

Essays in Applied Microeconomics with Policy

Implications

by

Christopher Geissler

Department of Economics
Duke University

Date: _____

Approved:

Andrew Sweeting, Supervisor

James Roberts

Frank Sloan

Patrick Bayer

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2013

ABSTRACT

Essays in Applied Microeconomics with Policy Implications

by

Christopher Geissler

Department of Economics
Duke University

Date: _____

Approved:

Andrew Sweeting, Supervisor

James Roberts

Frank Sloan

Patrick Bayer

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2013

Copyright © 2013 by Christopher Geissler
All rights reserved except the rights granted by the
Creative Commons Attribution-Noncommercial Licence

Abstract

My dissertation focuses on employing microeconomic techniques to study markets and questions that are important and complex, and also have potential policy implications. Two of my chapters analyze the health industry with an emphasis on hospitals, patient welfare, and regulation. The remaining chapter focuses on the housing market in Los Angeles and explores real estate flipping.

The second chapter of my dissertation studies the impact of state level regulations on hospital bed capacity decisions. The regulations are intended to decrease hospital investments without diminishing patient access. I find that the regulation decreases total hospital investment in bed capacity as expected. When running simulations to estimate how hospitals would behave differently were the regulatory policy changed, I find that total patient utility is negatively affected by the presence of the regulation as many patients get turned away from their preferred hospital due to overcrowding. This analysis has important policy implications as it suggests that the regulation has been ineffective in ensuring that patient welfare was unharmed by the restrictions.

The third chapter is based on joint research with Patrick Bayer and James W. Roberts and studies the housing market in the Los Angeles metropolitan area from 1988 to 2009. Using novel data, I identify which housing transactions involve flippers who aim not to live in the house, but rather to quickly resell it for financial gain. I find that flipper behavior varies based on how frequently I observe the individual engage in such behavior. Experienced flippers, who are observed to flip many houses

in the data, target homes being sold at below market value and earn their returns from buying them at a discount. Their effect on long term prices in the neighborhood is negligible. Inexperienced flippers who are less active, seek to earn their profits by timing the market and are more active when house prices were rapidly appreciating from 1999 to 2005. Their activity increases housing prices in the neighborhood in the short term, but decreases them in the long term. Such results are consistent with the claim that real estate flipping contributed to the housing bubble.

The fourth chapter of my dissertation again focuses on the hospital industry and looks at the question of how patient composition changes as a hospital becomes busier and has to turn patients away. I develop a theoretical model which predicts that hospitals are more likely to turn away less profitable patients. As a result, when a hospital becomes more full and therefore is more likely to have to turn patients away, its composition of patients will change and become more profitable on the whole. I test this theory by empirically analyzing the effect of hospital congestion on the composition of hospital patients using hospital discharge data. The findings are consistent with my theoretical model as when hospitals become more crowded, the fraction of uninsured patients and mental health patients (who are typically not profitable to a hospital) decreases. This result suggests that hospitals are more likely to turn away unprofitable patients while continuing to admit more profitable patients.

Contents

Abstract	iv
List of Tables	x
List of Figures	xii
List of Abbreviations and Symbols	xiii
Acknowledgements	xvi
1 Introduction	1
2 The Impact of State Regulations on Hospital Capacity Adjustment Decisions	7
2.1 Introduction	7
2.2 Background on CON	12
2.2.1 History of CON	12
2.2.2 CON Justification and Procedures	13
2.2.3 Debate Over CON's Impact	16
2.3 Data	19
2.3.1 State Inpatient Data	19
2.3.2 Hospital Data	20
2.3.3 Combined Dataset	20
2.3.4 Is State Level Data Representative?	21
2.4 Empirical Model	22

2.4.1	Turnaway Probability	23
2.4.2	Patient Utility	24
2.4.3	Hospital Capacity Decisions	28
2.5	Results	32
2.5.1	Patient Utility	33
2.5.2	Hospital Capacity Decisions	35
2.6	Simulations	39
2.6.1	Methodology	39
2.6.2	Fit of Simulations	41
2.6.3	Estimates	43
2.7	Conclusion	46
3	Speculators and Middlemen: The Role of Flippers in the Housing Market	48
3.1	Introduction	48
3.2	A Conceptual Framework	52
3.2.1	Flippers as Middlemen	52
3.2.2	Flippers as Speculators	54
3.3	Data	55
3.3.1	Flippers	59
3.3.2	Purchase Activity by Flippers	62
3.3.3	Flipper Holding Times	65
3.4	Measuring the Sources of Flipper Returns - Research Design	66
3.5	The Sources of Flipper Returns - Baseline Results	72
3.5.1	Flippers' returns	72
3.6	Robustness	77
3.6.1	Do Flippers Sell Winners and Hold Losers?	78

3.6.2	Are Results Driven by Flippers Buying Fixer-Uppers?	80
3.6.3	Selective Survival - The Dynamic Pattern of Returns	81
3.6.4	Neighborhood Targeting	82
3.7	Submarket Level Results	85
3.7.1	Which Submarkets are Targeted?	86
3.8	How Well-Informed Are Speculators?	89
3.9	Conclusion	93
4	Turnaways and Patient Composition: An Empirical Analysis of Hospitals	95
4.1	Introduction	95
4.2	Background Literature	97
4.3	Theoretical Model	99
4.3.1	Generalizing the Model	103
4.4	Data	103
4.4.1	Inpatient Data	104
4.4.2	Hospital Data	104
4.4.3	Final Dataset	105
4.5	Empirical Model	107
4.5.1	Estimation of Turnaway Probabilities	108
4.5.2	Regression Framework	110
4.6	Results	111
4.6.1	Admission Type	111
4.6.2	Payment Type	113
4.6.3	Race and Diagnosis Related Group	114
4.7	Extensions	116
4.8	Conclusion	118

5	Conclusions	119
A	Estimation of Hospital Turnaway Probabilities and Substitution Patterns	122
	Bibliography	125
	Biography	130

List of Tables

2.1	Summary statistics	21
2.2	National versus sample regressions	22
2.3	Patient utility regressions	33
2.4	Hospital capacity adjustment regressions	36
2.5	Mean values in data versus simulations	42
2.6	Regressions analyzing fit of simulations	42
2.7	Comparison of mean values for CON regulations and no CON	43
2.8	Comparison of simulation means by hospital characteristics	44
2.9	Utility comparison based on CON status	46
3.1	Transaction-level summary statistics	57
3.2	House summary statistics by flipper type	64
3.3	Holding times by flipper type	65
3.4	Regression coefficients for all flippers	73
3.5	Regression coefficients by flipper type	76
3.6	Sources of return by flipper type	77
3.7	Robustness checks	79
3.8	Flipper regression coefficients by flip number	83
3.9	Source of returns by flipper type at neighborhood level	85
3.10	Effect of flipper activity on neighborhood prices	87
3.11	Effect of earlier appreciation on neighborhood prices	87

3.12	Effect of flipper activity and earlier appreciation on neighborhood prices	87
4.1	Summary statistics	106
4.2	Regressions on admission type	112
4.3	Regressions on payment type	113
4.4	Regressions on race and diagnosis related group	115

List of Figures

2.1	CON application process	15
3.1	Transaction volume by county	58
3.2	Level of flipper activity over time	60
3.3	Flipper sales over time by type	63
3.4	Identification strategy	70
3.5	Map of voting districts used as submarkets	84
3.6	Houses held by speculators	91
3.7	Speculator activity based on expected appreciation	92

List of Abbreviations and Symbols

Symbols

Throughout my dissertation, I use a mixture of Greek and Roman letters to represent variables in my theoretical and empirical models. In many cases, these variables contain subscripts in order to index the observation to which the variable refers. For example, β_{ht} may represent the β value for hospital h in time period t . When introducing new variables, I explain what they refer to and how to interpret any subscripts used.

Abbreviations

Below I list the set of abbreviations used in my dissertation as well as a brief description of the term.

ADC	Average Daily Census, the average number of patients in a hospital for a given time period.
AHA	American Hospital Association, an organization that represents hospitals in the United States and collects annual data on these hospitals.
AHRQ	Agency for Healthcare Research and Quality, the federal organization that governs the Healthcare Cost and Utilization project.
BBL	Bajari, Benkard, and Levin, an empirical model introduced in Bajari et al. (2007) that includes two stages of estimation.
CON	Certificate of Need, a state level hospital regulation requiring hospitals to apply to a state health board before making large investments.

DRG	Diagnosis Related Group, a method of categorizing the diagnosis of a hospital patient.
FHFA	Federal Housing Finance Agency, an agency that regulates Fannie Mae and Freddie Mac.
FHA	Federal Housing Administration, a part of Department of Housing and Urban Development that provides mortgage insurance on loans made by approved lenders.
FFS	Fee for Service, a system of Medicare reimbursement where hospitals were reimbursed based on their costs.
GIS	Geographic Information Systems, a type of software that is used to convert geographic data to maps and estimate variables such as distances.
HCUP	Healthcare Cost and Utilization Project, a project run by the Agency for Healthcare Research and Quality that aims to improve the data available to researchers studying issues of health-care quality and cost.
HSA	Health Service Area, an estimate of a hospital market using patient flow data.
HUD	United States Department of Housing and Urban Development.
MPE	Markov Perfect Equilibrium, a Nash Equilibrium where the agent's decisions are determined by the state in the current period and do not consider the game's history.
MSA	Metropolitan Statistical Area, an estimate of metropolitan areas.
NBER	National Bureau of Economic Research, a national organization of academics based in Cambridge, MA, that conducts economics research.
NIS	National Inpatient Sample, a national sample of inpatients that samples at the hospital level meaning it includes national coverage, but does not include every hospital.
OLS	Ordinary Least Squares, an econometric specification that estimates regression coefficients after making assumptions about the distribution of the unobserved error term.
SID	State Inpatient Database, a state level dataset that includes every inpatient visit to a hospital within that state.

ZIP Zone Improvement Plan, a method to map specific locations to relatively small geographic zones.

Acknowledgements

I would like to thank my committee, Andrew Sweeting (chairman), James Roberts, Frank Sloan, and Patrick Bayer for their valuable and insightful comments throughout the process. I benefited greatly from talking with each of them about my research ideas.

I also appreciate comments from Daniel Xu, Charles Becker, Chris Timmins, and Marjorie McElroy made during presentations, especially for the work in the second chapter. The third chapter is coauthored with Patrick Bayer and James Roberts and I would like to thank them for all of their insight and work on this chapter.

I am grateful to a number of my classmates in the economics department for their feedback and encouragement during the dissertation process including Yair Taylor, Joe Mazur, Tim Schwuchow, and Ryan Brown.

I am indebted to Jean Roth and the National Bureau of Economic Research for all of their help in getting me the hospital and inpatient data used in the second and fourth chapters, as well as to Frank Sloan at Duke for helping to facilitate my access to this data.

Finally, I am incredibly appreciative of all of the support and feedback provided by my wife, Kimberley Geissler. I am solely responsible for any errors.

1

Introduction

My dissertation uses methods developed in the applied microeconomics literature to study markets that are large and complex, and also have potential policy implications and welfare effects. Using this criteria, my dissertation focuses on two industries - hospitals and the real estate market. With hospitals, I analyze the impact of state level regulations and their affect on hospital investment and patient welfare. Additionally, I study how hospital congestion and turnaways impact the composition of patients admitted to a hospital to test whether certain types of patients are disproportionately affected by hospital crowding. My analysis of the housing market looks at real estate flipping to better understand both how house flippers behave and earn their returns, and whether they contributed to the housing bubble in the late 2000s.

The second chapter of my dissertation looks at the effect of a state level regulation, Certificate of Need (CON), on hospital capacity investment. This regulation aims to limit health spending by requiring a hospital to apply to the state health board before making large expenditures, including the addition of hospital beds. There was an active literature in the late 1970s and early 1980s looking at the effect of the regulation on total health spending. Salkever and Bice (1976), Salkever and Bice

(1979), and Sloan and Steinwald (1980) found that total health spending was not affected by the presence of the regulation. More recent studies by Conover and Sloan (1998) and Rivers et al. (2010) have found similar results suggesting that CON has been ineffective in controlling health spending.

There has been less analysis looking at the impact of the regulation on patient welfare and the second chapter aims to address this topic using two datasets. The first is the collection of all inpatient visits to hospitals across four states over a fifteen year period. I observe detailed information about the patient's home zip code, length of stay, and the hospital he visits. The second dataset tracks all U.S. hospitals over time and includes variables for ownership type, location, and capacity. Using this capacity variable, I can track whether a hospital adds or subtracts beds in each year. Merging this data, I estimate the distance that the patient travels to attend the hospital, as well as the distance he would have traveled had he chosen another hospital.

I estimate patient utility using a multinomial nested logit model that includes hospital fixed effects to control for unobserved hospital quality and also allows patients to get more utility from hospitals closer to their home. My empirical model also estimates the probability that a hospital increases or decreases its capacity after controlling for numerous factors including demographic trends, market and hospital characteristics, and the regulatory environment. I find that the regulation decreases the probability that a hospital adds beds by 8 percent.

I run simulations which use the parameters estimated with the empirical model to predict how the hospital industry would evolve differently were regulations in place where they are currently not, and vice versa. This methodology lets hospitals adjust their capacity over a longer time horizon and allows me to estimate the effect of the regulation on patient utility. I find that the hospital industry would have more total capacity without the CON regulations. Additionally, I find that this difference

is most pronounced with smaller and for-profit hospitals which is consistent with the results in Hsia et al. (2012). When estimating the impact on patient welfare, I find that patient utility is decreased by the regulation and to be equally well off, all patients would have to see their distance traveled cut by five percent. These results suggest that the regulation has had a negative impact on consumers and as a result, policymakers should opt against allowing CON in their state.

The third chapter stems from research done jointly with Patrick Bayer and James W. Roberts and focuses on the housing market in the Los Angeles metropolitan area from 1988 to 2009. While there has been discussion in the media and among policymakers about real estate flipping, the literature studying this topic empirically has been very limited. Using a unique dataset that contains every housing transaction during this time period and includes variables for the house's location, the names of the buyer and seller, and the price, I study how real estate flippers behave, and analyze their effect on the housing market as a whole.

I use a repeat sales framework to estimate the time invariant quality of each house where I control for any flipper investment to the house by including transactions before observing the flipper buy the house and again after observing him sell the house. I identify flippers using the buyer and seller names in the data and classify an individual as a flipper if I observe him buying and then reselling multiple houses in less than two years during the sample period. Individuals who flip two or three houses are characterized as inexperienced whereas those who flip four or more houses are classified as experienced.

My results suggest that both of these flipper types earn large returns on their investments, but they do so in different ways. Inexperienced flippers earn their return primarily through market appreciation as they are most active when the housing market in Los Angeles was appreciating rapidly from 1999 to 2005. They tend to buy houses at roughly market value and then resell again near market value. Experienced

flippers are less concerned with timing the market. Instead, they focus on buying homes that are priced below market value. They are more active throughout the sample period since their returns are not tied to overall market appreciation.

I also study whether the flippers impact prices at the neighborhood level in the short or longer term. I find that more experienced flipper activity decreases neighborhood level prices in the short term, but does not affect them over a longer time horizon. Inexperienced flippers increase neighborhood prices in the short term, but decrease them in the long term. This finding is consistent with the argument that inexperienced flippers (who act as speculators) contributed to the housing bubble by driving prices up in the early- to mid-2000s before they fell later in the decade.

While this paper does not offer explicit policy recommendations, it serves as a detailed empirical analysis of how real estate flippers behave thus filling in a major gap in the current literature. Additionally, this analysis provides evidence that inexperienced flippers contributed to the housing bubble with their speculative buying. If policymakers aim to prevent housing bubbles in the future, they may wish to enact policies that will discourage such behavior through tax penalties, limiting access to credit, or other means. Doing so will be challenging, however, because more experienced flippers act as middlemen and provide liquidity. As a result, their presence may improve the market and any policies aimed at discouraging speculative buying should aim to continue to allow these middlemen to buy and sell.

The fourth chapter of my dissertation relates directly to the analysis in my second chapter. I again look at the hospital industry, but in this instance, study the effect of hospital congestion and turnaways on the composition of patients admitted to a hospital. Specifically, I ask whether a hospital's composition changes as it becomes busier and begins turning patients away. And if this composition does change, is it consistent with hospital utility maximization? While there is a literature analyzing the composition of patients across hospital ownership types, and the prevalence of

hospital turnaways, no papers currently study if hospitals turn away some types of patients while admitting others.

I develop a theoretical model of hospital behavior that analyzes how hospitals make patient turnaway decisions to maximize their utility. This model predicts that as a hospital becomes busier, it is more likely to turn away less profitable patients while continuing to accept those that are more profitable. As a result, the composition of patients changes as a hospital becomes more crowded with the fraction of less profitable patients decreasing.

I test my model's prediction empirically using the same datasets as in the second chapter - an inpatient dataset across four states and fifteen years and annual hospital data. The inpatient data includes detailed information about the patient's diagnosis related group, racial status, payment method, and admission type. Using these variables, I generate several measures of patient composition. To test my model, I use a fixed effects regression framework where the dependent variable is the fraction of patients in the given period admitted to the hospital who meet the selected criteria (for example, the fraction who are uninsured). The independent variables include hospital and time fixed effects to control for time and seasonal trends as well as the hospital's mean fraction of that specified patient type. Additionally, I include an estimated turnaway probability generated using the model outlined in Joskow (1980) that serves to estimate the number of patients that a hospital must turn away in that period due to congestion. I am interested in the coefficient on this turnaway parameter as it will indicate if a hospital's patient composition changes as it becomes more crowded.

My results indicate that hospital composition changes as facilities gets busier with less profitable patients seeing their share of admissions decrease. This finding is consistent with that predicted in my theoretical model. Among the types of patients who see their share drop with turnaways include uninsured patients, mental health

patients, and black patients. I also look at whether this behavior is specific to hospitals of a certain ownership type, and do not find any evidence that it is limited based on ownership.

These results suggest that less profitable patients are disproportionately affected by hospital turnaways. While they do not prove that hospitals are discriminatory in who they turn away, my results are consistent with such a story. I hope that these findings will help to inform policymakers about the distributional effects of hospital congestion as it is likely to become a bigger issue in the future with the aging population. Additionally, these findings inform researchers more generally about hospital behavior and utility and how healthcare facilities compare to traditional profit maximizing firms.

