice.

I assume that choices are determined by utility maximization. The utility that patient i gets from choosing alternative j in period t is given by:

$$u_{ijt} = g(X_{it}, Z_{jt}) + \sum_{k=1}^{J} \gamma_{kj} A_{ikt} + e_{ijt} \tag{B.2}$$

where X_{it} and Z_{jt} are characteristics of the consumer and the alternative in period t, respectively, $y_{i,t-1}$ denotes the alternative chosen by the patient in period $t-1$, and $A_{ikt} = \mathbb{I}[y_{i,t-1} = k]$.

This formulation is very general as it allows for general compatibility patterns: γ_{kj} is the utility premium the consumer gets from choosing alternative j if the alternative chosen in the previous occasion was k.

Consider the additional restriction that the consumer only receives a utility premium from choosing the same alternative as in the previous period. In other words, we have: 1) $k \neq j \Rightarrow \gamma_{kj} = 0$; 2) $\gamma_{jj} \geq 0$. Therefore, utility is given by:

$$u_{ijt} = g(X_{it}, Z_{jt}) + \gamma_{jj} \mathbb{I}[y_{i,t-1} = j] + e_{ijt} \tag{B.3}$$

Define the portion of utility that depends on current consumer and alternative characteristics as $u^*_{ijt} = g(X_{it}, Z_{jt}) + e_{ijt}$, and let $u^*_{it} = \text{argmax}_j \ u^*_{ijt}$. The assumptions above and transitivity imply that for any counterfactual period $t-1$ choice j, $h_t(j) \in \{j, u^*_t\}$.

Consider a consumer i with $y_{i,t-1} = j \neq k$. Then, we have:

$$u_{ijt} = u^*_{ijt} + \gamma_{jj} \geq u^*_{ijt} \ \text{ and } \ u_{ikt} = u^*_{ikt} \tag{B.4}$$

Then, there are two possibilities:

1. $u^*_{it} = j \Rightarrow u^*_{ijt} \geq u^*_{ikt} = u_{ikt} \Rightarrow u_{ijt} \geq u_{ikt}$
2. $u^*_{it} = q \neq j \Rightarrow u^*_{iqt} = u_{iqt} \geq u^*_{ikt} = u_{ikt} \ \ \forall k \neq j$

Then, a consumer who chooses alternative j in $t-1$ chooses either j or u^*_t in period t.

In my setting, the restrictions imply that, for $j \neq k \neq z$, we have:

1. $h_t(k) = j \Rightarrow h_t(j) = j$
2. $h_t(k) = j \Rightarrow h_t(z) \in \{z, j\}$

In the application, these restrictions are imposed ex-ante because of dimensionality issues, effectively restricting the set of possible types. However, whenever the number of alternatives and time periods is not restrictive we might want to treat them as regular identifying assumptions and evaluate their power independently of other restrictions.

Appendix C

Appendix 1 for Chapter 3

C.1. Definitions of pre and post-reform periods

The start of the new payment system for an LTCH is linked to the start date of the hospital's cost reporting period. Therefore, the pre and post-reform periods are hospital-specific. To define these periods, I use the 2016 cost reporting period of a hospital to determine the quarter in which the new payment system began for this facility. I refer to this quarter as the "switch" quarter. I refer to quarters before it as the pre-reform period and to quarters after it as the post-reform period. If the new system started the first day of the switch quarter, I include this quarter in the post-reform period; otherwise, I exclude it from the analysis. For example, suppose that the cost reporting periods of hospitals A and B start January 1st and February 1st, respectively. The first quarter of 2016 is the switch quarter for both hospitals. For hospital A, all discharges in the switch quarter are paid under the new payment system, so the post-reform period includes this quarter. For hospital B, some discharges in the switch quarter are paid under the new payment system while others are still paid under the old system. The post-reform period does not include this quarter.

In the main analysis, I use FY 2015 and 2017 for the comparison of pre and post-reform periods. In this case, the periods are the same for all hospitals. Moreover, all hospitals were subject to the same payment system in a given period. However, hospitals differ in terms of the amount of time they had been operating under the new payment system in the post-reform period.

I performed the analysis using two alternative definitions. In the first alternative, I consider the four quarters before and after the switch quarter. The length of the periods is the same for all hospitals. However, in a given calendar quarter there might be hospitals operating under different payment systems. In the second alternative, I consider the widest symmetric time window before and after the switch quarter of a hospital to define the pre and post-reform periods. For example, if a hospital was subject to the new payment system for six quarters during fiscal years 2016 and 2017, then the pre and post-reform periods for this facility consist of the six quarters before and after the switch quarter. In this case, the length of the periods differs across hospitals. As in the first

Table C.1: Determination of pre and post-reform periods

	Alternative 1		Alternative 2	
Cost report FY 2016	Pre-reform	Post-reform	Pre-reform	Post-reform
10/01/15-09/30/16	Q4 14 - Q3 15	Q4 15 - Q3 16	Q4 13 - Q3 15	Q4 15 - Q3 17
11/01/15-10/31/16	Q4 14 - Q3 15	Q1 16 - Q4 16	Q1 14 - Q3 15	Q1 16 - Q3 17
01/01/16-12/31/16	Q1 15 - Q4 15	Q1 16 - Q4 16	Q2 14 - Q4 15	Q1 16 - Q3 17
04/01/16-03/31/17	Q2 15 - Q1 16	Q2 16 - Q1 17	Q4 14 - Q1 16	Q2 16 - Q3 17

Notes: This table shows alternative definitions of pre and post-reform for hypothetical hospitals. The first column indicates the start and end dates for the FY 2016 cost reporting period of the hospital. The second and fourth columns indicate the pre-reform period under the two alternatives considered, while the third and fifth columns indicate the corresponding post-reform periods.

case, in a given calendar quarter we can have hospitals operating under different payment systems. Table C.1 shows four examples that illustrate the determination of pre and post-reform periods for the analysis.

C.2. Medicare LTCH payments

In this section, I describe in detail how Medicare payments to LTCHs are determined. This material is based on IPPS / LTCH PPS Proposed and Final Rules issued by CMS, MedPAC reports (MedPAC, 2012; MedPAC, 2015; MedPAC, 2016), and CMS (2016). I start by defining the different types of payments used:

C.2.1. Types of payments

MS-LTC-DRG payment. The MS-LTC-DRG payment is a fixed amount for all patients with a given MS-LTC-DRG, independently of length of stay, cost of care, diagnoses, procedures, or place of discharge. This payment is determined as:

$$p = \text{base payment} \times \text{DRG weight}$$

MS-LTC-DRG per diem payment. It is obtained by dividing the full MS-LTC-DRG payment by the average length of stay for the MS-LTC-DRG and multiplying the result by the actual length of stay of the case.