The Impact of State Regulations on Hospital Capacity Adjustment Decisions

2.1 Introduction

For decades, policy makers have debated how to best control the increasing costs of health care while maintaining high quality service for patients. During this period, numerous proposals have been discussed and enacted that attempted to keep health spending from growing too rapidly. One key policy aimed at decreasing spending is Certificate of Need (CON), a state level regulation that was first introduced in the 1960s. CON requires hospitals to apply to a state health board before making costly investments and is intended to limit hospital investment to instances where there is a demonstrated need from consumers that is not served by the hospital's local competitors. Its advocates argue that it limits wasteful and unnecessary spending by hospitals and these savings are passed along to consumers (New York State Department of Health (2012)). Adversaries claim that it decreases competition without lowering costs thus hurting patients through higher prices and lower quality (Department of Justice and Federal Trade Commission (2004)).

This paper studies how one type of investment covered under CON, hospital bed increases, impacts hospital capacity decisions. A hospital's ability to adjust capacity is important for two reasons. First, if a hospital is constantly full, it may wish to add beds to better meet patient demand. Second, demand may be endogenous to capacity decisions if patients prefer hospitals that have more beds.

There is considerable debate about CON's effectiveness in limiting spending, as well as its impact on the quality of care patients receive. This is best demonstrated by the fact that approximately half of all states currently have CON laws governing hospital investment. Additionally, health organizations such as the American Hospital Association (AHA) and American Health Planning Association (AHPA) endorse CON whereas the Department of Justice and Federal Trade Commission argue that it hurts consumers by limiting competition and should be uniformly repealed. In the continuing debate over CON's effectiveness as a policy, my analysis helps to illustrate how one key dimension of the regulation impacts patient welfare thus contributing to the debate over whether CON's total net impact is positive or negative.

More generally, this study addresses two classic questions in economics - when do markets perform suboptimally, and how can regulations improve welfare in such instances? As discussed in Mankiw and Whinston (1986), firms may invest more than the socially optimal level when unregulated because they do not consider that investing to increase demand steals business from competitors. This phenomenon may be more pronounced with hospitals than in other industries because the health care market is unique in several ways. First, the set of patients is relatively inelastic meaning that if a hospital receives an additional patient, it is likely stealing him from another hospital as opposed to increasing total market demand. Second, hospital prices do not fall when the hospital has excess capacity. Finally, researchers have argued that hospital utility increases in the number of patients served which suggests that hospitals are not simple profit maximizers and may supply more quality than

is socially optimal.¹

My analysis considers whether state regulations effectively curb investment spending by decreasing capacity without negatively impacting consumer welfare. The utility implications of this question in the health care market are greater than in many other industries studied because health spending comprised 17.9 percent of the United States' GDP in 2010 (World Health Organization (2013)). This indicates that health care represents a large portion of consumer spending and if hospital capacity is suboptimal, consumer welfare could be negatively affected in a significant way.

To address these questions empirically, I use a dynamic model of firm behavior that follows from the methodology outlined in Ericson and Pakes (1995). The model has two types of agents - patients and firms, and uses an infinite horizon with discrete time. Patients choose the hospital that provides the highest utility based on observable firm characteristics, though they may be turned away if the hospital is full. I use a nested logit framework to estimate patient preferences for hospital characteristics as is done in Gowrisankaran et al. (2010). The parameters for patient utility include measures of distance and hospital capacity and the nesting parameter is hospital ownership type and can be used to predict how patient demand for a facility would change based on its capacity decisions as well as those of its competitors.

Hospitals use Markov strategies to determine whether to keep capacity unchanged, or adjust it up or down in the following period. If they choose to change capacity, they must determine the size of this adjustment. These decisions depend on the hospital's current state, where the state space includes all variables that impact a hospital's returns from investment. Hospitals choose the capacity action in each period that maximizes their expected lifetime utility. Demand is estimated using the parameters generated by the nested logit model where I allow for patients to

¹ For more information on hospital utility functions, see Sloan (2000).

be turned away from their first choice hospital. I follow Benkard et al. (2010) to perform counterfactuals without estimating all of the underlying model parameters. This requires estimating the probability that a hospital increases or decreases its capacity conditional on the observed state. For hospitals that adjust capacity, I also estimate the magnitude of this change based on the hospital's state.

I use a simple queuing model for hospital admissions discussed in Joskow (1980) to estimate the probability that a patient is turned away based on the hospital's average daily census and total number of beds. Including this measure allows for a hospital's "crowdedness" to negatively affect patient welfare as some patients will be turned away from their first choice facility and instead will be admitted to their second choice where they receive less utility.

The estimated patient utility parameters indicate that patients prefer hospitals that are closer to their home, suggesting that travel is costly. Additionally, patients receive more utility when facilities are larger meaning that hospital size affects patient welfare. The hospital capacity decision parameters estimate that the presence of CON decreases the probability that a hospital adds beds by 8 percentage points. This finding indicates that CON either leads to proposed capacity increases being rejected by the state health board, or deters hospitals from applying to add beds.

I use the parameters for patient utility and hospital capacity decisions to simulate how the hospital industry would evolve differently over time if CON was implemented in states where it does not exist, or was discontinued where it is currently in place. The simulations assume that each hospital starts with its observed capacity in the first year of data. After the first year, I predict how hospital capacity would adjust in the following year conditional on the state variable in the current year under two scenarios. The first assumes that CON regulations are present and the second assumes they are not. The simulations' estimates illustrate how hospital capacity would evolve under each regulatory policy over a long time horizon.

The simulations indicate that total hospital capacity is smaller when CON regulations are present over a 25 year period. Their impact varies depending on hospital characteristics, however, as they predict that smaller and for-profit hospitals are more likely to see their capacity decreased under CON. This finding is consistent with qualitative evidence which suggests that these hospital types are disproportionately affected by the regulations (Yee et al. (2011)). While CON's effect on total capacity is modest, its impact on turnaway probabilities is more significant. The model predicts that turnaway probabilities increase on average from 1.5 percent to 2.2 percent when CON is present indicating that the regulations lead to fewer patients being admitted to their preferred hospital.

The simulations allow for the comparison of distances traveled and total patient utility across the regulatory policies over time. The presence of CON decreases total patient utility by 1.2 percent over the simulation period with the difference increasing over time. To make patient welfare equal in each of the two regulatory environments, patient distance traveled would have to decrease by 5.1 percent when CON regulations are present. These results suggest that CON regulations have a real impact on hospital capacity decisions, which ultimately hurts patient welfare.

These findings contribute to the debate about the merits of CON and whether it should be implemented or discontinued at the national level. While only analyzing a part of CON's overall impact on welfare, my findings indicate that the regulations decrease patient welfare. These results will be informative to state and federal policy makers in understanding how regulations in the hospital industry affect both patients and hospital capacity decisions.

The rest of the paper is organized as follows. Section 4.2 provides background information on CON. Section 4.3 discusses the data. Section 4.4 outlines the empirical model used to estimate patient utility and hospital capacity decisions. Section 4.5 estimates the results from the empirical model. Section 4.6 uses simulations to

analyze the impact of CON over a longer time horizon. Section 4.6 concludes.

2.2 Background on CON

2.2.1 History of CON

Real per-capita health spending in the United States increased by 44 percent between 1950 and 1960, and by 76 percent from 1960 to 1970 (Newhouse (1995)). As a result of this growth, policy makers sought to slow health spending. The first related policy related to CON was enacted in Rochester, New York in 1964 and required evaluation of community need prior to the approval of large health expenditures. In 1966, New York passed the first state CON law requiring hospitals to receive state approval before making large capital investments. Several states followed New York's lead and passed similar CON laws. The Federal government believed CON was limiting hospital expenditures and required every state to adopt CON by 1980 to be eligible for full Federal Medicaid reimbursements (Sagness (2007)). All 50 states eventually adopted the regulation.

CON supporters argued for the regulation for two reasons. First, hospitals were fully compensated for costs by insurers through the fee for service (FFS) payment system, giving hospitals little incentive to control costs (Conover and Sloan (1998)). Second, as most patients had health insurance, they were responsible for a small fraction of their total costs. As a result, patients made decisions regarding hospital choice based on quality of service rather than cost (Sagness (2007)).

In the mid-1980s, the Federal government stopped requiring states to have CON to receive Federal reimbursements. Early research on the regulation suggested that its impact on total health expenditures was minimal. Health spending was still growing rapidly and studies by Salkever and Bice (1976) and Salkever and Bice (1979) found CON had little to no impact on overall health spending. Sloan and Steinwald (1980) concluded that under CON regulations, hospitals shifted spending

from capital investments to labor, which was not regulated by CON policies.

The hospital reimbursement structure changed in the mid-1980s and this also may have impacted the government's decision to stop requiring that states have CON. Hospitals no longer received reimbursements for services provided based on costs and instead received a fixed payment based on the patient's diagnosis related group (DRG). This reduced incentives for hospitals to perform expensive procedures if they had little benefit to patients, as the hospital would not be fully reimbursed for the costs of these procedures. This gave hospitals a powerful incentive to provide more efficient care (Sagness (2007)).

Although the Federal government discontinued the CON policy requirements, a majority of states kept these regulations in place. Several states also chose to end their CON programs in the 1990s. At present, 36 states have CON regulations in place, although in some cases their regulations no longer govern hospitals and instead focus primarily on nursing homes.²

2.2.2 CON Justification and Procedures

The New York policy states, "The objectives of the CON process are to promote delivery of high quality health care and ensure that services are aligned with community need. CON provides the Department of Health oversight in limiting investment in duplicate beds and medical equipment which, in turn, limits associated health care costs." (New York State Department of Health (2012)). This is similar to the language used to justify CON regulation in many states. CON programs are designed to limit investment to proposals that improve the quality of care in a community. In theory, an application from hospital A for an expensive piece of equipment (e.g. catheterization lab) should be rejected if hospital B in the same market already has

² For more information on which states have CON programs and what is covered under each, see National Conference of State Legislatures (2013).

the equipment and there is not sufficient demand in the market for two sets of the equipment.

CON regulation applies to a wide range of hospital expenditures, with the specific policies varying by state. CON regulated hospital expenditures typically include expensive technologies that may not be necessary at each hospital in a market, such as a catheterization laboratory. Additionally, there is a total spending limit above which CON approval is required even if the investment is for a technology not specifically regulated under CON. This threshold varies by state but is typically in the hundreds of thousands or millions. For example, the threshold value requiring CON application in Washington is \$2.4 million (Washington State Department of Health (2013)).

If a hospital wants to increase its capacity by adding beds to its facility, CON approval is required. CON supporters worry that if hospital investment is unregulated, hospitals may expand capacity beyond the socially optimal level. Excess capacity is costly, and these expenses may be passed along to consumers through higher prices. While the rules and enforcement surrounding CON regulation of capacity increases vary by state, a common target used is 80 percent at either the hospital or market level.³

The CON application process varies by state, but is typically similar to the procedure outlined in Figure 2.1. When a hospital decides to make an investment regulated by CON, it must file an application with the state governing board. The application process requires the hospital to demonstrate a community or market need for the service provided by the investment. The need criteria vary by state and may not be transparent to hospitals (Yee et al. (2011)).

For a set period after an application is filed, competitors may contest the application by demonstrating that the investment duplicates services already available to

³ This value was recommended by the National Guidelines for Health Planning (Finch and Christianson (1981)).

Typical Certificate-of-Need (CON) Application Process

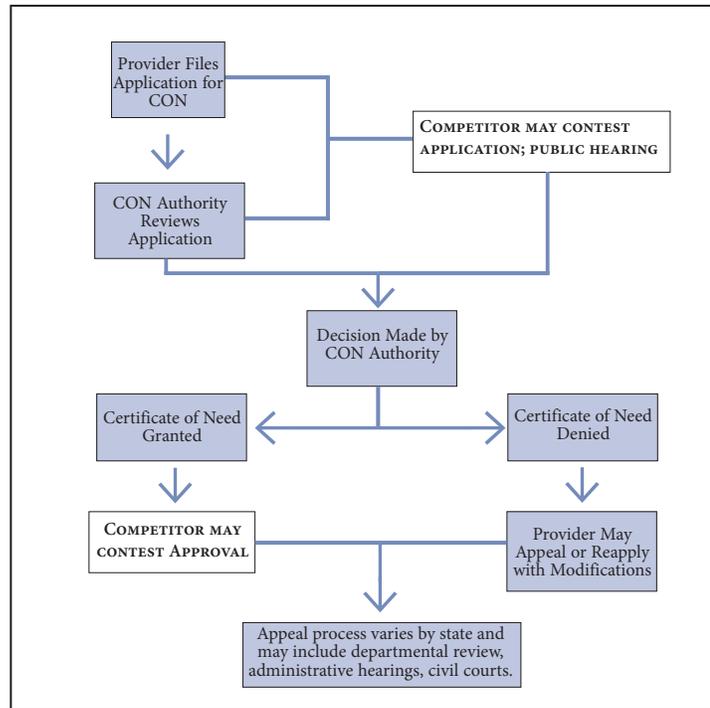


FIGURE 2.1: CON application process. Source: Yee et al. (2011).

patients in the community. Public hearings are also held where local citizens can voice their support or concerns with the proposed expenditure. If approved, competitors have a second change to appeal; if rejected, the applying hospital can appeal the decision or reapply after modifying its application. According to qualitative research in Yee et al. (2011) there may be politics involved in the approval process with wealthier and larger hospitals more likely to receive approval.

The presence of CON regulations is likely to impact the probability that a hospital changes its capacity in several ways. Most importantly, there is the possibility a hospital's application will be rejected, thus preventing the facility from adjusting capacity, even if it wishes to. Rejections are not common, but a recent study of six states shows that between 6 and 12 percent of applications were rejected in 2009 (Yee et al. (2011)). These percentages may not represent the full impact of CON, however, as a hospital that would have invested if no regulations were in place may choose not

to apply if it believes there is a high chance of its application being rejected.

There are also significant time and financial costs associated with CON. For example, the regulation imposes a time cost on applicants as they must wait to receive approval before making proposed changes. Even without the potentially length appeals process, the application process lasts several months. The CON process typically allows two rounds of appeals from competitors. Yee et al. (2011) finds it is common for rivals to appeal an application that is likely to be approved to delay its implementation, in some cases by several years.

A hospital must pay a fee to submit an application that varies by state and may also depend on the estimated cost of the proposed expenditure. For example, hospitals pay a fixed fee of \$40,470 in Washington whereas in New York, they pay a fixed cost of \$2,000 and then 0.3 percent of estimated the costs of the investment. These application fees increase the costs of adjusting capacity to a hospital. Larger hospitals may have a team of experts responsible for CON regulatory issues as substantial financial and legal resources are necessary to file a CON application and monitor the appeals process. Many smaller hospitals are unable to afford such a team as it is costly to maintain (Yee et al. (2011)). The financial costs associated with CON regulation may disproportionately affect smaller and for-profit hospitals that are less likely to have the resources necessary to file and navigate the application process.

2.2.3 Debate Over CON's Impact

Due to certain characteristics of the health industry, government regulation may be useful. In particular, consumers do not have full price information when making consumption decisions. This is due to the fact that patients may not receive information on hospital prices or charges until well after the episode of care. The services incurred, particularly in an inpatient setting, may not be known in advance meaning the patient will not know the cost of service even if he is aware of hospital prices.

Additionally, consumers may not be very price sensitive due to health insurance coverage which covers a large fraction of their expenses.

If the health industry operated as a more traditional market where consumers paid for services received in full, with complete information about prices and qualities, government regulation of the health industry could be harmful; regulation might prevent hospitals from setting prices or qualities optimally. If hospitals compete primarily on quality, the level provided may be above the socially optimal quality because patients do not face the full cost and therefore prefer a higher quality level than if they paid the full cost for the services provided.

CON's effectiveness can be measured on two primary dimensions: health care costs and quality of care. The original intent of policy makers in enacting CON was to lower costs without sacrificing quality, but early research indicated that it was ineffective in this capacity. A number of more recent studies have examined the effects of CON on health spending and concluded that CON has little to no impact on overall health spending (Conover and Sloan (1998); Rivers et al. (2010)). In fact, Lanning et al. (1991) found CON increased total hospital spending.

An additional body of literature focuses on how CON regulation impacts quality of care. Burda (1991) finds that CON decreases competition between hospitals and has negative effects on quality. This is consistent with the Department of Justice and Federal Trade Commission (2004) which concludes that CON enacts barriers to competition that may increase prices for consumers. Kessler and McClellan (2000) finds greater hospital competition in the 1990s led to better patient outcomes, suggesting that if CON restricts competition, quality is likely to suffer.

There is limited research studying CON's effects on hospital occupancy and capacity decisions. Joskow (1980) focuses on the impact of CON on hospital quality using a measure of the hospital's turnaway probability. The study considers this measure a proxy for quality of care, with higher quality corresponding to a lower

turnaway probability.⁴ Joskow uses hospital data from 1976 to analyze the impact of CON on turnaway probability and finds CON positively influences this value, thus decreasing his measure of hospital quality.

While limiting competition and capacity may decrease quality, some studies find that CON increases quality for specific procedures. CON has been shown to improve outcomes for two heart procedures requiring specialized equipment regulated by CON (Vaughn-Sarrazin et al. (2002) and Ross et al. (2007)).⁵ This finding may be explained by the volume-quality relationship found in many procedures (Peterson et al. (2004)). As fewer hospitals conduct these procedures in states without CON, the hospitals with the necessary equipment perform the procedure more frequently. If the quality of the procedures improves with volume, this may lead to better outcomes in states where CON is present. Additionally, states may only approve higher quality hospitals for the investments which may also improve the quality level of procedures performed. While these studies suggest that outcomes for selected interventions are better in states with CON, they are less informative in evaluating the effects of CON regulation on patient access.

Despite suggestions in the academic literature that CON's effect on spending is minimal, with mixed implications for quality, the regulation is still present in a majority of states. Organizations representing hospital interests, including the American Hospital Association (AHA) and American Health Planning Association (AHPA), support CON regulation. My analysis will therefore be useful in helping

⁴ While turnaways are rarely observed in hospital data, they do occur. This is evidenced by ambulance diversions where a hospital alerts local ambulances that their emergency department is overcrowded and patients should be transported elsewhere. This phenomenon is consistent with findings in Bagust et al. (1999) which uses simulations and finds that hospitals with an average occupancy of 85 percent begin to have problems with overcrowding, and as the average occupancy climbs to 90 percent, this issue becomes significantly worse.

⁵ The two procedures, percutaneous coronary intervention (PCI) and coronary artery bypass graft (CABG) both use a catheterization lab. These labs costs several million dollars to construct and require CON approval in many states.

to estimate CON's impact on hospital capacity decisions and patient welfare. These results will be informative to policy makers considering whether to implement or discontinue the regulations.

2.3 Data

I use two datasets for my analysis that provide information at the patient and hospital levels. Patient data are from State Inpatient Databases (SID) for selected states. The American Hospital Association (AHA) annual survey provides data on hospital characteristics. I discuss important characteristics of each dataset and the construction of the final dataset below.

2.3.1 State Inpatient Data

I use SIDs from Arizona, New Jersey, New York, and Washington for the period of 1995 to 2009.⁶ For the duration of my data, Arizona does not have CON regulations in place. The other three states have CON regulations governing major investment decisions, which include hospital capacity. The data are distributed by the Healthcare Cost and Utilization Project (HCUP). I use this subset of states and years because they include detailed information on the patient residence and were available to me through the National Bureau of Economic Research (NBER).

The data include observations for each inpatient visit to a hospital in the state excluding those that are Federally owned.⁷ The SID data contain the full universe of hospitals and patients within these hospitals, thus allowing me to fully characterize hospital and patient behavior in each state. I use inpatient visits for all diagnoses and admissions sources in my analysis.

⁶ The data for New Jersey start in 1997 and for Washington, they begin in 1999.

⁷ Federally owned hospitals are often run by the military or Department of Veteran Affairs and serve specific population subgroups. These hospitals are a small proportion of the total patient population for acute care hospitals.

The SIDs include detailed information about the patient admission including the length of stay, whether the visit was elective or an emergency, the hospital where the patient was admitted, patient ZIP code, payment method, and diagnosis and procedure codes. The data do not include outpatient visits to hospitals or other clinics, or emergency department visits not resulting in admission to the hospital.

The price paid by the patient is not available in the data. While it includes hospital charges, these values are not an accurate reflection of the price facing the patient. Insurance companies negotiate the fraction of charges to be paid with the hospital, and this value varies across insurance providers, hospitals, and procedures. Charges are often reduced substantially for patients without insurance; these patients may qualify for charity care or arrange payment plans for a negotiated percentage of the final bill. I discuss my strategy related to the lack of price data in more detail in Section 4.5.⁸

2.3.2 Hospital Data

The AHA survey data uses the hospital-year as the unit of observation and includes detailed information about each hospital including information on its location (ZIP code), ownership type, and number of beds.⁹ A hospital identifier allows me to track individual hospitals over time and link the hospital data to the SIDs.

2.3.3 Combined Dataset

I merge the two datasets using a hospital identifier variable included in each dataset. I then use geographic information systems (GIS) software to estimate the distance from the centroid of each patient's ZIP code to the centroid of the hospital's ZIP code. This provides a measure of the distance each patient traveled to receive care.

⁸ For more information on hospital charges and how they relate to prices, see Lane et al. (2001) and Dobson et al. (2005).

⁹ The beds variable corresponds to the number of beds in the facility. During periods of low demand, some of these beds may not be staffed but such changes are not observed in the data.

I reshape the data so that the unit of observation becomes the patient ZIP-hospital-year combination as is done in Berry (1994). I calculate the share of patients from each ZIP code admitted to each hospital. For my analysis, I consider any hospital that is further than 100 kilometers away, or receiving less than 5 percent of all patients from a ZIP code to be part of the outside option.¹⁰

Summary statistics are provided in Table 4.1. The distance traveled is generally similar across states, with patients in New Jersey traveling a shorter distance. New Jersey is more densely populated than the other states in the sample, making this an expected finding. The observed share is the fraction of patients from the ZIP code who select that hospital. Average observed shares are similar across the four states with their values being slightly higher on average in Washington. Hospitals in New Jersey and New York have more beds than those in Arizona and Washington.

Table 2.1: Summary statistics: Each observation in N represents a unique ZIP-hospital-year combination. Values in parentheses represent standard deviations. Total beds is calculated at the hospital-year level.

	All States	AZ	NJ	NY	WA
Distance (km)	23.5 (21.6)	25.5 (25.3)	14.9 (12.8)	24.2 (21.3)	28.0 (24.3)
Observed Share	0.235 (0.199)	0.224 (0.202)	0.237 (0.237)	0.229 (0.191)	0.273 (0.220)
Total Beds	272.1 (253.6)	190.7 (154.2)	312.6 (207.3)	325.9 (292.9)	131.0 (129.9)
Unique Hospitals	478	75	86	226	91
Unique Patient ZIP Codes	3,875	523	661	1995	696
N	145,876	20,683	21,653	84,863	18,677

2.3.4 Is State Level Data Representative?

I analyze AHA data on capacity changes to determine whether trends observed in the four states in the combined patient-hospital dataset are consistent with hospitals

¹⁰ Gowrisankaran et al. (2010) also use a distance measure to determine the outside option in their hospital patient utility framework. I check the sensitivity of this assumption to different distance and percentage cutoffs for the outside option and find similar results.

nationally. I estimate two probit regressions with the dependent variable equaling an indicator for whether the hospital adjusts capacity up (specifications (1) and (2)) or down (specifications (3) and (4)). I control for hospital and market characteristics available in the AHA data. Additionally, I include a dummy variable for whether the hospital is located in a state with CON regulations. I include two sets of regressions - the first use all hospitals and the second are restricted to the sample of states used in my primary analysis. If the coefficients for the CON variable are similar between all hospitals nationally and the sample, this result will suggest that the regulations' impact on hospital capacity can be generalized from my sample to all states.

Table 2.2: National versus sample regressions: The unit of observation is the hospital-year. Coefficients are marginal effects. Standard errors are in parentheses. Controls are included for changes in patient days at the state level, hospital occupancy level, hospital ownership type, and year.

	(1)	(2)	(3)	(4)
	Bed Increase		Bed Decrease	
	National	Sample	National	Sample
CON	-0.013 (0.004)	-0.014 (0.025)	-0.002 (0.004)	0.027 (0.026)
N	45,421	4,533	45,421	4,533

As demonstrated in Table 2.2, the impact of CON regulation on the probability that a hospital increases capacity nationally is similar in magnitude and direction to the effect in my sample. CON has a slightly larger impact on bed decreases in my sample, but neither estimate is significantly different from zero. As CON's impact on capacity decisions is similar in my sample to national estimates, my results likely generalize to all 50 states.

2.4 Empirical Model

My empirical model features a set of hospitals competing over an infinite horizon in discrete time and follows from the framework outlined in Ericson and Pakes (1995).

In each period t , hospitals make capacity decisions for $t + 1$. If a hospital elects to increase or decrease capacity, it must also determine the magnitude of this change. I introduce a state space that includes any variable which affects a hospital's returns from investment. Hospitals make capacity decisions that maximize their expected lifetime utility.

In each period, an exogenous set of patients require hospitalization and select the hospital that provides the highest utility. Each patient's static hospital decision is based only on hospital characteristics and an unobserved error term. In some instances, the patient's first choice hospital may be full in which case he attends the hospital that provides the second highest utility.

In the remainder of this section, I first outline the estimation of a hospital's turnaway probability that determines whether the patient is admitted to his first choice hospital. I then discuss a model of patient utility and how its parameters are estimated after allowing for patients to be turned away from their first choice hospital. Finally, I discuss the dynamic model that characterizes the hospital capacity decision and solving for the parameters that impact hospital decisions.

2.4.1 Turnaway Probability

Joskow (1980) estimates turnaway probability using a simple queuing model introduced in Shonick (1970) and used by state planning agencies studying hospital capacity (Joskow (1980)). To estimate the turnaway probability, he assumes that patients arrive to the hospital with a Poisson distribution of λ and their length of stay has a negative exponential distribution with a mean of $1/\mu$. If a hospital has b beds, then the probability that j of them are full at a specific time is given in equation (2.1).

$$P_j = \frac{(\lambda/\mu)^j/j!}{\sum_{h=0}^b (\lambda/\mu)^h/h!} \quad (2.1)$$

The distribution of P_j is not Poisson due to the truncation at b . However, as b becomes large relative to the mean number of patients, Poisson becomes a good approximation. In this distribution, the mean and variance are equal and when they become large, the Poisson can be well approximated by a normal distribution. Using these assumptions, the probability that a hospital is full can be rewritten in the following form where Φ represents a normal cumulative density function. The hospital's average daily census, ADC, is calculated using the patient arrival rate and average length of stay.

$$P_f = 1 - \Phi\left(\frac{b - \text{ADC}}{\sqrt{\text{ADC}}}\right) \quad (2.2)$$

While Joskow estimates the turnaway probability as a proxy for hospital quality, I use this variable for what it explicitly measures - the probability that a patient is turned away from the hospital. In the literature, patient utility parameters are typically estimated based on the hospital to which the patient is admitted. Incorporating an estimate of the turnaway probability allows these parameters to account for the possibility that some patients may not end up at their first choice hospital.¹¹

2.4.2 Patient Utility

Patient utility follows the nested multinomial logit form discussed in Berry (1994). This functional form uses patients' hospital choices to estimate patient utility parameters and the nest is the hospital ownership type. This is the same functional form used to measure patient utility parameters as in Gowrisankaran et al. (2010).¹²

¹¹ I estimate that a small number of hospitals have large turnaway probabilities. Throughout my analysis, I cap the estimated turnaway probability at 25 percent so that no hospital turns away more than a quarter of its patients. When I used other turnaway caps, the results do not significantly change.

¹² In Gowrisankaran et al. (2010), the authors include an unobserved time-varying hospital quality term, ξ_{jt} . But they assume that hospital shares are observed without measurement error which allows them to treat ξ_{jt} as the error term. This simplifies estimation by reducing the equation to a nested multinomial logit model. I make the same assumption in my estimation.

Patient i receives the following utility from visiting hospital j in period t .

$$u_{ijt} = \alpha_1 X_{ijt} + \alpha_2 Y_{jt} + \kappa_j + \epsilon_{ijt} \quad (2.3)$$

X_{ijt} is a vector of variables that are specific to the patient and hospital such as the distance between patient i 's home ZIP code and hospital j . Y_{jt} includes variables that impact the hospital's quality and vary over time such as the hospital's total capacity. κ_j is a time-invariant hospital fixed effect which captures the baseline quality of hospital j . Finally, ϵ_{ijt} is an error term containing two components as given below.

$$\epsilon_{ijt} = (1 - \sigma)\epsilon_{1ijt} + \epsilon_{2igt} \quad (2.4)$$

ϵ_{1ijt} is distributed extreme value and is i.i.d. ϵ_{2igt} is drawn from the distributional form that allows ϵ_{ijt} to be generated from an extreme value distribution.¹³ While patient i gets a separate draw of ϵ_{1ijt} for each hospital j , he only receives a draw ϵ_{2igt} for each hospital ownership type. This allows patients to have their unobserved preferences for hospitals be correlated by hospital ownership type. σ is assumed to be between zero and one and represents this correlation. A higher value of σ indicates that preferences between hospital ownership types plays a larger role in determining patient utility. When σ is equal to 0, equation (2.3) reduces to a standard multinomial logit as the nest-specific error term drops out of the equation. When σ is 1, the only error term for each patient occurs at the nest level.

Each patient chooses the hospital that provides the highest utility. If no hospital provides positive utility, the patient chooses the outside option and receives zero utility. The outside option still requires that each patient attends a hospital meaning I assume no patients skip or postpone medical treatment due to unsatisfactory

¹³ For more information on this distribution, see Cardell (1997).

hospital choices. Berry (1994) shows that the multinomial nested logit equation can be transformed into the following form.

$$\ln(s_{zjt}) - \ln(s_{z0t}) = \alpha_1 X_{ijt} + \alpha_2 Y_{jt} + \alpha_j D_j + \sigma \ln(\bar{s}_{j/g}) + \xi_{zjt} \quad (2.5)$$

This transformation converts the unit of observation from the patient to the ZIP-hospital-year level and while algebraically equivalent, it is less computationally intensive to estimate. s_{zjt} represents the fraction of patients in ZIP code z at period t that choose hospital j and s_{z0t} is equal to the fraction that choose the outside option. $\bar{s}_{j/g}$ is the share of patients from ZIP code z in period t that choose j , conditional on selecting a hospital with the same ownership type as j . $\bar{s}_{j/g}$ will likely be endogenous and I instrument using the distance to the closest hospital of j 's ownership type that is not j .

I calculate the patient shares which make up the dependent variable in equation (2.5) in two ways. The first method assumes that patients are never turned away and attend the hospital that provides the highest utility. To estimate the parameters under this assumption, I use the fraction of patients from ZIP code z who choose hospital j in period t .

The assumption that all patients attend their first choice hospital is problematic if hospitals turn away patients when they are full. If turnaways occur, the share of patients visiting each hospital does not represent patient preferences as some individuals may not be able to attend their preferred facility. If the data included information on patients' first choice hospitals (in addition to the hospital attended), the estimation would again follow equation (2.5) using the first choice hospital as the dependent variable.

I use the following identity to calculate the share of patients who consider a

hospital their first choice, s_f .¹⁴ s_o represents the observed share that visit the hospital in the data, s_t is the share that want to visit the hospital but are turned away, and s_s is the share that are admitted to the hospital after being turned away from their first choice hospital.

$$s_f = s_o + s_t - s_s \tag{2.6}$$

First choice shares are equivalent to the sum of the observed share in the data and the share who would have chosen the hospital had they not been turned away, less the share of patients who chose the hospital because they were turned away from their first choice. The only variable in identification (2.6) that is observed in the data is s_o meaning both s_t and s_s must be estimated to predict s_f .

To estimate the number of patients who are turned away from the hospital, s_t , I use equation (2.2) which generates a measure of the turnaway probability. The share of patients who choose a hospital because their first choice was full also depends on these turnaway probabilities. To predict s_s , I estimate substitution patterns between hospitals. More detail on how these values are calculated is provided in Appendix A.

Accurately predicting each hospital's turnaway probability and substitution patterns requires that the coefficients estimated from equation (2.5) are correct. However, to estimate these coefficients I need unbiased estimates of the turnaway probability and substitution patterns. To get unbiased estimates of both of these sets of parameters, and therefore of first choice shares, I use the iterative process outlined below. I begin by setting s_f equal to s_o .

(1) Given the current estimates of s_f , I calculate the values of $\ln(s_{zjt})$, $\ln(s_{zot})$, and $\ln(\bar{s}_{j/g})$, and use these values to estimate the coefficients in equation (2.5) via

¹⁴ My notation drops hospital and time subscripts for simplicity and assumes just one hospital.

instrumental variables regression.

(2) Conditional on the parameters estimated in (1), I estimate the turnaway probabilities and substitution patterns for each hospital. Appendix A provides more detail on these calculations.

(3) Using the turnaway probabilities and substitution patterns estimated in (2), I calculate new values of s_t and s_s and use these values and s_o to generate an updated estimate of s_f .

(4) I return to (1) until s_f converges.

The patient utility parameters will be informative in two ways. First, they will describe what characteristics patients prefer in hospitals including the relative importance between distance to the hospital and capacity. Second, variables estimated using the patient utility parameters such as turnaway probability and the hospital's "elasticity"¹⁵ will be used to estimate the probability that a hospital adds or subtracts capacity.

2.4.3 Hospital Capacity Decisions

I use an infinite horizon discrete time model where hospitals make capacity decisions in each period to maximize their expected lifetime utility. The model follows from the framework outlined in Ericson and Pakes (1995) and assumes that all hospitals simultaneously decide in period t whether to adjust capacity for $t + 1$. Hospital h 's decision in t affects its expected lifetime utility in two ways. First, this decision directly impacts utility in $t + 1$ as it determines h 's capacity and any associated adjustment costs in this period. Second, hospital h 's action in t affects the competitive environment from which it makes its capacity decision in $t + 1$, which may impact its capacity choice in $t + 1$ and affect its expected lifetime utility.

¹⁵ The "elasticity" is equal to the estimated change in the number of patients if the hospital adds one bed. A higher value therefore suggests that the number of patients is more sensitive to capacity changes.

I characterize this competitive environment with a state vector, ω_{ht} , which includes hospital and market variables that influence the hospital's returns from capacity. This vector includes key hospital variables such as its capacity in period t , ownership type, and turnaway probability. Additionally, it features market level variables such as the metropolitan statistical area (MSA) size and the number of hospitals in the health service area (HSA).¹⁶

The model assumes that hospital h receives an i.i.d. shock vector, γ_{ht} , in each period that is generated from a function $F(\cdot)$ and is not observed by its competitors or the econometrician. The vector represents unobserved factors that may impact the hospital capacity decision such as shocks to hospital adjustment costs. In each period, the hospital will choose a capacity action, a_{ht} , which characterizes whether the hospital increases or decreases its capacity, and the magnitude of any change.

Estimating hospital utility or adjustment cost parameters would require specifying a functional form.¹⁷ However, my estimation focuses on estimating the hospital capacity actions, which are observed in the data. As a result, I do not use a specific functional form for hospital utility and keep the equation general. Hospital h 's utility in period t is assumed to be a function of its action in this period, the current state, and its shock vector.

$$U_{ht} = U(a_{ht}, \omega_{ht}, \gamma_{ht}) \quad (2.7)$$

I define a Markov strategy for hospital h as a function, σ_h , that maps each potential state and unobserved shock into an action. I can replace the firm's action, a_{ht} , with its strategy function, $\sigma(\omega_{ht}, \gamma_{ht})$, in equation (2.7) to describe utility as a function of a hospital's strategy, state, and shock vector. A key feature of a Markov

¹⁶ HSAs were developed by the National Center for Health Statistics to estimate geographic markets for hospital care (National Cancer Center (2013)).

¹⁷ For example, in Ryan (2011), he specifies a functional form for firm adjustment costs and estimates the parameters for this form in the second stage of his estimation procedure.

strategy is that conditional on the state and shock vector, hospital h will choose the same action regardless of the period, thus allowing the time subscript to be dropped.