Estimated cost of the case. It is obtained by multiplying the LTCH's cost-to-charge ratio (CCR) by the Medicare allowable charges for the LTCH case. An LTCH's CCR is calculated by dividing the LTCH's total Medicare costs by its total Medicare charges.

IPPS comparable amount. To determine this amount, I first compute what the payment for the case would have been under the IPPS. The IPPS comparable amount is then calculated as a per diem by dividing the full IPPS payment by the average length of stay for that MS-DRG and multiplying the result by the actual length of stay of the case. The IPPS comparable amount is capped at the full IPPS payment.

C.2.2. Standard rate

The standard rate applies to all LTCH discharges in the pre-reform period and to post-reform discharges that meet the criteria for exclusion from the site-neutral rate. Two types of payments are used:

Short stay outlier. The short-stay outlier (SSO) payment applies to patients whose length of stay is less than 5/6th of the geometric mean length of stay for that MS-LTC-DRG in the previous fiscal year. Medicare pays an SSO case using the least of the following amounts:

1. The MS-LTC-DRG payment.

2. 120 percent of the MS-LTC-DRG per diem payment.

3. 100 percent of the estimated cost of the case.

4. For SSO discharges on or before December 28, 2012: a blend of the IPPS comparable amount and 120 percent of the MS-LTC-DRG per diem payment.

5. For SSO discharges on or after December 29, 2012, one of these amounts:

   (a) If the length of stay is greater than the IPPS-comparable threshold for the corresponding MS-LTC-DRG (defined as one standard deviation from the geometric average length of stay for the corresponding DRG under the IPPS): a blend of the IPPS comparable amount and 120 percent of the MS-LTC-DRG per diem payment.

   (b) If the length of stay is equal to or less than the IPPS-comparable threshold for the corresponding MS-LTC-DRG: the IPPS comparable amount.

C.2.3. Site-neutral rate

Post-reform discharges that do not meet the criteria from exclusion from the site-neutral rate are reimbursed at the lower of the estimated cost of the case and the IPPS comparable amount. During the transition period, the payment for these cases is a 50/50 blend of the standard and site-neutral rates.

C.2.4. High cost outlier payment

This payment covers part of the costs corresponding to a patient whose cost of care greatly exceeds the MS-LTC-DRG reimbursement. This payment applies if the estimated cost of the case exceeds the outlier threshold, which is defined as the MS-LTC-DRG payment plus a fixed loss amount. For patients in the high cost outlier category, Medicare reimburses 80% of the costs incurred above the threshold. Short-stay outliers are also eligible for outlier payments if their costs exceed the outlier threshold; the applicable short-stay outlier payment is used to determine the outlier threshold for these cases. I ignore these payments in the analysis.

C.3. Construction of the LTCH payment schedule

In the discharge data, I observe covered charges and the payment the hospital receives from Medicare for the actual length of stay of the patient. For the analysis, I need to determine the payment for alternative length of stays to recover the payment schedule faced by the LTCH. I construct the payment schedule directly from the payment policy described above. The elements required to construct the payment schedule are the DRG to which the LTCH stay is grouped, the corresponding DRG weights and average length of stay under the LTCH-PPS and the IPPS, the IPPS threshold, and the estimated cost of the case. The DRG is indicated in the discharge record of the patient. The DRG weights and average length of stay are obtained from CMS. To compute the estimated cost of the case, I use the CCR from the Impact Files and assume that covered charges are a linear function of the length of stay of the patient. The rates required to compute the payments (for example, the base payment for the MS-LTC-DRG payment) are obtained from the Long-Term Care Hospital PPS PC Pricer and the Inpatient PPS PC Pricer[1].

C.4. Criteria for exclusion from the site-neutral rate

The criteria for exclusion from the site-neutral payment rate are (CMS, 2016):

1. The LTCH discharge does not have a principal diagnosis relating to a psychiatric diagnosis nor to rehabilitation. An LTCH discharge meets this criterion if it is not assigned to one of the following MS-LTC-DRGs: 876, 880, 881, 882, 883, 884, 885, 886, 887, 894, 895, 896, 897, 945, and 946.

2. The LTCH admission must have been immediately preceded by a subsection (d) hospital stay. This criterion is met if admission to the LTCH occurred either on the date of or the calendar date after the discharge from the preceding subsection (d) hospital.

[1] For more details, see: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PCPricer/index.html

3. The patient discharged from the LTCH must have either:

    (a) Spent at least 3 days in an intensive care unit during the immediately preceding subsection (d) hospital stay (ICU criterion). The LTCH discharge fulfills the requirements of the ICU criterion if the claim from the subsection (d) hospital stay that immediately preceded the admission to the LTCH indicates receipt of at least 3 days of care in an ICU using revenue center codes 020X or 021X.

    (b) Received at least 96 hours of respiratory ventilation services during the LTCH stay (ventilator criterion). The LTCH discharge fulfills the requirements of the ventilator criterion if the hospital reports procedure code 5A1955Z (International Classification of Diseases, 10th Revision, Procedure Coding System (ICD-10-PCS)) on its LTCH claim.

Appendix D

Appendix 2 for Chapter 3

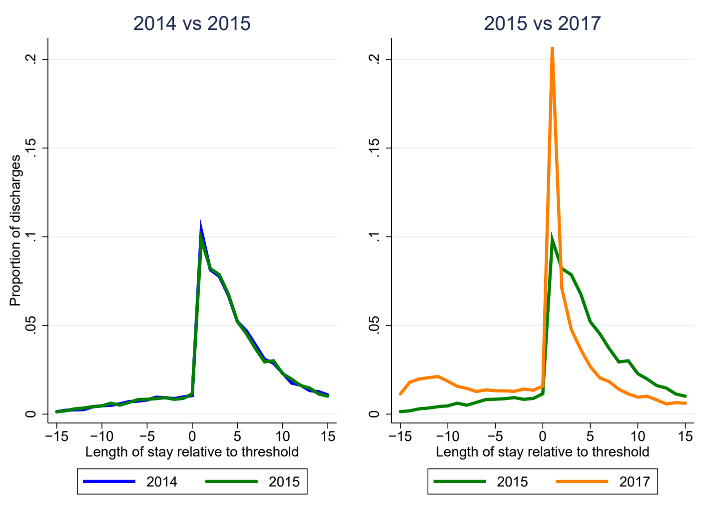
D.1. Additional figures of discharge patterns

This Appendix contains additional figures that support the analysis of discharge patterns in the main text. Figures D.1 through D.4 show discharge patterns for site-neutral cases, by type of hospital. The left panel of each figure compares the discharge patterns in FY 2014 and 2015. The right figure of each panel reproduces the comparison of discharge patterns in FY 2015 and 2017 discussed in the main text. During FY 2014 and 2015, all hospitals were subject to the old payment system. During FY 2017, all hospitals were subject to the new payment system. In all cases, the discharge patterns correspond to patients discharged downstream.