I construct a value function for hospital h using equation (2.7) which gives h 's expected lifetime utility conditional on its current state and Markov strategy. The value function assumes that all hospitals have a common discount factor β , and conditional on the current state and hospital h 's action in t , the state variable for h in $t+1$ is generated from a probabilistic function P . These assumptions are necessary to ensure the hospital capacity action and transition between states depend only on the current state variable and are not influenced by hospital actions or states in previous periods.

$$V(\omega_h, \sigma_h, \gamma_h) = U(\sigma_h(\omega_h, \gamma_h), \omega_h, \gamma_h) + \beta \int \int V(\omega'_h, \sigma_h, \gamma'_h) dP(s' | \sigma_h(\omega_h, \gamma_h), \omega_h) dF \quad (2.8)$$

The value function consists of two parts. The first is equal to the utility that the hospital receives in the current period based on its strategy function, the state vector, and the shock vector. This is simply the per-period hospital utility function defined in equation (2.7). The second part of the equation represents the expected lifetime utility that h receives beginning in the following period, conditional on using its strategy function, σ_h , in the current period. Because the state and shock vectors in the following period are unknown, I integrate over each of their distributions to generate an expected value of the value function beginning in the following period.

Let σ be a set of Markov strategies for all hospitals, $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_H)$. σ represents a Markov Perfect Equilibrium (MPE) if no hospital h has an incentive to deviate from its strategy function, σ_h , assuming all other hospitals play their equilibrium strategies. A set of strategies, σ , is therefore an MPE if for all states, ω , in the state space, and all hospitals, $h \in [1, \dots, H]$, the following condition holds.

$$E(V(\omega_h, \sigma_h)) \geq E(V(\omega_h, \sigma'_h)) \quad \forall h \in [1, \dots, H], \forall \omega_h \in \Omega \quad (2.9)$$

Following the methodology in Benkard et al. (2010), I make three assumptions about the MPEs for hospitals that are necessary to solve the model and estimate its parameters. First, I assume that at least one pure strategy MPE exists. Second, all hospitals play the same MPE, even if it is not unique. Third, we can think of firms as making decisions under two regulatory environments - one where CON regulations are present and another where they are not. I assume that the state vector captures how returns to capacity vary in each environment and hospitals play the same Markov strategy in each case. This assumption is consistent with Benkard et al. (2010) which assumes that airline merger strategies are not impacted when other mergers are approved or rejected.

Using the value function in equation (4.1), I can solve for the strategy function, σ , which maps states and γ 's into hospital actions. This estimation is difficult for several reasons. First, one must specify functional forms for hospital utility and capacity adjustment costs. If the form of either of these equation is misspecified, the estimates will be incorrect. Secondly, such estimation requires distributional assumptions about the unobserved vector, γ and requires that this shock value is one dimensional. Finally, estimating these parameters requires multidimensional integration, which is computationally very challenging for my analysis.

Instead, I follow the estimation strategy outlined in Benkard et al. (2010) which estimates the probability of each hospital's capacity action directly from the data, conditional on the current state. The primary advantage of this approach is that it does not require any functional form assumptions about hospital utility or capacity adjustment costs. Additionally, this method is less computationally demanding because the probabilities can be estimated without integration. The primary drawback

to this methodology is that it does not generate estimates of second stage parameters such as hospital utility and adjustment costs that would be included in the Bajari et al. (2007) (BBL) framework. These parameters are not needed to predict hospital capacity decisions or patient utility, however, and are therefore not necessary for my analysis.¹⁸

I estimate the hospital choice probabilities using a simple two-step framework. The first stage uses a probit model where the dependent variable is an indicator for whether the hospital increased capacity between t and $t + 1$ and the independent variables are the set included in the state vector at time t , ω_{ht} . These estimates generate predicted probabilities that the hospital increases capacity conditional on the state. Second, I estimate the amount by which the hospital changes capacity, conditional on adding beds. In this OLS regression, I restrict the sample to hospitals that add capacity between t and $t + 1$ and use the same set of independent variables; the dependent variable is the magnitude of the capacity change. For both the probit and OLS regressions, I include a dummy variable for whether CON is present among the independent variables to control for the regulatory environment. After estimating these parameters for capacity increases, I repeat this process with hospital capacity decreases.

2.5 Results

Based on my empirical model, I present two sets of results: parameters associated with patient utility and those corresponding to hospital capacity adjustments.

¹⁸ However, if I wanted to expand the analysis to estimate how total hospital spending was affected by the presence of CON, including such parameters would be necessary.

2.5.1 Patient Utility

Table 2.3 uses equation (2.5) to estimate the patient utility that using two specifications. The first specification, labeled “No Turnaways” assumes that each patient is admitted to his first choice hospital. The second specification labeled “Turnaways” allows a patient to be turned away from his first choice hospital if it is full. When this occurs, he chooses his second choice hospital. The calculation of these turnaway probabilities and substitution patterns is outlined in more detail in Section 2.4.2. The primary difference between these regressions is that I now use estimated first choice shares rather than observed first choice shares. Both regressions include hospital fixed effects to control for unobserved time-invariant hospital quality as well as other hospital level variables such as ownership type.

Table 2.3: Patient utility regressions: Standard errors are clustered by hospital and in parentheses. Regressions include hospital and year fixed effects. N corresponds to ZIP-hospital-year combinations. σ is the correlation of the nesting parameter.

	No Turnaways		Turnaways	
Distance (km.)	-0.025	(0.003)	-0.023	(0.002)
Distance Squared	0.0002	(0.0001)	0.0001	(0.00004)
Distance Cubed	-3.85e-07	(6.03e-07)	1.31e-07	(2.96e-07)
Closest Hospital	0.242	(0.039)	0.307	(0.021)
Distance*Gov't	-0.003	(0.004)	-0.003	(0.001)
Distance*Nonprofit	-0.005	(0.002)	-0.005	(0.001)
Log(Beds)	-0.047	(0.036)	-0.054	(0.023)
Log(Beds)*Gov't	0.064	(0.031)	0.063	(0.010)
Log(Beds)*Nonprofit	0.138	(0.022)	0.125	(0.011)
σ	0.515	(0.052)	0.446	(0.034)
N	145,876		145,876	

The results in the two specifications are similar suggesting that patient turnaways do not substantially impact the patient utility parameters. This is not surprising given that I estimate that the mean hospital turnaway probability in the data is 1.2 percent suggesting that a large majority of patients are accepted at their first choice hospital. While the distance coefficients are difficult to interpret directly due

to the squared and cubic terms, they indicate patients prefer hospitals closer to their homes. For example, an increase in the distance traveled to a nonprofit hospital from 10 to 15 kilometers corresponds to utility decrease equaling 24 percent of the total value associated with travel distances. The dummy variable for being admitted to the closest hospital is also positive and significant indicating that patient utility increases when the patient is admitted to the closest hospital to his home.

While the coefficient on the log of beds is negative, when it is interacted with a dummy variable for being a nonprofit hospital, the overall effect is significant and positive. This suggests that for nonprofit hospitals, patient demand is endogenous to capacity and a nonprofit facility should expect more patients if it adds beds.

The correlation for the nesting parameter is close to one half and significantly different than zero indicating that some patients do have a preference for certain hospital ownership types. These estimates are comparable to the correlation coefficient found in Gowrisankaran et al. (2010) which estimates patient preferences among rural hospitals.¹⁹

As mentioned earlier, I do not include a measure of hospital prices. By including a hospital fixed effect, I capture the relative time-invariant prices of hospitals. Additionally, patients rarely know the price of a procedure or visit when choosing a hospital, and in many cases, only pay a small fraction of the total price with insurance covering the rest.²⁰ For these reasons, I am confident that omitting prices does not substantially impact my estimates.

In these regressions, I have treated all patients as having homogenous preferences for hospital characteristics (e.g. distance, ownership type, capacity). However, in

¹⁹ It is also possible that the correlation parameters is picking up any differences in insurance coverage for patients across ownership types because a patient may prefer hospitals of a certain ownership type because his insurance covers care at these facilities.

²⁰ I ran similar regressions to those in Table 2.3 which limited the sample to Medicare patients and found similar results. As all Medicare patients have substantial coverage of the hospital costs, they may have lower sensitivity to prices than other patient subgroups.

practice, patients who are hospitalized for an emergent health problem may have different preferences for distance traveled than those admitted for an elective procedure. Likewise, my analysis is restricted to general hospitals, whereas some patients may have a preference for certain hospitals based on the type of procedure performed (e.g., heart surgery). While my nesting parameter may capture a portion of this effect if hospital ownership types are correlated with the characteristics desired by patients, future research will address this in more detail by allowing patient preferences to vary across hospitals based on patient characteristics.²¹

2.5.2 Hospital Capacity Decisions

Using probit regressions, I estimate hospital choice probabilities analyzing the effect of variables in the state vector on the likelihood the hospital increases capacity in the following year. Separately, I estimate these probabilities for contractions in capacity. Conditional on making an adjustment, I estimate the magnitudes of the capacity adjustment in separate OLS regressions for both increases and decreases.

One of the key variables of interest is an indicator for the presence of CON in the state where the hospital is located, which does not change over the time period analyzed. I therefore cannot include hospital fixed effects in these regressions as they be collinear with the CON dummy variable thus eliminating the estimated effect of the policy on hospital capacity decisions. I include numerous variables to control for hospital characteristics and the competitive environment such as market characteristics and local demographic trends.

Both the probit and OLS results are presented in Table 2.4. The specifications in the first two columns are probit regressions where the dependent variable is a dummy for whether the hospital increased or decreased its capacity in the following year.²²

²¹ For example, we may think that patients seeking treatment for an emergency will be more sensitive to distance traveled than those receive an elective procedure.

²² I also combine these first two probit regressions using an ordered probit specification equaling

Table 2.4: Hospital capacity adjustment regressions: “Adjust” regressions give probit marginal effect coefficients whereas the “How Much” regressions provide OLS coefficients. Robust standard errors are in parentheses. All regressions include time fixed effects. \hat{p} refers to the estimated probability from specification (1) or (2) that the hospital adjusts capacity and is included in two of the “How Much” specifications. N corresponds to hospital-years. For the probit regressions, the value given for R^2 refers to the pseudo-R2.

	(1)	(2)	(3)	(4)	(5)	(6)
	Adjust?		How Much?			
	Up	Down	Up	Up	Down	Down
Log(Beds)	0.062 (0.008)	0.072 (0.008)	-0.112 (0.016)	-0.203 (0.078)	0.065 (0.012)	0.220 (0.061)
CON	-0.082 (0.022)	0.016 (0.019)	0.007 (0.024)	0.128 (0.101)	0.036 (0.036)	0.072 (0.037)
Turnaway Prob.	0.311 (0.106)	0.132 (0.108)	-0.059 (0.094)	-0.560 (0.397)	0.149 (0.105)	0.426 (0.150)
Pop Change	0.00002 (0.00001)	-0.00002 (0.00001)	-6.52e-07 (7.19e-06)	-0.00003 (0.00002)	5.80e-06 (9.19e-06)	-0.00004 (0.0002)
Elasticity	0.003 (0.001)	0.0005 (0.0006)	7.16e-06 (0.0003)	-0.005 (0.004)	-0.002 (0.001)	-0.002 (0.001)
Market Hosps	0.034 (0.018)	-0.006 (0.018)	-0.0006 (0.003)	-0.048 (0.039)	-0.008 (0.004)	0.028 (0.024)
Market Beds	-0.006 (0.003)	0.001 (0.003)	0.006 (0.018)	0.009 (0.007)	0.049 (0.023)	-0.003 (0.004)
\hat{p}				1.46 (1.07)		-2.14 (0.759)
	Probit	Probit	OLS	OLS	OLS	OLS
R^2	0.036	0.047	0.211	0.216	0.097	0.111
N	5,101	5,101	1,113	1,113	1,114	1,114

The third through sixth specifications use an OLS framework where the dependent variable is the difference between the log of the capacity total after the adjustment and that from before the change. For the capacity increase specification, the dependent variable will always be positive, and for the capacity decrease specification, it

0 if the hospital decreases capacity, 1 if it makes no change, and 2 if it increases capacity. The coefficient estimates were similar, but the overall fit was weaker. This finding was likely due to some variables that affected the probability of increasing and decreasing capacity in the same direction. For example, larger hospitals are more likely to both increase and decrease capacity and this is difficult to capture in an ordered probit model.

will be negative.

The coefficients in the first column indicate a hospital is more likely to add capacity if it has more beds in the current period. Additionally, the presence of CON in the state where the hospital is located decreases the probability that a hospital adds beds. This result indicates that regulations affect hospital capacity decisions and hospitals in states with CON are less likely to add beds which is consistent with the notion that CON makes capacity increases more costly or improbable.

The probit regression on capacity increases also indicates that a hospital with a higher turnaway probability is more likely to increase capacity. The elasticity measure, which indicates how many additional patients a hospital is expected to receive if it increases capacity by one bed, is positive; this suggests hospitals that are better suited for business stealing via capacity investments are more likely to add beds. The coefficients on the market's total number of hospitals and total beds suggest a hospital is more likely to increase capacity when there are more local competitors, but as these competitors get larger, this likelihood decreases.

The coefficients for the model estimating the probability a hospital adjusts capacity downwards are less definitive as the only statistically significant coefficient is the log number of beds. As with capacity increases, a hospital is more likely to decrease its capacity when it is larger. The presence of CON in the hospital's state does not significantly affect the probability that a hospital adjusts capacity downwards.

While the pseudo- R^2 values for my probit regressions indicate that my models only predict a small fraction of the variance in capacity decisions between hospitals, they are comparable to the values found using this sort of dynamic model of firm behavior such as Ryan (2011).²³ For the simulations, I use similar regressions, but include additional variables and interactions to more fully characterize what impacts

²³ When looking at the probability of entry, Ryan (2011) reports likelihood ratio test values under 0.03 when excluding region fixed effects.

the probability that a hospital adjusts its capacity. More specifically, I interact the dummy variable for the presence of CON with a number of other hospital variables such as size and ownership type to allow the regulation to affect hospitals differently based on observable characteristics. In these probits, the pseudo- R^2 's increase to 0.067 and 0.056 respectively.

While the results become more difficult to interpret with the additional variables, the estimated impact of CON regulations on a hospital with mean characteristics is similar to those values estimated in Table 2.4. For example, the average marginal effect of CON on the probability that a hospital increases capacity is -0.059, whereas the value in Table 2.4 is -0.082. Because I allow the impact of CON to vary based on hospital characteristics, some hospitals will be more affected by the presence of regulations than others. For example, if a hospital decreases from having the median capacity (214 beds) to the 25th percentile (110 beds), but keeps all other characteristics constant, CON decreases the probability that it adds capacity by an additional 0.6 percent.

I include four regressions estimating the amount by which a hospital adjusts its capacity conditional on having changed capacity between t and $t + 1$. Specifications (3) and (4) estimate the effect of variables on the amount by which a hospital increases capacity, conditional on adjusting upwards. Specifications (5) and (6) focus on cases where the hospital decreases capacity.

\hat{p} is a variable equal to the estimated probability that the hospital increases or decreases its capacity using the coefficients estimated in the probit regressions. For example, in specification (4), \hat{p} is equal to the probability that the hospital increased its capacity as estimated using the probit regression in specification (1). This variable allows the probability to impact the amount by which hospitals change capacity. The inclusion of \hat{p} makes the remaining coefficients more difficult to interpret so I also include specifications that exclude this variable. Specifications (3) and (4) indicate

that as a hospital becomes larger, the log of its capacity change decreases. This result is explained by the definition of the dependent variable in the OLS regressions. The change in capacity is equal to the difference in log capacity values. As a result, an increase of 5 beds will yield a larger dependent variable when the original capacity is small than when it is larger.²⁴ Similarly, the results suggest that the measure of capacity decreases by a larger amount when they the hospital’s original capacity is smaller. In specifications (4) and (6), the coefficients on \hat{p} indicate that as a hospital becomes more likely to adjust capacity, the expected magnitude of the change is greater, even after controlling for hospital observables.

2.6 Simulations

2.6.1 Methodology

I use simulations to predict how the U.S. hospital market would evolve over an extended period of time both with and without CON regulations. They are informative in two ways. The first benefit of these simulation model is that instead of estimating how CON regulations impact the likelihood of a capacity change between periods t and $t + 1$, simulations allow for hospitals to adjust to the regulatory policy over several periods. This is important for assessing the impact of the policy over a longer time horizon.

The second benefit of the simulations is that I am able to estimate how patient welfare changes from the hospital capacity decisions associated with each regulatory policy. This is done by simulating patient hospital choices conditional on the hospital capacity decisions in the simulations for each regulatory policy. Estimating CON’s impact on patient decisions is critical because I am ultimately interested in assessing the regulation’s effect on patient welfare; this framework will allow for the estimation of total patient distance traveled and total patient utility.

²⁴ For example, $(\log(25) - \log(20)) > (\log(100) - \log(95))$.

I assume that in the first year of the simulations, hospitals have the capacity observed in the data. Beginning in the next period, however, I consider two policy environments: one in which the hospital is subject to CON regulations, and a second where it is not. Conditional on having an estimate of the hospital's capacity in t , the simulation model estimates capacity in $t + 1$.

Conditional on capacity in t , I estimate values for each variable in the state vector for the empirical model of hospital capacity decisions. These include the hospital's turnaway probability, elasticity, market characteristics, and local demographic trends. I use these values and the corresponding coefficients estimated in the empirical model to estimate the probability that a hospital adjusts its capacity up or down.²⁵ For each hospital, I draw a random number from a uniform distribution and use this value and the estimated probability to determine whether the hospital increases or decreases capacity in this instance.²⁶ If I predict that the hospital adjusts capacity, I estimate the magnitude of this change using the estimated variables and coefficients from the OLS regressions. Having simulated hospital capacities in $t + 1$, I repeat this process to estimate capacities in $t + 2$.

I estimate capacity for each period observed in the data, which ends in 2009. I continue simulating hospital decisions for an additional 10 years to get a longer time horizon in which to evaluate the CON regulation's impact. This requires assumptions about how the number of patients in each ZIP code will change over the final 10 years. I assume that for each patient ZIP code, the annual patient growth rate is equal to the rate over the past three years observed in the data. Additionally, I assume that during this time period, there is no hospital entry or exit.²⁷

I differentiate between the two regulatory policies using the dummy and interac-

²⁵ In a very small number of cases, the sum of these two probabilities exceeds one. In these cases, I deflate the probabilities so that their ratio is unchanged, but they sum to one.

²⁶ All of the simulated values reported in this section take the mean value from across 50 simulations.

²⁷ This assumption is reasonable given the low levels of entry and exit observed in the data.

tion variables associated with whether the hospital is in a state with CON. In the first simulation, I set the CON dummy variable equal to one and turn on all of the interaction variables that include CON status. In the second, the CON dummy and interaction variables are set to zero as the regulation is assumed to be eliminated. This leads to different probabilities that the hospital increases or decreases capacity, even if all of the other hospital and market variables are equal.

After running the simulations, I compare how hospital capacity decisions vary based on whether CON regulations are in place to determine the impact of CON regulations on hospital capacity. Additionally, I estimate the share of patients from every ZIP code who choose each hospital under both policy scenarios using the patient utility parameters and estimates of each hospital's turnaway probability. These predicted shares are used to estimate total patient utility.

2.6.2 Fit of Simulations

As discussed in the results section, the pseudo-R² values for the probit regressions are comparable to other papers using dynamic models of firm behavior. However, I study the fit of the simulations using two additional methods.

In the first, I compare the mean values of hospital capacity and turnaway probability for hospitals in the observed data with that in the simulation that uses the regulatory policy that is in place. For example, when looking at hospitals in New York, I compare the observed data to values from my simulations that assume CON is present because New York has CON regulations. If the simulations accurately reflect how hospitals make capacity decisions, these mean values should be similar to those observed in the data. The mean values are presented in Table 2.5.

Overall the mean values capacity and turnaway probabilities are similar for the observed data and the model predictions in both regulatory environments. In particular, the total capacity levels are nearly identical between the data and the simulations.

Table 2.5: Mean values in data versus simulations: Observations are at the hospital-year level. All values are means. This table compares the mean values observed in the data with those predicted in my model and separately consider hospitals subject to CON and those which are not regulated.

	CON		No CON	
	Data	Model	Data	Model
Log(Beds)	5.27	5.27	4.88	4.87
Turnaway Probability	1.2%	1.4%	0.9%	1.2%

While the mean values from the simulations show that the models appear to fit the data well, this does not necessarily indicate that the hospitals that are observed to adjust capacity in the data are the same who are predicted to adjust capacity in the simulation models. It is possible that the simulations accurately predict aggregate trends in capacity, but are less effective in predicting which hospitals are choosing to add or subtract capacity. To address this, I sum the total number of times a hospital is observed to increase capacity in the data, as well as in the number of times it increases capacity in the simulations using the observed regulatory environment. I then run regressions where the unit of observation is the hospital, the dependent variable is the number of times that a hospital increases or decreases capacity in the simulations, and the only independent variable is the number of capacity increases or decreases observed in the data. If my simulations correctly predict which hospitals are adding or subtracting capacity, the coefficients on the variables will be positive.

Table 2.6: Regressions analyzing fit of simulations: The observations occur at the hospital level. In the first specification, the dependent variable is the number of times in the model that the hospital increases capacity. The independent variable is the number of times this hospital increases capacity in the data. Robust standard errors are given in parentheses.

	Observed Data Changes	
	Increase	Decrease
Model Total Changes	0.382 (0.050)	0.143 (0.051)

The coefficients suggest that while the simulations do not perfectly predict hos-

pital capacity behavior, there is a strong positive correlation between the number of capacity increases observed in the data and the simulations. As a result, I am confident that my simulations do not just fit the data on an aggregate level and instead also are able to predict with some accuracy which hospitals choose to add or subtract capacity. Based on the coefficient sizes, the model appears to better fit capacity increases than decreases. These results hold when I run separate regressions for hospitals in states with CON, and without it indicating that the fit holds for both regulatory environments.

2.6.3 Estimates

Table 2.7 uses simulated data and compares the mean estimated capacities and turnaway probabilities under each regulatory policy. I look at the entire sample as well as breaking the analysis into years observed in the data and those that are simulated. This table will describe how the regulation affects aggregate capacity levels and turnaway probabilities over multiple times periods.

Table 2.7: Comparison of mean values for CON regulations and no CON: All values are means. The unit of observation is the hospital-year. “Observed Years” refers to estimates from 1995 through 2009.

	Observed Years		2010-2019		Combined	
	CON	No CON	CON	No CON	CON	No CON
Log(Beds)	5.19	5.19	5.01	5.05	5.12	5.13
Turnaway Probability	1.5%	1.4%	3.1%	1.7%	2.2%	1.5%
N	5,578	5,578	3,890	3,890	9,468	9,468

In the observed years, both regulatory policies have similar values for overall capacity and the turnaway probability indicating that if a state added or dropped regulations, the impact on these variables would be minimal in this time period. When looking at the projected impact from 2010 to 2019, however, hospitals have more capacity when they are not subject to CON regulations. While this difference in capacity is relatively small, the subsequent change in mean turnaway probabilities

is larger, with CON regulated hospitals turning away an estimated 82 percent more patients.²⁸

While the regulation’s overall impact on hospital capacity is small, it is possible that CON’s impact differs across hospitals based on their characteristics. Table 2.8 considers this possibility and looks at the mean predicted capacity and turnaway probability values based on the hospital’s size, ownership type, and whether it is in a large metropolitan area.

Table 2.8: Comparison of simulation means by hospital characteristics: Values in the first four columns are means. The fifth column gives the percent difference between the values in the third and fourth columns. The unit of observation is the hospital-year. The categorization by log beds in the first two rows uses the number of beds observed in the data. For observations after 2009, I use the observed bed value from 2009. Using this criteria, any hospital with 245 or more beds is categorized as large. A hospital is categorized as being in a small MSA if it is not in an MSA, or if it is in an MSA with a population less than 1,000,000. It is categorized as being in a large MSA if it is an MSA with a population greater than 1,000,000.

	Log Beds		Turnaway Probability		
	CON	No CON	CON	No CON	% Diff.
Log(Beds) > 5.5	5.91	5.86	2.5%	2.1%	18.1
Log(Beds) < 5.5	4.55	4.61	2.0%	1.1%	81.6
Local Gov’t Owned	4.64	4.65	1.1%	0.5%	118.5
Nonprofit	5.27	5.28	2.4%	1.8%	33.4
For-Profit	4.53	4.62	3.4%	1.3%	158.7
Small MSA	4.65	4.63	2.1%	1.8%	18.5
Large MSA	5.56	5.60	2.1%	1.3%	64.9

The results suggest that CON regulation affects the average capacity of both large and small hospitals, but in opposite directions. Large hospitals actually have slightly higher capacity when CON is in place whereas small hospitals have lower

²⁸ The divergence in turnaway probabilities later in the sample occurs most noticeably in Arizona. During this time period, the number of patients in Arizona is estimated to grow rapidly and my simulations predict that under these conditions, CON regulations prevent the hospitals from increasing capacity sufficiently to meet the growing demand. While the patient growth rate in Arizona may be larger than in other states, this finding indicates that CON regulations may be problematic nationally if the population continues to age and demand more hospital services going forward.

capacity under CON regulation. Turnaway probabilities are lower without CON for both large and small hospitals, though the disparity is greater for smaller hospitals with turnaways being 81.6 percent more likely when the regulation is in place. When looking at hospitals by ownership type, the difference in capacity with and without CON regulation is minimal for local government and nonprofit hospitals. For-profit hospitals are most affected and have a higher capacity when no CON regulations are in place. The turnaway probabilities are again larger for all ownership types when CON regulations are in place with the greatest difference being a 158.7 percent change in turnaways for for-profit hospitals. These findings are consistent with the qualitative research from Yee et al. (2010) suggesting that smaller hospitals are disproportionately affected by CON.

CON also impacts patient welfare through its effect on hospital capacity decisions. I therefore estimate this impact and present the results in Table 2.9. I measure utility in several ways. I first look at the difference in the distance traveled under the two regulatory policies. The percentage difference between the policies is small with patients traveling 0.3 percent further over the full simulation period when CON regulations are present. I then consider how total patient utility changes based on the regulatory environment. Calculating total utility using the patient utility parameters, I estimate that utility is 1.2 percent higher when there are no CON regulations in place. The majority of this difference occurs in the last ten years of the simulation period. This percentage is not informative by itself, however, as utility is not expressed in a meaningful unit measure. I therefore estimate the necessary decrease in distance traveled for all patients when CON regulations are in place to get the same total utility as when no regulations are present. This value is estimated to be 5.1 percent, indicating that if all patients decreased their distance traveled by this amount they will, on average, be as well off as when no regulations are in place.

Given that the average patient travels roughly 15 miles to attend a hospital, this

Table 2.9: Utility comparison based on CON status: All values are percent changes. In the first two rows, the percent difference represents the change when going from having CON regulations to no regulations. The third row represents the decrease in distance traveled for patients when CON is present so that total patient utility is equal to that when there are no regulations.

	Utility Increase w/o Regulations		
	1995-2009	2010-2019	All
Travel Change (%)	0.2	-1.0	-0.3
Utility Change (%)	0.1	2.4	1.2
Compensating Distance Decrease (%)	0.7	10.2	5.1

indicates the average patient receive the equivalent of nearly one mile less travel in additional utility if CON does not exist. This demonstrates that the regulation’s effect on patient welfare are sizable and negative. These results strongly indicate that CON should be discontinued in states where it is currently in place.

2.7 Conclusion

I use patient and hospital data to build an empirical model that estimates the probability that hospitals adjust capacity. The model analyzes the impact of CON regulations on hospital capacity decisions and finds that the presence of CON decreases the probability that a hospital adds capacity. Additionally, I estimate patient demand by using a multinomial nested logit model to estimate how parameters such as distance traveled and hospital size affect patient hospital preferences.

Using the parameters for hospital capacity decisions and patient utility estimated in the model, I run simulations that predict how hospital capacity would evolve differently between CON requirements and an unregulated environment. The simulation results suggest that CON has a limited effect on mean hospital capacity and turnaway probabilities in the short term, but over time, hospitals have more capacity and turn away fewer patients when there are no regulations in place. Additionally, the simulation results indicate that the differences in capacity are greater for smaller

and for-profit hospitals. The simulations allow for the estimation of patient welfare using the patient utility parameters estimated in the empirical model. They find that the presence of CON regulations decreases patient welfare considerably and, to make patients as well off under CON as when no regulations in place, each patient would have to decrease his distance traveled by 5.1 percent.

This analysis is informative in detailing how CON regulations' impact hospital capacity, and ultimately how this affects patient welfare. My findings suggest that patients are made worse off by CON regulations. In future research, I plan to develop a model that also estimates the impact of CON on hospital investment costs which will allow for a complete welfare comparison between the regulatory policies. However, given the extensive literature suggesting that CON's affect on total spending is limited, I am confident that I will again find that the regulation decreases welfare. As debate over CON's impact continues, this research will help to inform policy makers about a major component of CON regulations and how it impacts both patients and hospitals, as well as total welfare.

Speculators and Middlemen: The Role of Flippers in the Housing Market

By Patrick Bayer, Christopher Geissler, and James W. Roberts

3.1 Introduction

Local, regional, and even national housing markets periodically experience substantial booms and busts that have significant impacts on the real economy. Yet our understanding of the root causes of these phenomena remains incomplete. Recent research on the US housing market during the 2000s shows that the share of homes purchased by investors, i.e., those purchasing a home to hold or rent not as a primary residence, increased very rapidly in exactly those metropolitan areas that experience the largest housing bubbles during this period. The primary goal of this paper is to study the behavior of these investors in much greater detail.

Housing provides a prime example of a thin market for high value, durable goods. In such markets, intermediaries often play a number of important economic roles. In housing markets, these intermediaries are often referred to as “flippers,” a term

used to describe individuals or firms that buy homes with no intention to reside in or rent the property, but simply to quickly resell at a profit. Flippers may serve as *middlemen*, purchasing from sellers with substantial holding costs who cannot afford to wait for the right buyer, thereby providing liquidity in this heterogeneous market. They may also make important *physical investments* in houses, if improvements are optimally made when a house is empty or can be done at relatively low cost.¹ And, they may operate as *speculators*, seeking to exploit arbitrage opportunities made possible through either superior information about market fundamentals, or by exploiting deviations from the fundamentals resulting from naïve decision-making on the part of other market actors.

The goal of this paper is to identify the activity of flippers operating in these three distinct economic roles and to study their impact on the housing market. Our analysis is based on comprehensive micro-level house transaction data from the Los Angeles metro area from 1988-2009 that allows us to identify flippers and study their activity. We introduce a novel research design to decompose observed returns on each flipped home into four components: (i) any discount on the transaction price at the time of purchase, (ii) any premium on the price at the time of sale, (iii) market returns during the holding period, and (iv) any physical improvements made to the property by flippers. This last source of returns presents an econometric challenge for us since improvements are not directly observable in the data. It is here that our research design exploits the panel nature of the data to identify these components of the return. In particular, by examining sales prices in transactions between pairs of non-flippers both prior to and following the period where a flipper buys and sells the property, we are able to control for any persistent changes in the houses unobservable

¹ In this role, flippers may not only generate returns for themselves (if they can make these improvements at relatively low cost), but may serve to maintain and restore the housing infrastructure of a community, leading to positive externalities in terms of house values and the local tax base for neighboring homeowners.

quality that may have been due to flipper investment.

We establish that middlemen and speculators (i) follow very distinct strategies for when and where to buy and (ii) generate returns from almost completely distinct sources.² Middlemen hold properties for very short periods of time (a median of six months) and earn most of their return by buying houses relatively cheaply. Flippers in this role tend to be professionals; the same individuals are observed transacting numerous properties throughout the sample period. Market timing is not an important source for their returns; they operate throughout booms and busts in the housing market and target submarkets that, if anything, are appreciating more slowly than the rest of the metro area.

By contrast, speculators tend to enter the housing market at an increasing rate as prices rise. They do not buy at much of a discount or sell at a premium, but instead earn almost their entire return through timing the market, i.e., earning the average market return. They operate only during boom times and target submarkets of the Los Angeles area that experience both an above average rate of appreciation in the short term (next 1-2 years) and a sharp decline in the intermediate term (3-5 years). In this way, entry by speculative flippers is strongly associated with the amplification of local housing price cycles.

Interestingly, there are a number of signs that the speculators that entered the market during the most recent housing boom may not be particularly sophisticated. First, perhaps fueled by access to equity in their primary residence as prices rise, speculators tend to be amateurs that are not particularly experienced at flipping houses. Secondly, many speculators continue to purchase properties at a rapid rate through the point that market reaches its peak and hold a large fraction of their pur-

² Although not the primary focus of the paper, our research design also enables us to measure the impact that flippers have on the market through investment in physical home improvements. We estimate that flippers of both types (speculators and middlemen) invest little more than the typical homeowner in their homes, implying that their impact on the market comes primarily from their roles in transacting and holding properties.

chased properties well past the peak, thereby experiencing substantial losses.^{3,4} Here, our paper is related to Brunnermeier and Nagel (2004) who examine the potentially destabilizing force of hedge funds during the technology bubble. While their findings suggest hedge funds may be more sophisticated than the speculators that we study (a fact which should not be terribly surprising), like their paper, ours too calls into question a central tenet of the efficient markets hypothesis: that it is always optimal for rational speculators to attack a bubble.

Taken together, our analysis provides strong evidence that flippers play multiple economic roles in the housing market. As middlemen, flippers may significantly enhance welfare by providing liquidity to high holding cost sellers. As investors in durable good quality, they may promote neighborhood gentrification. Finally, as speculators, they are strongly associated with, and likely contribute to, increased volatility in local housing markets, which has serious economic and social consequences. Given these roles' dueling implications for welfare, and the fact that much of their activity occurs over short holding periods, it may be difficult and suboptimal to target flippers with anti-speculative policy prescriptions such as transaction taxes, which have been suggested in other speculative markets,⁵ or by limiting their ability to finance investment.⁶

³ While our research design allows us to decompose the returns for a particular flipped house, it does not allow us to measure the returns to operating in the flipping business *per se*. In particular, because the same individuals that flip certain properties may hold others for a very long time (presumably as rental units), the lack of information on rents in our dataset precludes us from calculating returns for individual flippers. Moreover, our analysis provides only an *ex post* estimate of the particular realization of returns for Los Angeles over the study period, rather than an *ex ante* measure of expected returns. For these reasons, we confine the focus of our paper to a study of the behavior and decomposition of the sources of returns for the distinct types of flippers described above.

⁴ Whether the current boom and bust cycle was in fact a bubble in housing is a subject of intense debate. See Himmelberg et al. (2005) for a detailed discussion.

⁵ See, for example, Tobin (1974), Tobin (1978), Eichengreen et al. (1995) or Summers and Summers (1988).

⁶ A 2006 HUD regulation preventing FHA financing for houses sold within 90 days of purchase likely had this effect. More generally there is reason to believe a wave of "anti-flipper" sentiment

The paper proceeds as follows: Section 3.2 presents a simple theoretical discussion of the economic roles of flippers as middlemen and speculators. Section 3.3 describes the unique dataset used and the definition of a flipper. Section 3.4 outlines the research design that will allow us to identify flipper returns and investment. Section 3.5 gives our primary empirical results as well as robustness checks for these findings. Section 3.7 extends our analysis to the neighborhood level and focuses on what types of neighborhoods flippers target and their impact on these neighborhoods. Section 3.8 investigates whether some flippers were caught holding houses when the housing market turned. Section 3.9 concludes.

3.2 A Conceptual Framework

To frame the empirical analysis, it is helpful to present a conceptual discussion that highlights the potential economic roles of flippers as middlemen and speculators.

3.2.1 Flippers as Middlemen

Housing markets are a classic example of a thin market for high-valued durable goods and, as a result, the home-selling problem is generally modeled in a search theoretic framework.⁷ When selling a home, a household lists the property for sale and waits for offers from buyer(s) to arrive, determining its reservation price (i.e., minimum acceptable offer) as a function of market conditions and its motivation to sell or holding costs. In general, holding costs for comparable properties vary across sellers depending on how quickly they need to relocate, their consumption value from residing in the house (if they continue to do so), and their borrowing costs.

Flippers who purchase a property with plans to immediately put the house back on the market face an analogous home-selling problem to that of other home-owners.

may lead to further, similar legislation.

⁷ For example, see Goetzmann and Peng (2006).

As a result, flippers will be able to profitably bid above the seller's reservation price only when their holding costs are lower than that of the seller. The holding costs of flippers will generally be governed by their borrowing costs or, more generally, their cost of capital.

Because flippers do not receive consumption value from residing in the home, their holding costs will generally be greater than those of a large fraction of sellers who can continue to reside in their home while waiting for offers to arrive and face little pressure to sell quickly. A motivated seller, however, may have a holding cost that exceeds those of flippers if, for example, the seller needs to relocate to a new city or sell a house quickly to settle a divorce.⁸ When transaction costs are sufficiently low, a flipper's maximum bid will exceed the reservation price of sufficiently motivated sellers and flippers will be able to purchase the property with the intention to immediately re-list it for sale, waiting more patiently than the existing home-owner for a strong offer to arrive.