Figures D.5 through D.12 show discharge patterns for site-neutral cases by period, but were constructed reweighting post-reform cases so they reflect the DRG and hospital composition of the pre-reform cases.

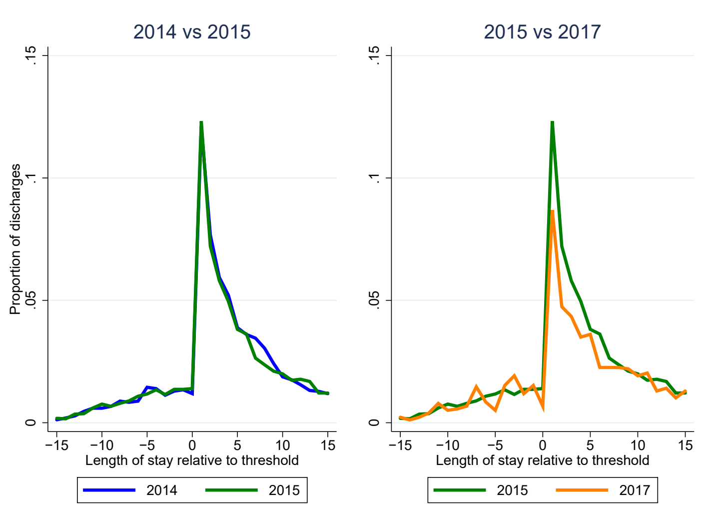
Finally, Figures D.13 through D.20 are similar to Figures 3.6 through 3.13 discussed in the main text: they show discharge patterns by type of patient and period, but using the year before and after the start of the new payment system of a hospital for defining pre and post-reform periods (this definition corresponds to the first alternative discussed in Section C.1).

Figure D.1: Proportion of discharges, site-neutral cases, other for-profit hospitals



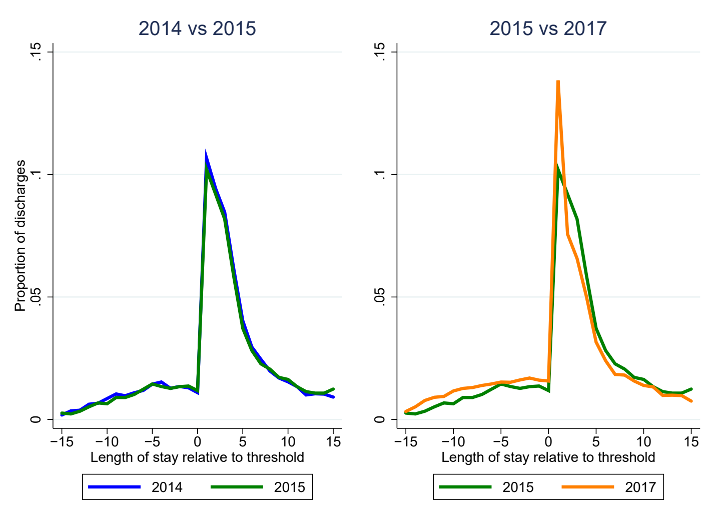
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.2: Proportion of discharges, site-neutral cases, Select hospitals



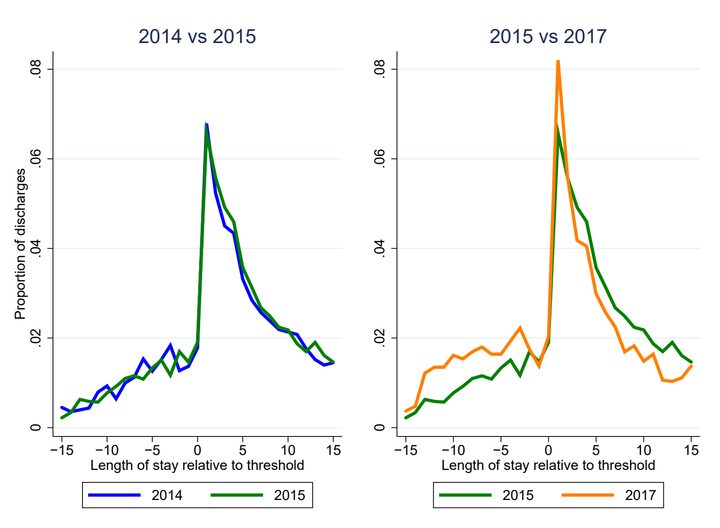
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.3: Proportion of discharges, site-neutral cases, Kindred hospitals



Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.4: Proportion of discharges, site-neutral cases, nonprofit hospitals

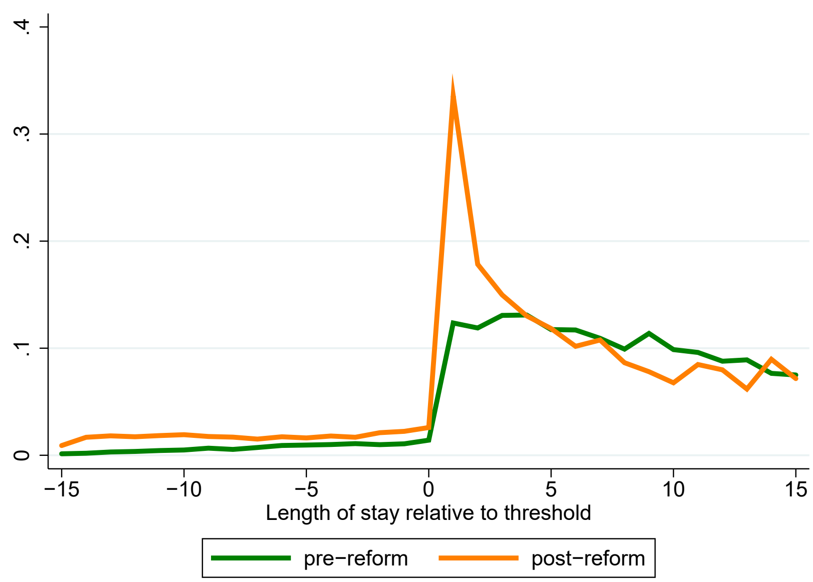


Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.5: Proportion of discharges, site-neutral cases, other for-profit hospitals

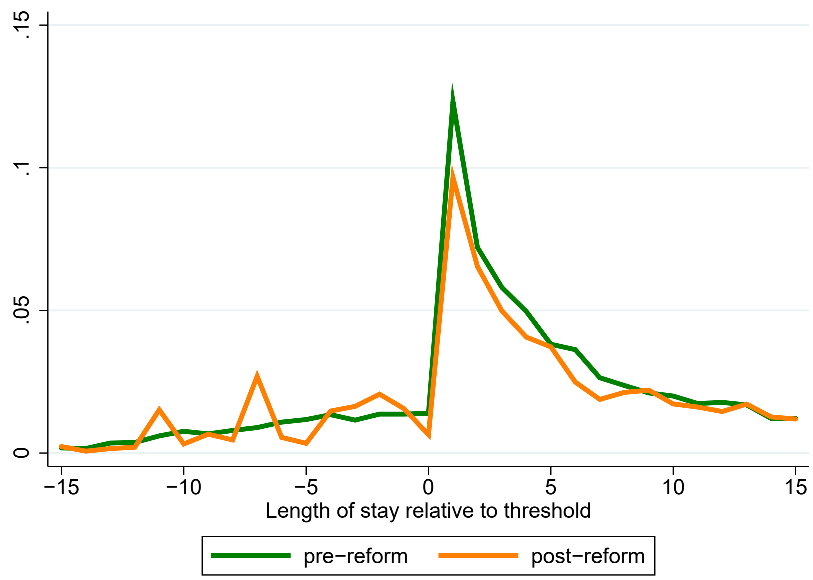
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.6: Discharge hazard rate, site-neutral cases, other for-profit hospitals



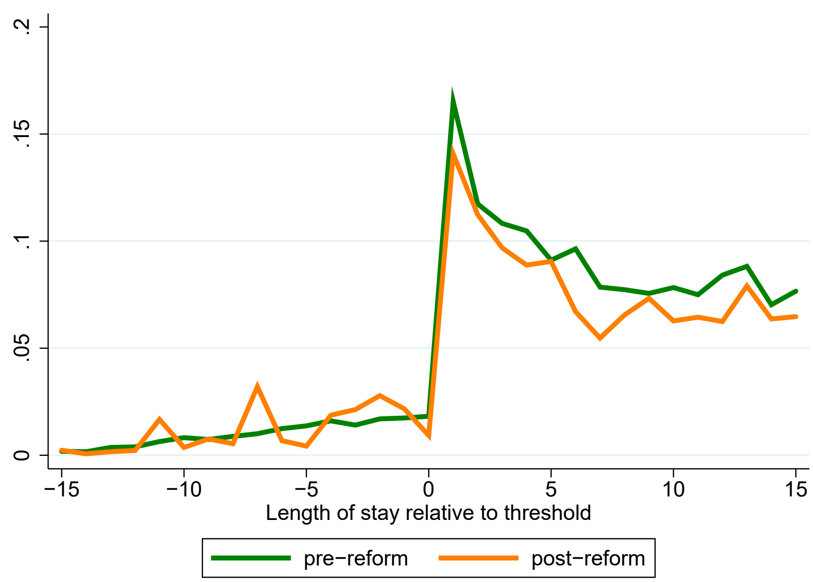
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure D.7: Proportion of discharges, site-neutral cases, Select hospitals



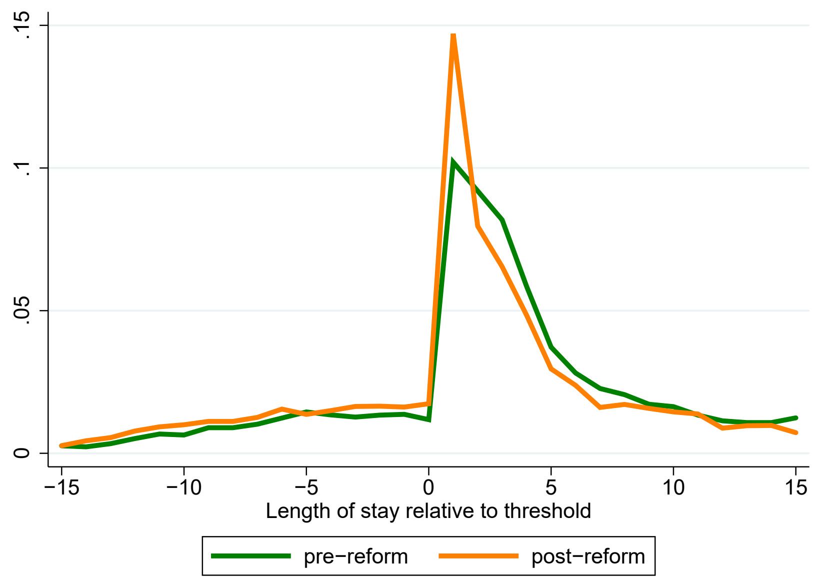
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.8: Discharge hazard rate, site-neutral cases, Select hospitals



Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure D.9: Proportion of discharges, site-neutral cases, Kindred hospitals

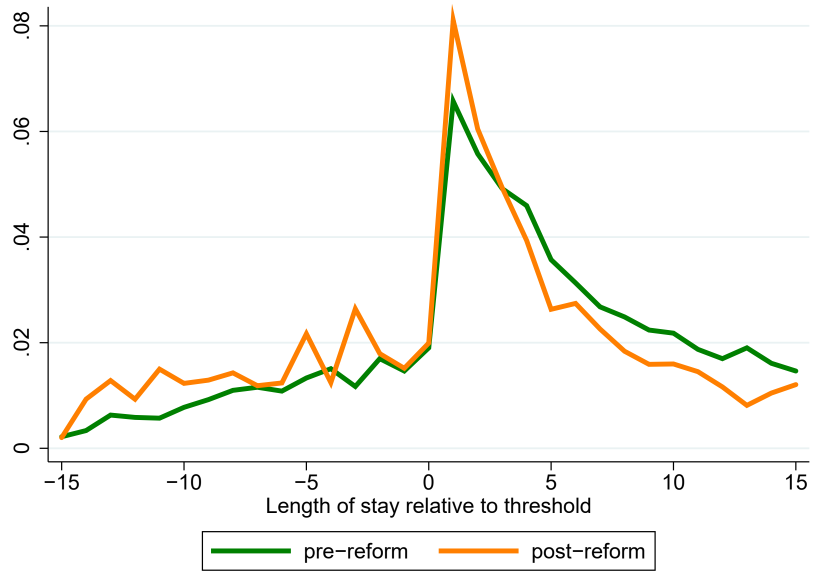


Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.10: Discharge hazard rate, site-neutral cases, Kindred hospitals

Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure D.11: Proportion of discharges, site-neutral cases, nonprofit hospitals

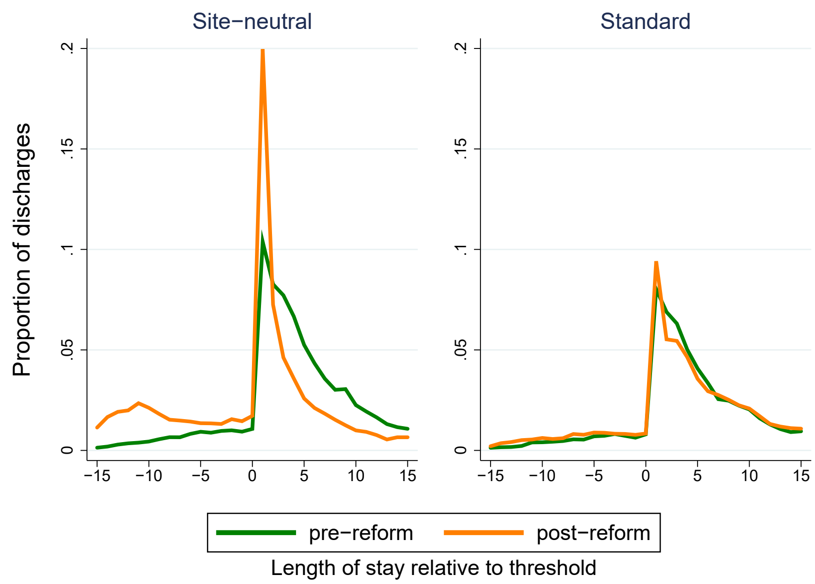


Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold) .