The economic function of flippers that buy properties from especially motivated sellers, hold them for a short period, and then sell them to a buyer that places a sufficiently high value on the property is that of a middleman. When flippers operate as middlemen, motivated sellers are dynamically matched to future buyers that place a higher value on the property (on average) than those who the seller would have sold to in the absence of flippers. In this capacity, flippers provide liquidity to the market, essentially providing a price floor that is a function of their cost of capital and market conditions, and their presence generally improves the economic efficiency of the market.

⁸ Springer (1996) finds that distressed sellers deal more quickly and sell for less than other sellers. Glower et al. (2003) find that when a seller takes a new job, he sells faster than average, indicating he likely has a higher holding cost.

3.2.2 Flippers as Speculators

The theoretical finance literature supports (at least) two broad rationales for the existence of speculators in the housing market. Most obviously, efficient market theory admits an economic role for speculators that have access to better information than the broad set of agents participating in a market. Given the decentralized nature of the housing market, with many individuals participating in the home buying or selling process only a handful of times during their lives, it is easy to imagine that some market professionals might be especially well-informed or be able to process information in a sophisticated way that generates arbitrage opportunities. In the classic theory of efficient markets, speculators, acting on the basis of their superior information, serve to align prices more closely with market fundamentals, generally improving the efficiency of the market.

Modern finance theory admits a wider range of strategies for speculators and a much more ambiguous understanding of their impact on welfare and efficiency.⁹ One starting point for much of modern finance theory is the presence of a set of naïve market actors, noise traders, who are subject to expectations and sentiments that are not fully justified by information about market fundamentals. By following simple strategies, such as chasing trends, or by sticking to rules of thumb, noise traders can create distortions between prices and market fundamentals.

In this setting, potential arbitrageurs face multiple risks. Even if they are fully aware that prices have temporarily deviated from the fundamentals, there is a risk that they might deviate further in the short-run (depending on the beliefs and activity of the noise traders) before eventually falling back in line with the fundamentals. It is not always optimal, therefore, for arbitrageurs to simply take a short position on any observed market deviations from the fundamentals.

⁹ See Shleifer and Summers (1990), Barberis and Thaler (2003) and Shiller (2003) for summaries of this literature.

In fact, it can be optimal to pursue a much wider range of strategies. If, for example, noise traders engage in positive feedback trading - i.e., have a tendency to extrapolate or to chase the trend, it can be optimal for rational speculators to jump on the bandwagon (DeLong et al. (1990)). By buying as noise traders begin to get interested in a market, speculators actually fuel the positive feedback trading that motivates the noise traders. And, by selling out as the market nears a peak, speculators speed the return of the market to the fundamentals. In this case, rational speculators take advantage of the noise traders by strategically selling out before the noise traders realize the bubble is about to burst. In this way, it is easy to see that the welfare consequences of the existence of speculators need not be positive. To the extent that their actions fuel bubbles and increase volatility in the market, speculators tend to decrease welfare and market efficiency.

Finally, it is important to keep in mind that the relatively high holding costs of flippers, operating as either speculators or middlemen, limit their overall impact on the market. Because flippers generally do not reside in the property while holding it, they will only purchase properties when their expected returns, whether achieved by buying low from motivated sellers or speculating on market appreciation, exceed their expected holding and transactions costs. For middlemen, opportunities to buy may occur under any market conditions, provided they are able to identify especially motivated sellers (those with higher holding costs than their own). Speculators will require expected market appreciation to be sufficiently high to justify their purchases and, therefore, will be active in only those times and places where conditions are right.

3.3 Data

The main goal of our paper is to study the behavior of flippers operating as either middlemen or speculators in a housing market over a long period of time. The pri-

mary data set that we have assembled for our analysis is based on a large database of housing transactions from Dataquick. For each transaction, the Dataquick data contain the names of the buyer and seller, the transaction price, the address and property identification number, the transaction date, and numerous characteristics including, for example, square footage, year built, number of bathrooms and bedrooms, lot size and whether the house has a pool.

The data set that we use for our analysis includes the complete census of housing transactions in the five largest counties in the Los Angeles metropolitan area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura counties), between 1988 and 2009. Dataquick collects information from two sources. Its transaction variables, which include the date, price, and names of the buyer and seller are based on publicly available data and thus cover every transaction in the Los Angeles metropolitan area during the study period. Its housing attribute variables are drawn from a second public source, the local tax assessor's office. A key drawback is that Dataquick only maintains a current assessor file, overwriting historical information, so that only the attributes of a house are only known as of 2009, rather than the date a sale took place. This prevents researchers from tracking major home improvements by using changes in house characteristics over time. This drawback motivates the research design that we introduce below to address the possibility of unobserved improvements to properties.¹⁰

From the initial set of transactions we drop a small number of observations in order to ensure that data used in the analysis is of especially high quality. In particular, we drop observations if a property was subdivided or split into several smaller

¹⁰ A research design to address the possibility of unobserved improvements to properties would be necessary even if Dataquick kept track of housing attributes on a continuous basis, as many home improvements (e.g., a renovated kitchen or bathroom) would not generally affect the more basic attributes of the home (e.g., lot size, square footage) collected by the tax assessor.

properties and resold, the price of the house was less than \$1,¹¹ the house sold more than once in a single day, the price or square footage was in the top or bottom one percent of the sample, there is a potential inconsistency in the data such as the transaction year being earlier than the year the house was built, or the sum of mortgages is \$5,000 more than the house price as this may indicate that the buyer intends to do substantial renovations. Table 3.1 provides summary statistics of our primary data set.

Table 3.1: Transaction-level summary statistics: The table shows transaction-level summary statistics for data that cover five counties in the LA area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura). Based on 3,544,615 transactions from 1988-2009. Loan to value is measured relative to the price paid.

	Mean	Std. Dev
Price	280,823	195,478
Square Footage	1,605	615
Transaction Year	1999.8	4.99
Year Built	1970.2	21.2
Has Loan?	0.908	0.289
Loan to Value	0.786	0.288
Number of Transactions	2.20	1.17

Table 3.1 provides summary statistics for our primary data set based on a full sample of over 3.5 million transactions between 1988-2009. Homes in the Los Angeles area are relatively expensive, especially considering their relatively small size, and recently built, compared to those in many other American cities. The vast majority of buyers take out a mortgage, with an average loan-value ratio of approximately 80 percent. Finally, for the homes that we observe sell at least once, the mean number of transactions during the sample period is 2.20, indicating that these homes sell on average once every 9-10 years.

Figure 3.1 shows the basic dynamics of prices and transaction volume for the

¹¹ A price of zero suggests that the seller did not put the house on the open market and instead transferred ownership to a family member or friend.

Los Angeles metropolitan area over the study period. Following a rapid increase in prices in the late 1980s, the early 1990s were a “cold” market period for Los Angeles, with prices declining by roughly 30 percent between 1992 and 1997 and transaction volume average only a little more than 30K houses per quarter during this period. Starting in the late 1990s and continuing until early 2006, the Los Angeles housing market experienced a major boom, with house prices more than tripling and volume roughly doubling to just short of 60K homes transacted per quarter. Most of the house price appreciation experienced during the boom, had already been conceded back by the end of 2008, with transaction volume falling to record low levels (less than 20K per quarter) over this period in late 2007 and early 2008.

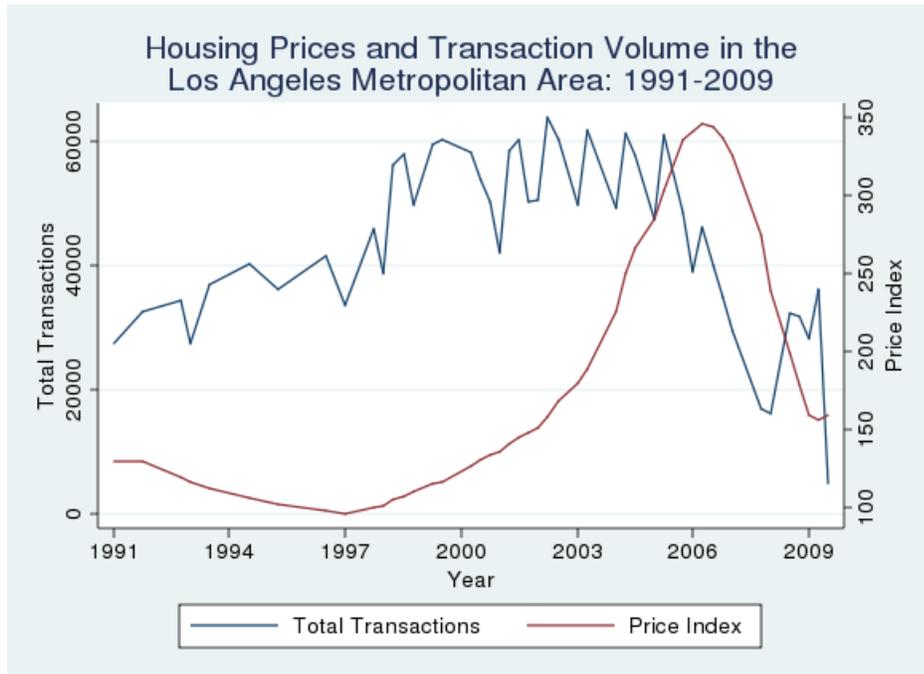


FIGURE 3.1: Transaction volume by county: Annual transaction volume cover five counties in the LA area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura), from 1988-2009.

In our analysis below, we will find it helpful to distinguish three key market periods: the “cold” market period in the early 1990’s (1992-1998), the “hot” or boom market period in the late 1990’s and early 2000’s (1999-2005) and the “post-peak”

period (2006-2009).

3.3.1 Flippers

A basic measurement challenge for anyone wishing to study the behavior of market actors that we might refer to as flippers, speculators, middlemen, or investors is how to identify them in the data. A clever approach utilized by Haughwout et al. (2011) is to examine credit reports, looking for cases where the same individual is observed to hold mortgages on multiple properties. While some instances of second home purchases may be motivated by reasons other than pure investment (e.g., vacation properties, first homes purchased for children), by carefully documenting the pattern of new home purchases by individuals who own multiple properties, these authors are able to provide a reasonable proxy for the amount of investor activity in a market at a given point in time. Haughwout et al. (2011) documents that a large fraction of new mortgage originations (upwards of 50 percent in some markets) during 2004-2006 in the states that experienced the largest housing booms/busts were made to individuals who already owned at least one house.

Figure 3.2 reports the time series for three distinct proxies for investor and flipper behavior in the Los Angeles market between 1991-2009 derived from our transaction data set. The first of these, labeled “Second Homes” is constructed in the spirit of Haughwout et al. (2011). In particular, we categorize a home as a second home if the buyer’s name matches that of an individual that we also observe to be simultaneously holding another property in our dataset. A fundamental problem with this definition, of course, is that for an individual to be observed as a home-owner at all, they need to have purchased a home since the beginning of our study period in 1988. Thus, our measure of “Second Homes” is likely to substantially understate the amount of actual second home purchases, especially near the beginning of the sample period. For the latter reason, it is important not to over-interpret the trends in the measure.

However, even subject to this limitation, our measure of second home purchases tracks that of Haughwout et al. (2011) very closely, rising to a peak of nearly 30 percent of the market in 2006.¹²

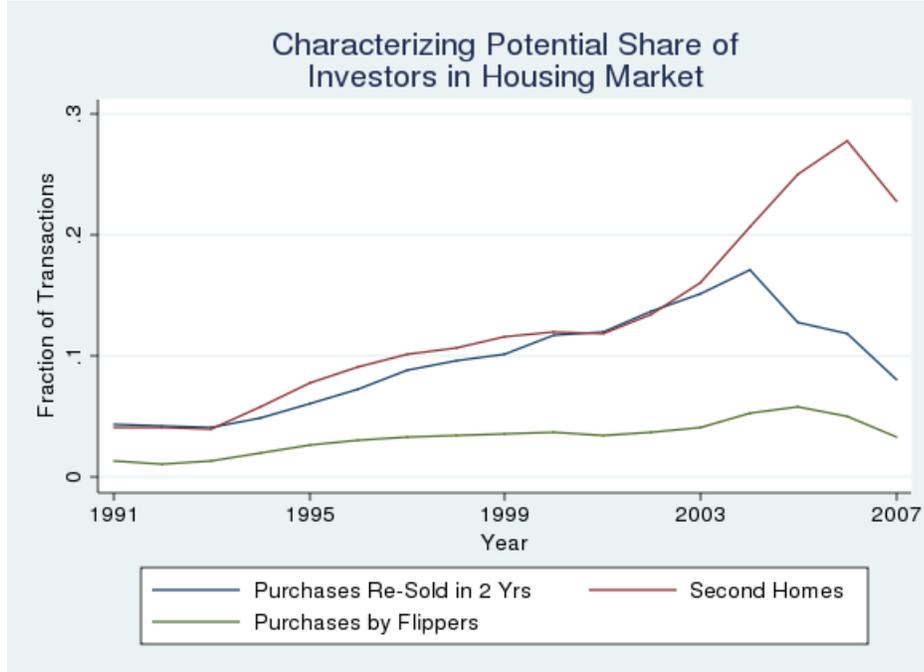


FIGURE 3.2: Level of flipper activity over time: The figure plots three data series that serve as proxies for investor and flipper activity. “Second Homes” plots the fraction of new purchases by individuals with names matching those of a current home-owner in the data set. “Re-Sold within Two Years” plots the fraction of all homes purchased in a given quarter that are res-sold within two years. “Flipper Purchases” depicts the fraction of homes purchased by individuals that are identified as flippers for the purposes of our analysis. The data cover five counties in the LA area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura), from 1988-2008.

The broad category of housing market investors, of course, includes many that would not be consider flippers operating in the role of middleman, speculator, or physical improver of a property. Instead, many investors purchase properties with

¹² A second limitation of our definition of second home purchases is that it is based on name matches and, therefore, might be overstated because of false matches of different individuals with the same name. The qualitative pattern of a sharp peak in the presence of second home purchasers in 2004-2006, however, is not affected by the exclusion of the most common names observed in the data set.

the plan to rent the property to a tenant and hold it for an intermediate or long period of time. To identify flippers, therefore, we look for evidence that an individual is generally engaged in a strategy of purchasing homes with the intention of re-selling the property after a relatively short holding period.

A second time series shown in Figure 3.2, “Purchases Re-Sold within Two Years” simply reports the fraction of all homes (regardless of the buyer) purchased in a given quarter that are re-sold within two years. Well over 15 percent of all homes purchased near the peak of the boom in 2003-2005 were re-sold within two years, a rate that is more than triple the corresponding rate for the cold market period of 1991-1994, when home prices were declining. While certainly a portion of the buyers that re-sell homes within two years of purchase are owner-occupants rather than investors, this time-series provides a proxy for flipper-like behavior in the market throughout the cycle, providing a sense that this activity is both highly pro-cyclical and economically significant.

For the vast majority of our analysis, we focus not on flipped homes per se, but on a set of individuals and firms that we identify as “flippers.” We identify flippers using two pieces of information in our dataset: the period of time that a house was held and the names of buyers and sellers. We define a flipper to be anyone that we observe buying and selling at least X different properties while holding for less than Y years. For the vast majority of our analysis we set $X=2$ and $Y=2$, (i.e., flippers are those who have bought and sold at least two properties, each with a holding period of less than two years), but we also explore in detail how flipper behavior, strategy, and returns are affected by variation in X and how robust our results are to variation in Y .

Limiting our definition of flippers to individuals that we observe buying and selling multiple homes after a short holding period provides a conservative measure of flipper activity, as we certainly miss any individuals who engage in this activity

only once during the sample period or who tend to hold properties for slightly longer periods of time. We do so to make sure that we avoid (as much as possible) counting normal owner-occupants as flippers. The final data series, “Flipper Purchase” shown in Fig 3.2 reports the fraction of housing transactions in each quarter that were made by individuals that we define as flippers. Note that this measure includes all homes purchased by flippers regardless of how quickly these homes are re-sold. This time series generally tracks housing market conditions, peaking at over 5 percent of all purchases in 2006, a rate that is 4-5 times higher than the rate of flipper activity in the early 1990s.

Overall, the three broad metrics of investor or flipper activity shown in Fig 3.2 show a consistent pattern of pro-cyclical behavior, with purchases by these agents reaching a maximum at the peak of the housing boom, at levels that are roughly three times the level activity observed during the market trough in the early 1990s.

3.3.2 Purchase Activity by Flippers

In the analysis that follows, we document consider heterogeneity in flipper behavior, strategy, and outcomes that is strongly associated with experience, i.e., the number of times that we observe them buying and selling homes after holding for a short period. Figure 3.3 shows the percentage of all homes purchased in a given quarter by flippers in four experience categories. In particular, we define the category Flipper 1 as those flipping 2 or 3 houses, Flipper 2 as those flipping 4-6 houses, Flipper 3 as those flipping 7-10 homes, and Flipper 4 as those flipping 11 or more homes. For the purposes of this definition, we count a purchase as a flipped home if it was re-sold within two years and we categorize flippers on the basis of their activity over the full sample period. The sum of all four data series presented in Figure 3.3 produces the total count of flipper purchases shown in Figure 3.2.

Figure 3.3 shows a very different pattern of activity on the basis of experience.

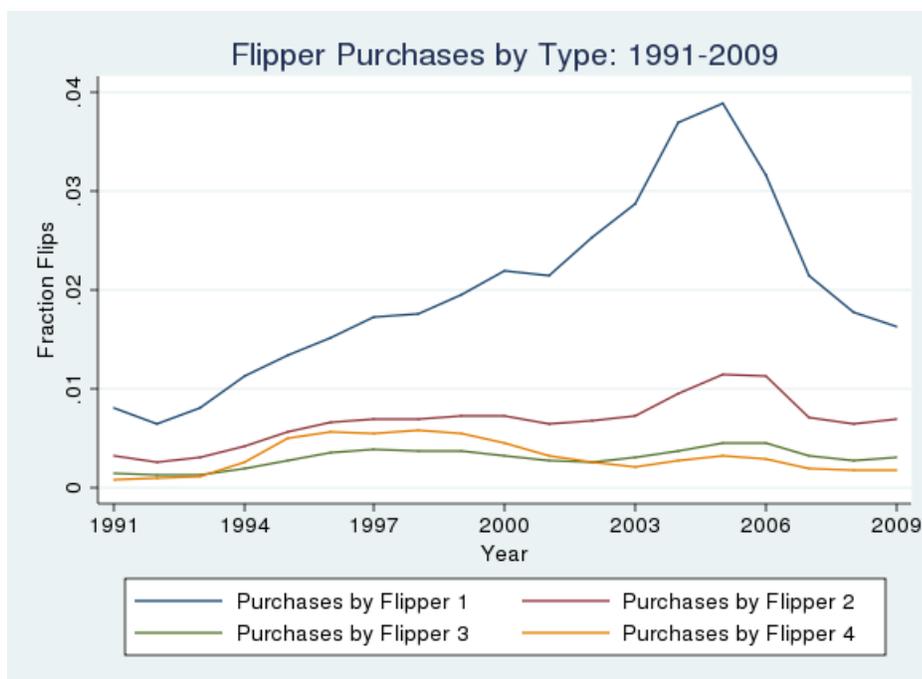


FIGURE 3.3: Flipper sales over time by type: The figure plots flipper sales by year. Flipper 1’s: 2-3 flips, flipper 2’s: 4-6 flips, flipper 3’s: 7-10 flips, flipper 4’s: 11 or more flips. The data cover five counties in the LA area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura), from 1988-2008.