Figure D.12: Discharge hazard rate, site-neutral cases, nonprofit hospitals

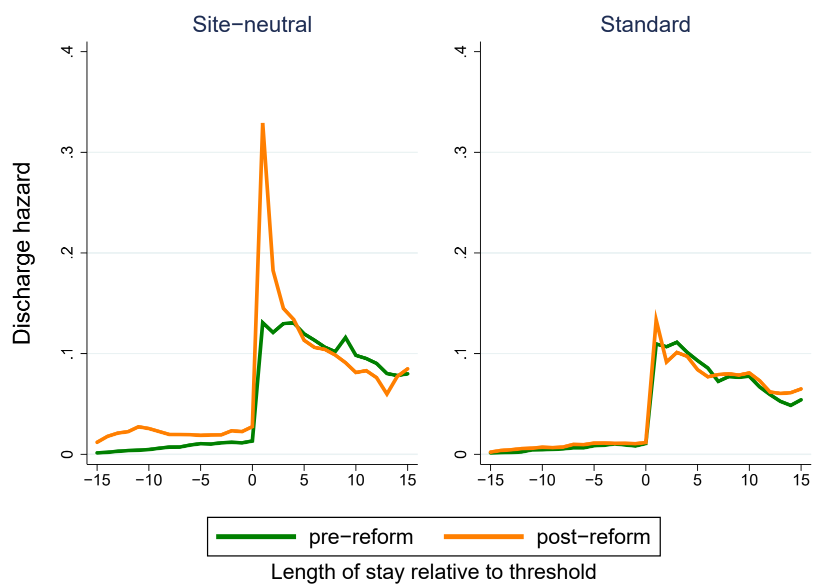
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure D.13: Proportion of discharges, other for-profit hospitals



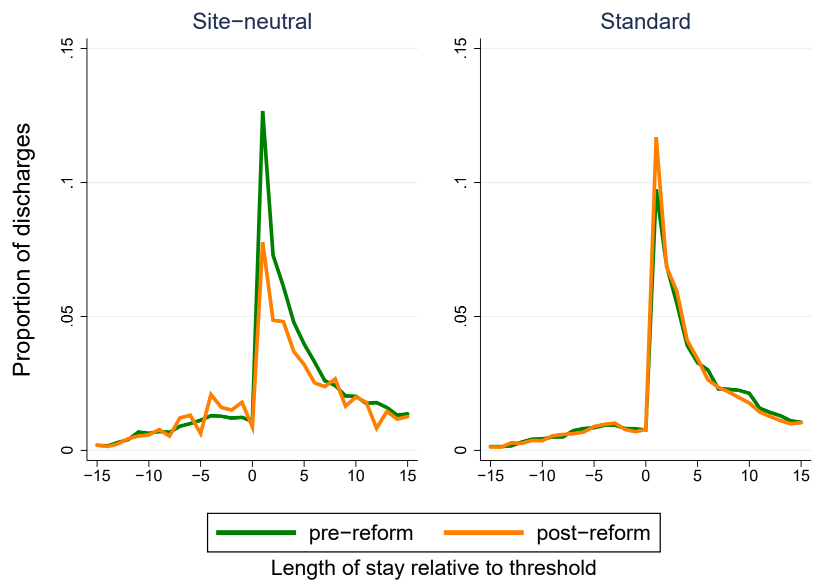
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.14: Discharge hazard rate, other for-profit hospitals



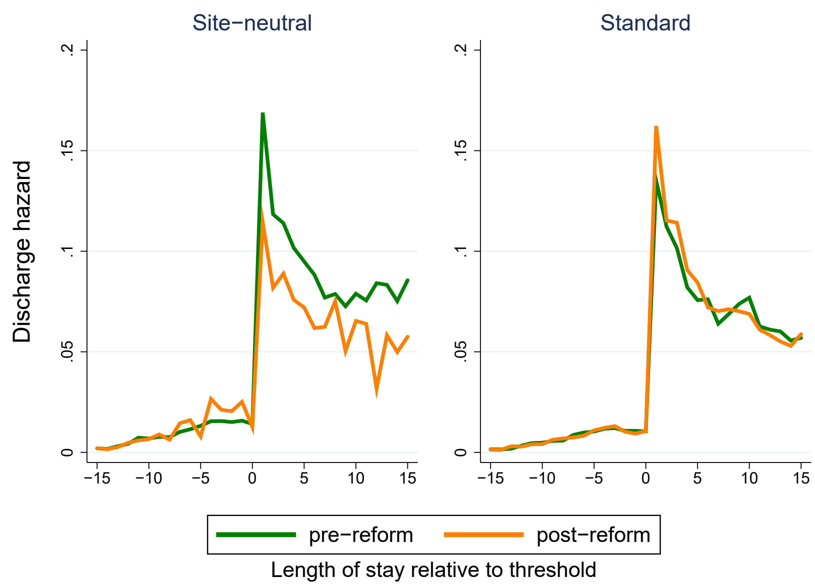
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure D.15: Proportion of discharges, Select hospitals



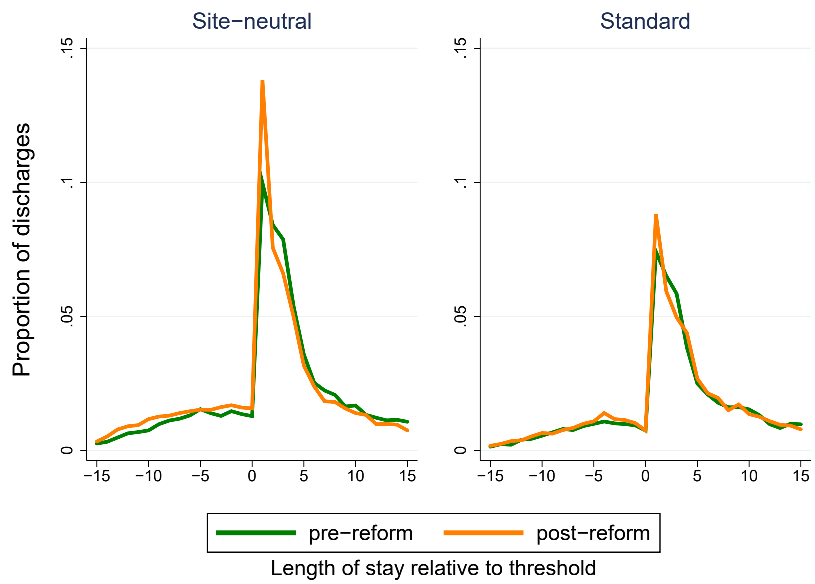
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.16: Discharge hazard rate, Select hospitals



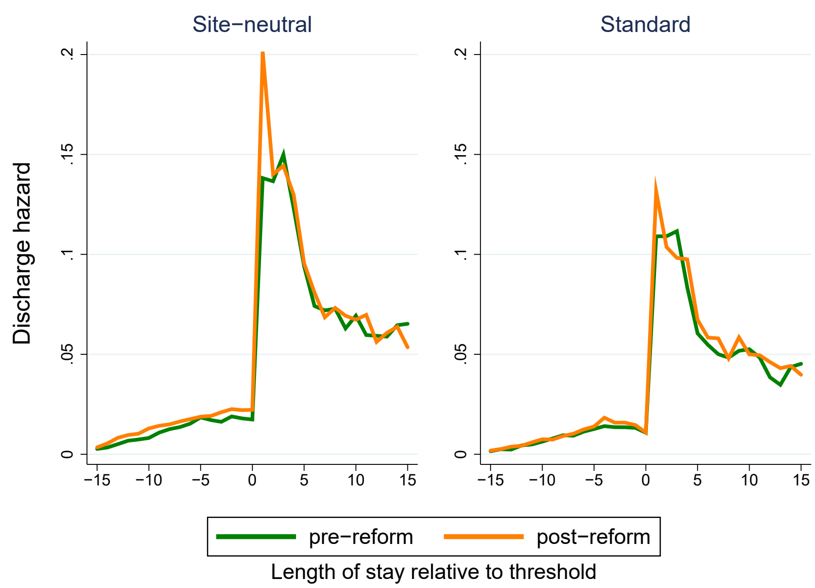
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure D.17: Proportion of discharges, Kindred hospitals



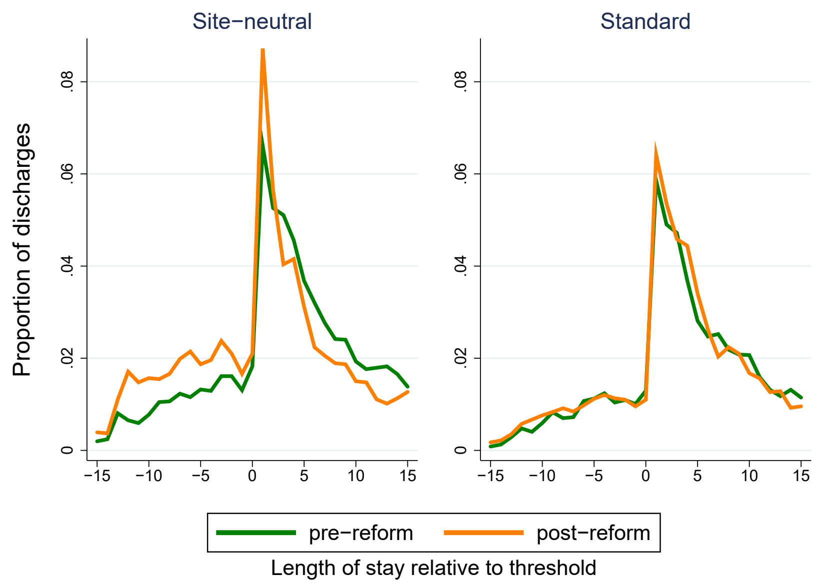
Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold).

Figure D.18: Discharge hazard rate, Kindred hospitals



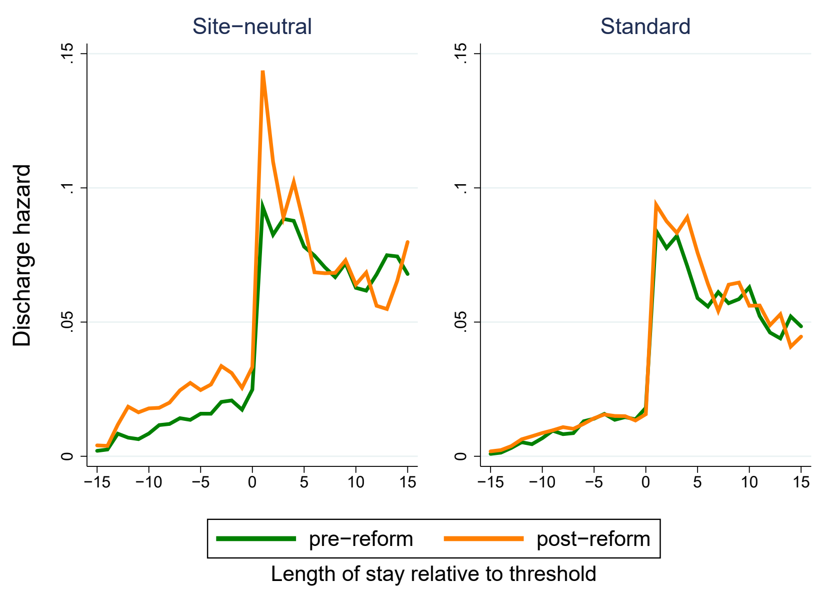
Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Figure D.19: Proportion of discharges, nonprofit hospitals



Notes: This figure shows the proportion of discharges as a function of length of stay (relative to the SSO threshold) .

Figure D.20: Discharge hazard rate, nonprofit hospitals



Notes: This figure shows the discharge hazard rate as a function of length of stay (relative to the SSO threshold) .

Appendix E

Appendix 3 for Chapter 3

In this Appendix, I expand the scale of previous work on the discharge behavior of LTCHs during the pre-reform period. First, I compare the discharge patterns of Medicare and privately insured LTCH patients in the states of Florida and Texas. Second, I analyze the discharge patterns of Medicare LTCH patients in selected DRGs; in particular, I show how changes in the SSO threshold across years affect discharge behavior.