The purchase activity (as a percentage of all homes sold) by more experienced flippers (Flipper 3 and Flipper 4) is relatively constant over the study period, actually peaking in the colder market period of the mid-1990s. This pattern of activity is consistent with the view that the more experienced flippers tend to operate as middlemen, looking for opportunities to buy from very motivated sellers with higher holding costs than their own, opportunities that are just as (or perhaps more) likely to arise in cold versus hot market conditions.

The purchase activity by inexperienced flippers (Flipper 2 and especially Flipper 1) is highly pro-cyclical, rising from a very small percentage of the overall market in the early-mid 1990s to almost 5 percent of the market in 2004-2006. This pattern of activity is consistent with the view that many inexperienced investors were drawn into the market during the boom period. While a measure of activity is not enough

to establish the motives of these flippers, the timing of their purchases is certainly consistent with a view that they are seeking to make a quick speculative gain on the basis of market appreciation.

It is worth noting at the outset that that our definition of flipper experience is far from perfect. In particular, our measure of experience is based on activity over the full study period. Thus, many of the flippers that we categorize as inexperienced may, in fact, ultimately become more experienced if they continue to flip homes after our study period ends. Moreover, we certainly expect survival in the flipping business to be non-random, with those that perform better and can operate profitably lasting long enough in the business to reach the higher experience categories. In each section of our analysis below, we will explicitly address these and other issues that arise due to our definition of flipper experience.

A final aspect of flipper purchase activity that is important to describe at the outset of our analysis is the heterogeneity in the attributes of homes purchased by flippers of each type. To this end, Table 3.2 summarized some basic characteristics of the homes purchases by flippers of each type.

Table 3.2: House summary statistics by flipper type: The table shows house-level summary statistics by type of flipper for data that cover five counties in the LA area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura). Standard deviations in parentheses. The right hand column includes repeat sales (homes that sell at least twice), which form the basis for our analysis below.

Flipper Type	All Flips				All Houses
	1	2	3	4	
Year Built	1965.1 (23.2)	1959.8 (23.6)	1955.4 (23.6)	1949.6 (22.6)	1971.0 (20.9)
Year Bought	2001.9 (3.19)	2001.4 (3.56)	2001.0 (3.71)	1999.6 (3.81)	2000.2 (5.00)
Square Feet	1,504 (592)	1,438 (513)	1,360 (516)	1,284 (440)	1,563 (592)
N	25,181	5,678	2,322	2,596	2,187,081

As the table makes clear, flippers, especially experienced flippers, generally pur-

chase properties that are somewhat older and smaller than the homes that sell in the market as a whole over our study period. The research design that we present below for estimating the sources of flipper returns is motivated in large part by the very real possibility that flippers may systematically purchase older homes or “fixer-uppers” that can benefit from substantial renovations or improvements before being re-sold. We also take additional steps to ensure that we compare the sources of returns for flippers for comparable houses - in particular, at the mean of homes that are sold in the market during the sample period.

3.3.3 Flipper Holding Times

Before turning to our analysis of the sources of flipper returns, we present a final descriptive characterization of the heterogeneous behavior of flippers at each experience level. In particular, Table 3.2 reports the fraction of homes purchased by flippers of each type within 1-4 years of the purchase. The table reports these statistics for our main study period of 1992-2005 and separately for purchases made in the cold market period of 1992-1998 and the hot market period of 1999-2005. To measure holding periods of up to four years, it is, of course, necessary to restrict attention to homes that were purchase at least four years from the end of the sample in 2009.

Table 3.3: Holding times by flipper type: The table reports the fraction of the homes purchased by flippers of different types sold with 1, 2, 3, and 4 years, respectively, in the LA area (Los Angeles, Orange, Riverside, San Bernardino, and Ventura).

Flipper Type	All 1992-2005				Cold 1992-1998				Hot 1999-2005			
	1	2	3	4	1	2	3	4	1	2	3	4
Sell Within:												
1 Year	0.263	0.269	0.321	0.571	0.227	0.312	0.429	0.684	0.308	0.302	0.328	0.562
2 Years	0.447	0.370	0.410	0.636	0.367	0.379	0.487	0.738	0.531	0.429	0.438	0.633
3 Years	0.512	0.443	0.474	0.676	0.425	0.435	0.532	0.763	0.597	0.504	0.507	0.679
4 Years	0.556	0.490	0.519	0.705	0.486	0.488	0.576	0.789	0.643	0.560	0.566	0.714

The figures reported in Table 3.3 show that flippers of all types hold a significant fraction of the properties that they purchase for more than four years. This may

reflect the fact that these investors intend to hold some properties as rental units or may reflect the fact that one of the purchases that we observe in the data is the flippers primary residence.

Table 3.3 also reveals significant heterogeneity in holding periods by both flipper type and market conditions. Experienced flippers, in particular those in category Flipper 4, are much more likely to re-sell homes after very short holding periods. In fact, they sell close to 57 percent of all of the homes they purchase within the first year and more that 70 percent within four years. During the cold market condition, this pattern is even more pronounced as Flipper 4's sell almost 70 percent of their purchases within a year and almost 80 percent within four years. This pattern is consistent with the notion that Flipper 4's purchase many homes with the intent to put them immediately back on the market and is consistent with the notion that these experienced flippers may be serving the economic function of a middleman, seeking to buy low from motivated sellers and re-sell quickly rather than hold properties during periods of rapid market appreciation.

By contrast, the figures for inexperienced flippers are qualitatively very distinct. Flipper 1's, for example, sell only 26 percent of their purchases within a year of purchase, a figure that steadily rises to 56 percent by the four year mark. This pattern of behavior is more consistent with a strategy of buying properties with the intention of capturing market appreciation – a strategy which, of course, requires a reasonable holding period.

3.4 Measuring the Sources of Flipper Returns - Research Design

Having documented time series pattern of purchase activity by flippers and experiences, we turn next to an analysis of the sources of their returns. At the outset, it is important to note several key limitations of this analysis that shape the interpretation of the results of our analysis. In particular, we do not observe (i) whether

a home is rented to a tenant during a holding period, (ii) any transactions costs that a flipper might pay while buying and selling a house, and (iii) the borrowing costs (or, more broadly, the cost of capital) that a flipper faces when procuring a mortgage (or cash investments) in order to purchase a property. Thus, we will not be able to calculate the profit, return, or rate of return that a flipper receives on each investment.

Instead, we will focus on only on the components of the returns that are associated directly with the purchase, holding, and sale of the property. In particular, we seek to identify (i) the discount that flippers get (relative to the average sales price in the market in the corresponding period at the time of purchase, (ii) the market return that they earn over the period that they hold the property and (iii) the premium that they get at the time of sale (again relative to the average sales price in the market at the time). By measuring these sources of flipper returns, we seek to categorize flippers on the basis of their motivation and strategy, i.e., whether they appear to be operating as middlemen or speculators.

An important complicating factor is that flippers may systematically make physical improvements to the properties that they purchase, improvements which are unobserved in our data set for the reasons mentioned in Section 3.3. A major concern is that a naive analysis of the sources of flipper returns from buying, holding, and selling a property might wind up counting money that flippers invested in improving a property as part of their return. If, for example, a flipper purchased a house for 150k, put 40k into it and re-sold it for 210k, we would want to measure this as a gain of 20k not 60k.

To address this problem, we develop a research design that aims to uncover the sources of flipper returns from buying, holding, and selling a property in the (potential) presence of unobserved investment. The method is based on a repeat sales index which we first review.

Case and Shiller (1987) introduced the repeat sales regression to generate a price index:

$$\log(p_{it}) = \alpha_1 yq_t + \alpha_2 id_i + \varepsilon_{it} \quad (3.1)$$

In equation 3.1, yq_t represents a quarter fixed effect and id_i is a house-level fixed effect on house i . Exponentiating the coefficients on the time fixed effects gives the price index for each quarter, which can be normalized to 1 in any quarter. This framework requires that quality is constant for each house across sales. Additionally, it assumes that the market evolves homogeneously across different regions of a metropolitan area.

We modify this framework by first introducing controls for whether the buyer or seller is a flipper. If the coefficient on the *Flipper Buyer* dummy is negative, it suggests that flippers buy houses below their expected value. A positive sign on the *Flipper Seller* coefficient would indicate that flippers sell houses for more than their expected value.

$$\log(p_{it}) = \alpha_1 yq_t + \alpha_2 id_i + \beta_{1k} b_{kit} + \beta_{2k} s_{kit} + \varepsilon_{it} \quad (3.2)$$

In equation (3.2), b_{kit} is a dummy for if the buyer is a flipper of type k and s_{kit} is a dummy equaling one if a flipper of type k is the seller. This estimated coefficients related to flipper activity will provide estimates of the discount that flippers get when buying and the premium they command when selling, provided that house quality is constant over time. If, however, flippers purchase houses and then invest heavily to improve them before putting them back on the market, these parameter estimates will be biased. In particular, we would expect β_{1k} to be negative because the true house quality in this period would be less than the estimated quality. Similarly, β_{2k} would likely be positive because the true quality in this period would be greater than the quality estimated. The researcher may, therefore, infer that flippers are buying

at a discount and selling at a premium when they are simply investing more than the average homeowner.

Because of this concern, we adapt this framework to control for the possibility of unobserved investment in the property by the flipper. To do so, we add an additional term to the regression:

$$\log(p_{it}) = \alpha_1 yq_t + \alpha_2 id_i + \beta_{1k} b_{kit} + \beta_{2k} s_{kit} + \beta_{3k} a_{kit} + \varepsilon_{it}. \quad (3.3)$$

Here we introduce a_{kit} , which is equal to one if, in any previous period, we see a flipper of type k purchase house i . This variable, therefore, controls for any improvements made by the flipper that extend beyond average homeowner investment since β_{3k} captures the change in house quality between when the flipper purchased and sold the home.

In any repeat sales framework, a property must sell at least twice in order for that house to be useful in identifying the underlying pattern of price appreciation in the market, i.e., be helpful in identifying any coefficients other than the corresponding house fixed effect. As in the standard repeat sales framework, all houses that sell at least twice will be useful in identifying the time series of market appreciation α_1 in equation 3.3. The identification of the coefficients corresponding to the sources of flipper returns and investment, $\beta_{1k}-\beta_{3k}$, however, require homes to sell at least four times, with at least one non-flipper to non-flipper transaction before and after a flipper buys and sells the house.

Figure 3.4 provides a visual illustration of how this structure controls for unobserved investment. In particular, consider a house that sells at four transaction times: A, B, C and D. At A both transacting parties are non-flippers. At B the house is sold to a flipper by the non-flipper. At C the flipper sells the house to a non-flipper. At D it is sold to a non-flipper by the non-flipper. The observation before the flipper buys is used, in effect, to identify the original house quality and

the observation after the flipper sells is used to identify the new house quality.

The left figure shows a flipper who buys below market price in period B and is able to sell above market price in C without making any improvements. In the right figure, the flipper makes improvements as can be seen by the fact that p_D continues to stay above its expected price, conditional on p_A . If we did not account for this improvement, it would appear that the flipper sold the house for above market value when in fact he sold it for exactly market value.

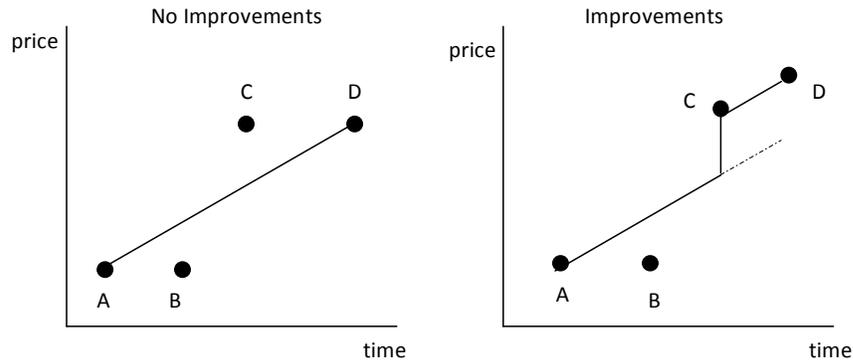


FIGURE 3.4: Identification strategy: Left: A case where the flipper did not make improvements between periods B and C. Right: An instance where the flipper did make improvements between B and C.

Several important features of this research design are worth noting. First, our estimates of the sources of flipper returns will be based on houses that have sold at least four times during the sample period and fit this ABCD structure. This means that, by construction, the period of time that the previous owner held a property before selling to a flipper is limited (as the sale at point A must be within the study period). This excludes a set of houses that may have been neglected over a long period of time by an owner (i.e., “fixer-uppers”) from contributing to our estimates of the sources of flipper returns.¹³ While flippers, especially those seeking to make

¹³ In fact a comparison of the housing attributes of homes that meet the ABCD structure reveals considerably less heterogeneity in the houses that flippers purchase versus the average homes that

significant physical improvements, may in fact target such homes for purchase, they will not generally be the ones that identify the sources of returns given our research design.¹⁴

A related concern is that flipper improvements may be underestimated if these improvements depreciate significantly under the care of the next home-owner. Of course, once again by construction, the length of time between when a flipper sells the house and when the house is re-sold by the subsequent buyer is limited by the fact that the sale at point D needs to take place within the study period. This provides a limited window for any physical improvement made by the flipper to have depreciated between points C and D. In robustness checks below, we also report results for a sub-sample that restricts the window of time between A and B and/or between C and D to less than three years.

In the analysis that follows, we report results for two slight adjustments to the specification shown in equation 3.3. First, we include a series of dummy variables for how many times we have seen a given property previously transacted in the study period. In general, sellers make some home improvements at the time of a sale so that a house will show well. Thus, we include these additional sales number dummy variables in order to make sure that we do not systematically overstate the performance of homes that meet the ABCD structure simply because they sell at least four times during the study period.

Secondly, as we show below, flippers (especially experienced flippers) tend to purchase homes that are slightly older and smaller than the average homes that are

sell in the market as a whole. The average year built of the homes purchased by Flipper 4's increases from 1949 to 1956, for example, when the sample is limited to just homes that meet the ABCD structure.

¹⁴ For the analysis of the sources of flipper returns (but, importantly, not the counts presented throughout the paper), we drop any purchases from banks or firms that might be associated with a foreclosure. We do this because of concerns that these homes may have been systematically run-down by the previously owners or vandalized, leading to large real declines in house quality between sales at points A and B, even if the time period between points A and B is short.

sold in the market. Therefore, to ensure that we are comparing apples to apples, we report results for a second specification of equation 3.3 that interacts the three key flipper variables with de-meaned measures of housing attributes, reporting the flipper coefficients at the mean attributes of the homes sold in the study period. This ensures that all comparisons of sources of returns are done for the same type of property, even though flippers with different levels of experience purchase properties that are a little heterogeneous.

Finally, it is worth stressing that while only flipped houses that sell at least four times and meet the ABCD structure will be helpful in identifying the three key flipper coefficients in equation 3.3, all of the counts presented in the paper are based on the full set of homes purchased by flippers. This is important because the set of homes that fit the ABCD structure will systematically result in a flipper purchase and sale closer to the middle of the study period (so that at least one sale can occur before and after the flipper's holding period).

3.5 The Sources of Flipper Returns - Baseline Results

3.5.1 *Flippers' returns*

We now provide estimates of the sources flippers' returns using the research design above. Our baseline results are presented in Table 3.4. To be included in the time period, the flipper must purchase the house within the time period mentioned. We exclude the end of the sample from these regressions because we need to follow the property for at least two years from the purchase date to estimate its return. By limiting the sample from 1992 to 2005, we also can easily split it into cold (1992-1998) and hot (1999-2005) periods of equal length. As mentioned above, for each sample period, results are presented for a basic specification and for one that interacts the key flipper variables with de-meaned housing attributes to ensure that the estimates are reported for comparable houses.

Table 3.4: Regression coefficients for all flippers: Standard errors in parentheses. Interacting house characteristics indicates that the mean house characteristics for the sample are subtracted from individual house characteristics and these values are interacted with the flipper dummies. The dummy for 5th sales or greater were omitted.

	(1)	(2)	(3)	(4)	(5)	(6)
Flipper Buyer	-0.058 (0.003)	-0.053 (0.003)	-0.129 (0.014)	-0.107 (0.013)	-0.035 (0.003)	-0.036 (0.003)
Flipper Seller	0.054 (0.003)	0.054 (0.003)	0.087 (0.006)	0.063 (0.007)	0.045 (0.003)	0.048 (0.003)
Flipper Investment	0.029 (0.004)	0.006 (0.004)	0.070 (0.014)	0.040 (0.013)	0.013 (0.004)	-0.008 (0.004)
First Sale	-0.132 (0.003)	-0.130 (0.003)	-0.141 (0.003)	-0.140 (0.003)	-0.144 (0.003)	-0.142 (0.003)
Second Sale	-0.075 (0.003)	-0.074 (0.003)	-0.081 (0.003)	-0.081 (0.003)	-0.082 (0.003)	-0.082 (0.003)
Third Sale	-0.042 (0.002)	-0.041 (0.002)	-0.047 (0.002)	-0.047 (0.002)	-0.047 (0.002)	-0.047 (0.002)
Fourth Sale	-0.020 (0.002)	-0.020 (0.002)	-0.023 (0.002)	-0.023 (0.002)	-0.023 (0.002)	-0.023 (0.002)
Interact House Characteristics?	No	Yes	No	Yes	No	Yes
Time Period	All Years	All Years	1992-1998	1992-1998	1999-2005	1999-2005
N	2,187,081	2,187,081	2,187,081	2,187,081	2,187,081	2,187,081
R ²	0.947	0.948	0.947	0.947	0.947	0.947

Controlling for unobserved investment, the estimates reported in the first column of Table 3.4 imply that flippers purchase homes at a discount of about 5.8% over the full sample period. That is, they purchase the house for approximately 6% less than its expected market price. Flippers also earn a premium of 5.4% when they sell the property (after controlling for investment). When the mean-differenced value of house characteristics are interacted with the flipper dummies to account for potential differences in the types of homes purchased by flippers, the magnitude of these coefficients changes very little, as shown in column (2).