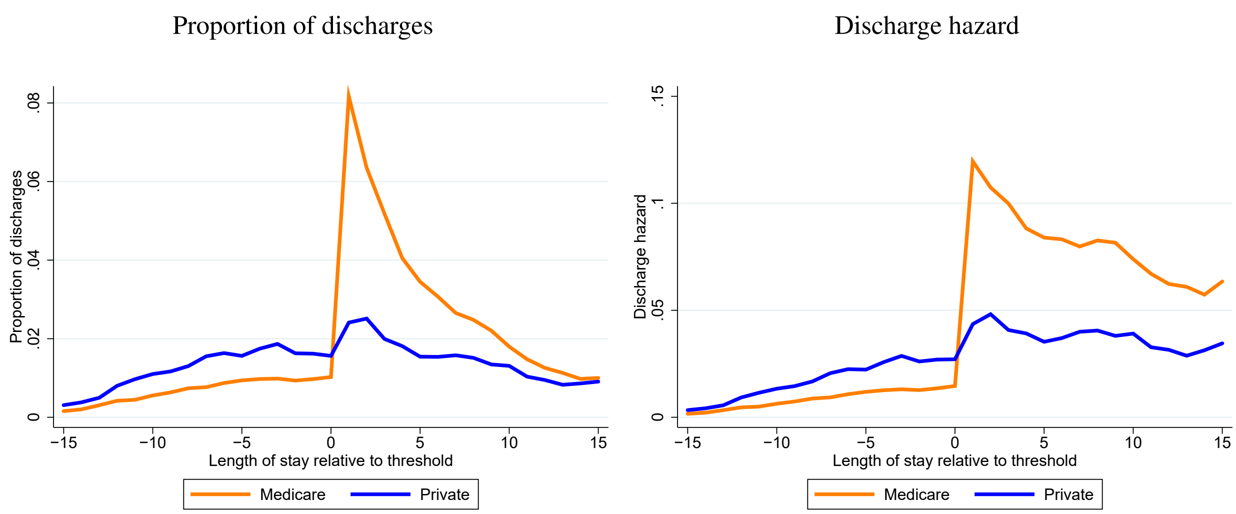
E.1. Discharge patterns: Medicare and privately insured patients

I use state hospital discharge databases for the periods from 2008 to 2015 (Florida) and from 2008 to 2011 (Texas). These datasets contain information on the universe of discharges from hospitals in the corresponding state during a calendar year. The main difference with the MEDPAR dataset is that the state databases contain information for all discharges, independently of the type of insurance of the patient. I consider discharges from hospitals that are certified as LTCHs by Medicare.

As in the main analysis, I determine the length of stay of the patient relative to the corresponding SSO threshold. Then, for each day I compute the fraction of patients discharged and the discharge hazard rate, by discharge destination (downstream, upstream, and death). I recover the discharge patterns separately for Medicare patients and patients with commercial insurance. Here, I report the results corresponding to downstream discharges. While I know the form of the payment schedule LTCHs face for Medicare patients, I do not have any information about the payment schedule for privately insured patients.

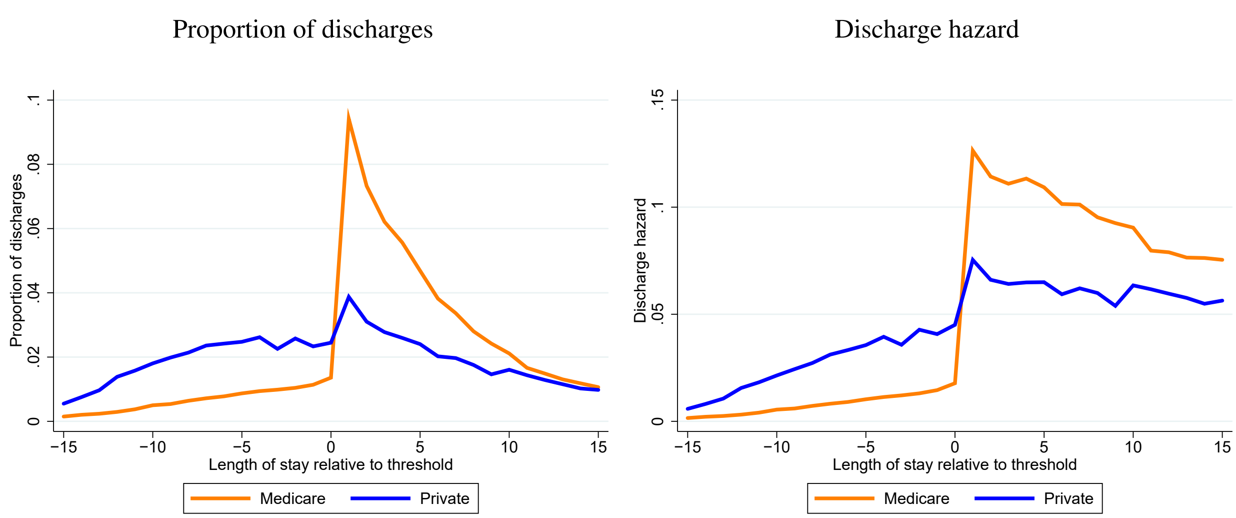
Figure E.1 shows the results for Florida, while Figure E.2 shows the results for Texas. As in the case of Medicare discharges documented by Einav et al. (2017) and Eliason et al. (2016) and in Chapter 3, for Medicare patients in these states there is a large spike in discharges at the SSO threshold. In the case of privately insured patients, the jump in discharges at the SSO threshold is much smaller.

Figure E.1: Discharge patterns, Florida, by type of insurance



Notes: This figure shows the proportion of discharges and the discharge hazard rate, by length of stay and type of insurance. The discharge patterns correspond to downstream discharges.

Figure E.2: Discharge patterns, Texas, by type of insurance



Notes: This figure shows the proportion of discharges and the discharge hazard rate, by length of stay and type of insurance. The discharge patterns correspond to downstream discharges.

E.2. Discharge patterns: Changes in thresholds across years

I examine the discharge behavior of hospitals for Medicare patients assigned to selected DRGs. In particular, I consider six DRGs for which the SSO threshold changes from year to year. Table E.1 shows the evolution of the SSO threshold for these DRGs during the period from FY 2011 to FY 2015.

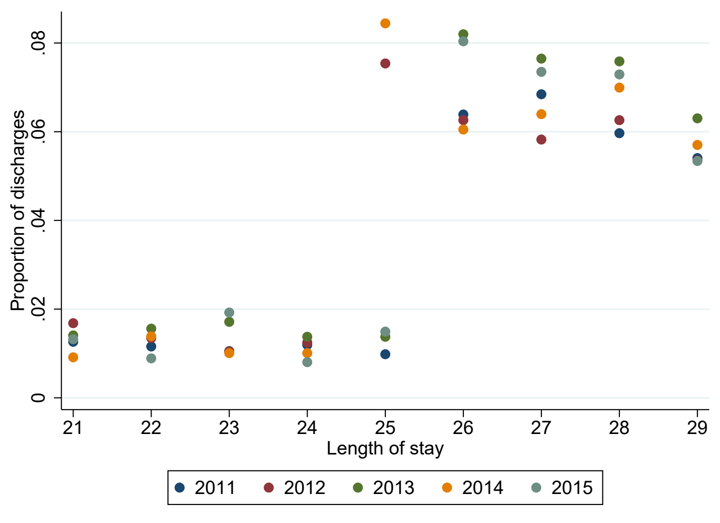
I examine discharge patterns separately by DRG and year. Specifically, I determine the proportion of discharges in a given year that take place each day, but I do not normalize length of stay relative to the SSO threshold. Figures E.3 - E.8 show the discharge patterns. These results provide clear evidence that LTCHs modify treatment decisions in response to financial incentives.

Table E.1: Evolution of the short-stay outlier threshold, selected DRGs and years

DRG	2011	2012	2013	2014	2015
539	25.7	24.9	25.0	24.8	25.2
208	18.6	18.5	17.8	18.2	17.7
559	21.3	22.2	22.1	21.7	22.0
870	26.4	26.3	24.7	25.6	24.8
862	21.1	21.1	20.9	21.1	20.6
004	36.4	37.3	35.8	36.3	35.9

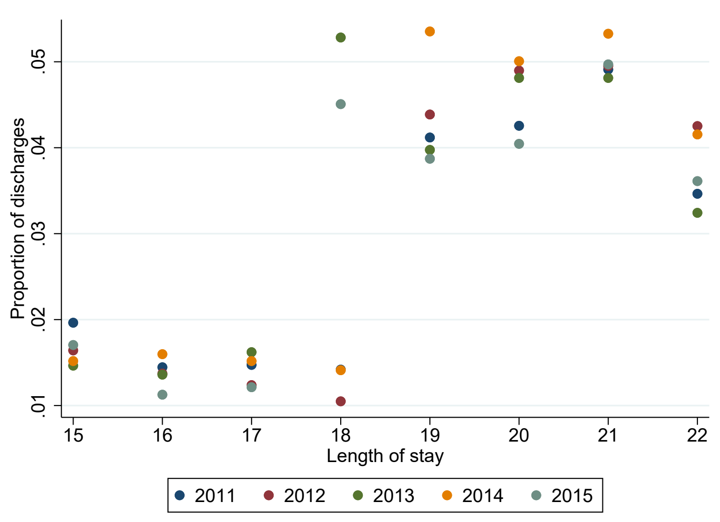
Notes: This table shows the evolution of the SSO threshold during the period from 2011 to 2015 for selected DRGs.

Figure E.3: Discharge patterns, DRG 539



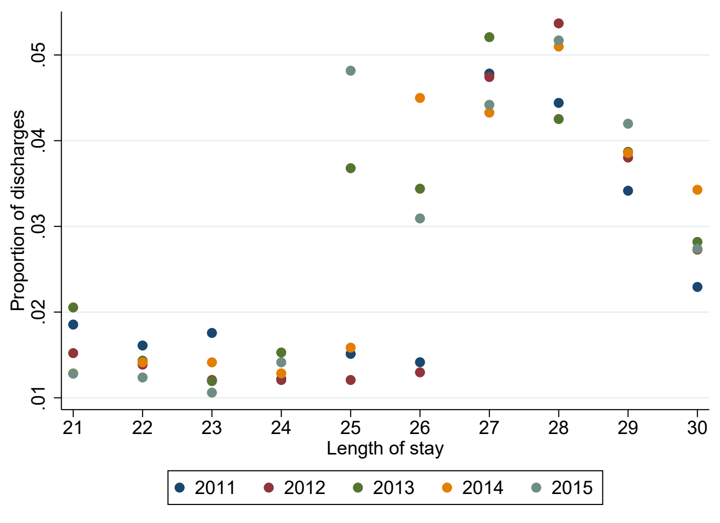
Notes: This figure shows the proportion of discharges that take place each day, by fiscal year.

Figure E.4: Discharge patterns, DRG 208



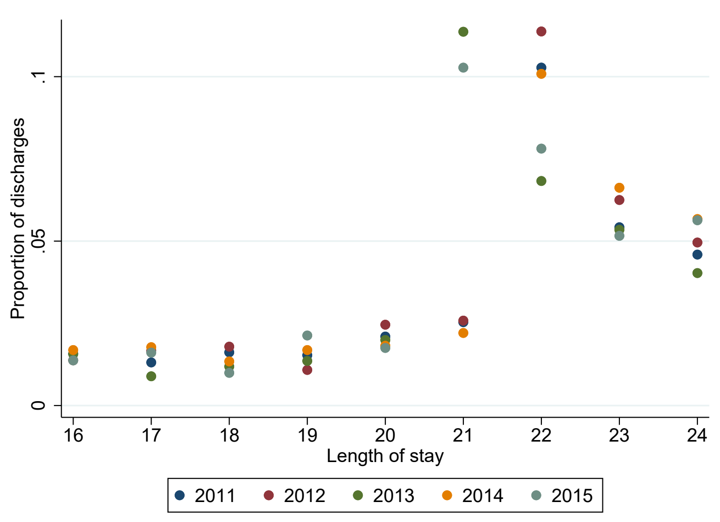
Notes: This figure shows the proportion of discharges that take place each day, by fiscal year.

Figure E.5: Discharge patterns, DRG 870



Notes: This figure shows the proportion of discharges that take place each day, by fiscal year.

Figure E.6: Discharge patterns, DRG 862

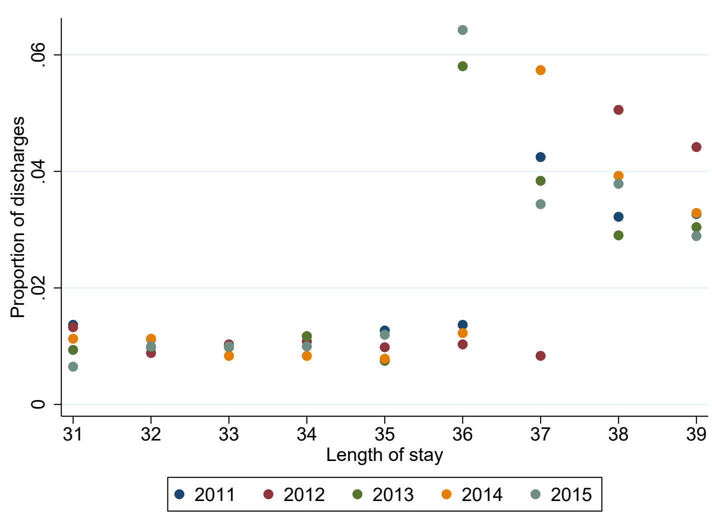


Notes: This figure shows the proportion of discharges that take place each day, by fiscal year.

Figure E.7: Discharge patterns, DRG 559

Notes: This figure shows the proportion of discharges that take place each day, by fiscal year.

Figure E.8: Discharge patterns, DRG 004



Notes: This figure shows the proportion of discharges that take place each day, by fiscal year.