Specifications (3) and (4) restrict the sample period to the cold market period (1992-1998), which was characterized by lower transaction volume and declining or flat housing prices. In general, flippers purchase homes at a much steeper discount and sell at a greater premium during this period. This is consistent with the idea

that flippers need to make their return by operating as middlemen during the cold market period, buying low and selling at a premium, relative to the average sales price in the market at the time.

Specifications (5) and (6) restrict the sample to the hot market period (1999-2005) in which prices were increasing rapidly and sales volume was much greater. In hot market conditions, by contrast, flippers have the potential to make returns by purchasing houses at times and in locations where expected market appreciation is high. Thus, as the parameter estimates in Table 3.4 show, flippers on average do not get a particularly low price when buying or a particularly high price when selling during this period.

The coefficients on *Flipper Investment* reported in the first column suggest that flippers are not investing much more than three percent of a house's value. This number falls to less than one percent in column (2), which reports results at mean house characteristics. Taken together, the results presented in columns (1) and (2) that flippers make substantial investments for houses that are especially old (and presumably in poor condition). The coefficients related to the order of sale reported in the lower half of the table make clear, however, this result may mask result on flipper investment may understate, to some extent the improvements that flippers make. These coefficients show a clear monotonic pattern of improvements, with all houses that sell multiple times typically selling at an increasing premium relative to market prices on later sales. Houses that sell four or five times, which flipped homes are more likely to be, typically generate a premium of upwards of 10-15 percent higher than the expected market price. Thus, some of the investment that flippers make in the properties that they buy and re-sell quickly is being captured by the inclusion of these control variables.¹⁵

¹⁵ In fact, the estimated coefficients corresponding to flipper investment in Table 3.4 are about 3 percentage points higher when the order of sale dummy variables are excluded from the analysis.

We now investigate the differential sources of returns across flipper experience levels, using the same four categories defined above in Figure 3.3. Table 3.5 presents parameter estimates for a set of specifications that correspond directly to those reported in Table 3.4 but that allow each flipper coefficient to vary by flipper experience. The heterogeneity in the sources of returns by flippers with different levels of experiences is immediately obvious. Looking across flipper types, it is clear that while all flippers buy relatively cheaply, more experienced flippers buy at a deep discount relative to expected market prices. For the sample period as a whole, Flipper 4's get a discount at purchase of approximately 20 percent and this discount is well over 30 percent in the cold market period. Steep discounts at the time of purchase are consistent with these experienced flippers operating as middlemen, buying very low at the purchase and operating during any market conditions. Inexperienced flippers, on the other hand, generally do not buy at much of a discount, especially in hot market conditions. This, again, is consistent with the idea that they are generally seeking profit as speculators rather than middlemen. None of the flipper types is associated with more than 1-2 percentage points worth of physical improvements in properties with mean housing attributes, over and above what is captured by the fact that houses bought and sold by flippers sell more often than other properties.¹⁶

Using the results from the estimates of the specifications reported in Table 3.5, which also include return measures of the Los Angeles house price index, we can report the source of a flipper's return for each flipper type: breaking this into the fraction that stems from buying cheaply, selling high, and simply earning the market return during the holding period. These results are in Table 3.6. We include estimates of flipper rates of return based on time held, market growth, and the residuals. Again, it is important to emphasize, that these estimates of sources do

¹⁶ The sale order dummy variables are included in the specifications reported in Table 3.5 but the parameter estimates (which are similar to those reported in Table 3.4 are not reported for exposition convenience).

Table 3.5: Regression coefficients by flipper type: Standard errors in parentheses. Interacting house characteristics indicates that the mean house characteristics for the sample are subtracted from individual house characteristics and these values are interacted with the flipper dummies.

	(1)	(2)	(3)	(4)	(5)	(6)
Flipper 1 Buyer	-0.035 (0.003)	-0.034 (0.003)	-0.062 (0.016)	-0.058 (0.015)	-0.020 (0.003)	-0.023 (0.003)
Flipper 2 Buyer	-0.075 (0.007)	-0.070 (0.007)	-0.175 (0.031)	-0.160 (0.031)	-0.050 (0.008)	-0.051 (0.008)
Flipper 3 Buyer	-0.123 (0.013)	-0.129 (0.012)	-0.188 (0.052)	-0.232 (0.060)	-0.090 (0.014)	-0.096 (0.014)
Flipper 4 Buyer	-0.210 (0.015)	-0.181 (0.018)	-0.330 (0.051)	-0.334 (0.071)	-0.153 (0.016)	-0.134 (0.019)
Flipper 1 Seller	0.049 (0.003)	0.051 (0.003)	0.061 (0.008)	0.053 (0.008)	0.045 (0.004)	0.047 (0.004)
Flipper 2 Seller	0.060 (0.007)	0.060 (0.008)	0.098 (0.016)	0.073 (0.021)	0.048 (0.008)	0.055 (0.008)
Flipper 3 Seller	0.072 (0.011)	0.050 (0.014)	0.137 (0.018)	0.098 (0.026)	0.046 (0.014)	0.034 (0.017)
Flipper 4 Seller	0.090 (0.010)	0.055 (0.014)	0.137 (0.013)	0.072 (0.026)	0.054 (0.013)	0.050 (0.016)
Flipper 1 Investment	0.019 (0.004)	0.002 (0.004)	0.083 (0.017)	0.046 (0.016)	0.005 (0.004)	-0.010 (0.004)
Flipper 2 Investment	0.042 (0.009)	0.012 (0.009)	0.049 (0.032)	0.027 (0.031)	0.028 (0.010)	-0.001 (0.009)
Flipper 3 Investment	0.047 (0.015)	0.016 (0.016)	0.080 (0.053)	0.041 (0.060)	0.037 (0.017)	0.003 (0.018)
Flipper 4 Investment	0.037 (0.016)	0.019 (0.019)	-0.022 (0.051)	-0.048 (0.070)	0.043 (0.018)	0.019 (0.021)
Interact House Characteristics?	No	Yes	No	Yes	No	Yes
Time Period	All Years	All Years	1992-1998	1992-1998	1999-2005	1999-2005
N	2,187,081	2,187,081	2,187,081	2,187,081	2,187,081	2,187,081
R ²	0.948	0.948	0.947	0.947	0.947	0.947

not account for flippers' transaction or holding costs, meaning actual profits are almost certainly much smaller. Nonetheless, the estimated average rates of returns earned by all types of flippers during our sample period suggest that they, in fact, can operate quite profitably in the market.

Table 3.6 highlights the distinction between flipper types and provides strong evidence that some flippers act as speculators while others operate as middlemen. First, there is a large disparity in time held. Flipper 4's quickly resell their houses

Table 3.6: Sources of return by flipper type: The table shows the sources of returns by flipper type. The discounts, premiums, and market growth are calculated from specification (2) of Table 3.5 and quarters held is simply the mean number of quarters held. The nominal rate of return is generated by dividing the mean total return (premium - discount + market growth) by the mean years held.

	Nominal Rate of Return	Buyer Discount	Seller Premium	Market Growth	Quarters Held	N
Flipper 1	0.234	-0.034	0.051	0.150	4.01	25,181
Flipper 2	0.294	-0.070	0.060	0.109	3.25	5,678
Flipper 3	0.374	-0.129	0.050	0.089	2.86	2,322
Flipper 4	0.531	-0.181	0.055	0.053	2.17	2,596

while flipper 1's hold them almost twice as long. Second, flipper 1's do not buy at an especially low price and, as a result, their (nominal) rate of return is primarily driven by overall market growth: 64% of their return stems from market growth. Flipper 4's, on the other hand, earn most of their return by buying at prices below average market prices (purchasing cheaply generates 63% of their return) and quickly reselling so that only 18% of their return stems from overall market growth. Taken together, the evidence on purchase activity, holding times, and sources of returns paints a very consistent picture: experienced flippers generally act as middlemen and inexperienced flippers as speculators in the Los Angeles housing market over our study period.

3.6 Robustness

In this section, we examine the robustness of the results presented above to a number of the assumptions that underlie our analysis. In so doing, we also address a number of additional questions regarding the behavior of flippers and foreshadow the analysis of the next section, which explores how middlemen and speculators target particular locations for their purchases.

3.6.1 Do Flippers Sell Winners and Hold Losers?

In the results presented in Section 3.5, we examined the sources of returns for houses that were re-sold in less than two years. Of course, the timing of the decision to re-sell the property is an endogenous choice made by the investor likely influenced by the appreciation of the property and the cost of capital. By limiting the sample to only those homes that were re-sold in the first two years, we may be inadvertently focusing on a very select sample of homes that performed very well in terms of market appreciation. As a simple check on the sensitivity of our results to the definition of flipped homes as those sold within two years, we consider the effect of adjusting this time period. In Table 3.7, the first specification is the baseline, which uses the estimates from specification (2) in Table 3.4. Specifications (2)-(4) in Table 3.7 vary the amount of time required from eighteen months to four years.

Table 3.7: Robustness checks: Standard errors in parentheses. Specification (1) is set to specification (2) from Table 3.4. Specification (2) is similar to the baseline but changes the required holding time from 2 years to 18 months. Specifications (3) and (4) increases the maximum holding time to 3 and 4 years, respectively. Specification (5) requires that the time between transactions A and B is less than 3 years. Specification (6) requires that the time between transactions C and D is less than 3 years. Specification (7) requires the conditions in both (6) and (5).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Flipper 1 Buyer	-0.034 (0.003)	-0.046 (0.003)	-0.031 (0.003)	-0.030 (0.003)	-0.007 (0.004)	-0.020 (0.005)	0.003 (0.007)
Flipper 2 Buyer	-0.070 (0.007)	-0.080 (0.008)	-0.065 (0.007)	-0.060 (0.007)	-0.044 (0.012)	-0.045 (0.012)	-0.012 (0.015)
Flipper 3 Buyer	-0.129 (0.012)	-0.143 (0.013)	-0.117 (0.011)	-0.111 (0.011)	-0.087 (0.018)	-0.106 (0.026)	-0.061 (0.031)
Flipper 4 Buyer	-0.181 (0.018)	-0.199 (0.018)	-0.173 (0.017)	-0.163 (0.016)	-0.165 (0.029)	-0.177 (0.035)	-0.172 (0.053)
Flipper 1 Seller	0.051 (0.003)	0.054 (0.004)	0.049 (0.003)	0.048 (0.003)	0.051 (0.007)	0.052 (0.004)	0.054 (0.007)
Flipper 2 Seller	0.060 (0.008)	0.063 (0.008)	0.058 (0.007)	0.056 (0.007)	0.083 (0.016)	0.063 (0.010)	0.082 (0.019)
Flipper 3 Seller	0.050 (0.014)	0.057 (0.015)	0.049 (0.013)	0.048 (0.013)	0.046 (0.031)	0.058 (0.017)	0.067 (0.039)
Flipper 4 Seller	0.055 (0.014)	0.054 (0.014)	0.056 (0.013)	0.054 (0.013)	0.037 (0.028)	0.086 (0.019)	0.059 (0.034)
Flipper 1 Investment	0.002 (0.004)	0.000 (0.004)	0.003 (0.004)	0.003 (0.004)	-0.012 (0.007)	-0.002 (0.005)	-0.008 (0.008)
Flipper 2 Investment	0.012 (0.009)	0.014 (0.010)	0.010 (0.008)	0.012 (0.008)	-0.027 (0.015)	0.004 (0.011)	-0.005 (0.017)
Flipper 3 Investment	0.016 (0.016)	0.013 (0.017)	0.018 (0.015)	0.017 (0.014)	0.006 (0.030)	0.013 (0.024)	-0.005 (0.035)
Flipper 4 Investment	0.019 (0.019)	0.015 (0.020)	0.011 (0.018)	0.011 (0.018)	-0.005 (0.032)	0.001 (0.029)	-0.026 (0.041)
Robustness Check	Baseline	Flip in < 1.5 Years	Flip in < 3 Years	Flip in < 4 Years	A to B < 3 Years	C to D < 3 Years	Specifications (5) and (6)
N	2,187,081	2,187,081	2,187,081	2,187,081	2,187,081	2,187,081	2,187,081

The results presented in columns (1)-(4) of Table 3.7 reveal that the conclusions drawn from the baseline results regarding the sources of returns - that more experienced flippers earn a large fraction of their return by buying at especially low prices, while less experienced flippers do not - are not at all sensitive to the choice of threshold holding time. When the threshold is set to four years instead of two, for example, the estimated discount relative to expected market price that Flipper 4's get at purchase is 16.4 percent versus 18.1 percent, while the estimate for Flipper 1's remains very low 3.0 percent versus 3.4 percent.

3.6.2 Are Results Driven by Flippers Buying Fixer-Uppers?

As we discussed in detail above, a broad challenge in examining the sources of flipper returns is the possibility that flippers invest significant amounts of money to improve properties, investment that is unobserved to the researcher. If, for example, flippers purchase fixer-uppers at what might appear to be below market prices and then bring them back up to standard market conditions, we might improperly infer that they were making substantial returns by buying at low prices relative to market.

Several aspects of our baseline analysis have been designed to minimize this concern. In particular, our focus on the ABCD structure for identifying the sources of flipper returns, not only provides a way to estimate the amount of unobserved investment that flippers put into properties (versus typical home-sellers) but also naturally limits the identification of returns to properties that were transacted within a reasonably small period both before and after the flipper bought and sold the property.

Specifications (5)-(6) in Table 3.7 take the logic of this one step further, limiting the time between sales at point A and B, and C and D, respectively, to less than three years. Specification (7) combines these restrictions. By limiting the times between A and B and C and D, specifications (5)-(7) not only address the potential concern that flippers may buy homes which have been run down by their previous owners,

but also that flippers invest in houses, only to have their investment depreciate by period D.

The results presented in final three columns of Table 3.7 again strongly support the the conclusions drawn from the baseline results regarding the sources of returns for more and less experienced flippers, respectively. When the time between the transactions preceding and subsequent to the flippers holding of the property are both limited to three years, for example, the estimated discount relative to expected market price that Flipper 4's get at purchase is 17.2 percent versus 18.1 percent for the baseline case. In fact, the estimated discount at purchase for Flipper 1's falls all the way to zero, implying that these inexperienced flippers essentially purchase houses at expected market prices.

It is worth emphasizing that nothing in our analysis implies that flippers do not indeed often purchase fixer-uppers that could be physically improved in a profitable way. It is just that our research design ensures that such properties do not contribute to the identification of the sources of flipper returns.

3.6.3 Selective Survival - The Dynamic Pattern of Returns

Another potential concern with our baseline results is that our examination of the heterogeneity in flipper returns is not based on a time-invariant attribute of flippers but instead on their experience. At the outset of this discussion, it is important to keep in mind that we are not interested in identifying the effects of flipper experience per se. Instead, as it turns out, cutting the data by experience revealed a striking difference in the typical patterns purchases, holding times, and sources of returns for experienced versus inexperienced flippers that maps almost perfectly into the roles of middlemen and speculators, respectively. This does not imply, however, that some of the inexperienced flippers are middlemen just launching their careers or that some of the inexperienced speculators might not eventually become more experienced if

they survive in the profession.

One way to examine the dynamics of experience and flipper type is to examine the dynamic pattern of sources of returns for the more experienced flipper types. In particular, instead of just examining the sources of returns for all houses flipped by Flipper 4's, for example, we consider separately the sources of returns for the first three flipped houses that they flipped, the 4th-6th houses, 7th-10th houses, and all houses after their first ten, respectively. These results are presented in Table 3.8.

The results reveal a consistent pattern of sources of returns for the more experienced flippers observed in the data. On their first three houses flipped in the data Flipper 3's received a discount at purchase of 13.4 percent and premium at sale of 9.5 percent relative to expected market prices. For Flipper 4's these numbers were even higher, a 25 percent discount and 17 percent premium. The magnitude of these numbers reflects, of course, the fact that an experienced flipper's first three flipped homes were more likely to have occurred during the cold market period (1992-1998), during which flippers required larger margins on purchase and sales prices in a market with declining home prices. Taken as a whole, table 3.8 supports the notion that the more experienced flippers observed in our dataset have been acting in the economic role of middlemen throughout the sample period, while the ranks of inexperienced flippers is dominated by those pursuing a more speculative strategy in their limited careers in the market to date.

3.6.4 Neighborhood Targeting

For our baseline results, we estimated a single housing price index for the Los Angeles metropolitan market and used that to measure the rate of market appreciation during the holding times, as reported in Table 3.6. A concern with using a single aggregate price index for our analysis is that flippers might be able to identify and target submarkets or neighborhoods that appreciate faster than the metropolitan

Table 3.8: Flipper regression coefficients by flip number: Standard errors in parentheses. The rows correspond to the flipper type (total number of flips) and the columns correspond to the flip number. For example, Flipper 1's only have coefficients for 1 to 3 because the maximum number of flips for this type is 3.

Flip Number	Nominal Rate of Return 1 to 3	Buyer Discount 4 to 6	Seller Premium 7 to 10	Market Growth 11 and Up
Buyer Coefficients				
Flipper 1	-0.035 (0.003)			
Flipper 2	-0.080 (0.011)	-0.072 (0.009)		
Flipper 3	-0.134 (0.034)	-0.123 (0.018)	-0.113 (0.020)	
Flipper 4	-0.254 (0.074)	-0.208 (0.035)	-0.173 (0.034)	-0.216 (0.019)
Seller Coefficients				
Flipper 1	0.049 (0.003)			
Flipper 2	0.046 (0.010)	0.073 (0.010)		
Flipper 3	0.095 (0.021)	0.063 (0.015)	0.072 (0.024)	
Flipper 4	0.173 (0.032)	0.102 (0.022)	0.086 (0.021)	0.069 (0.013)
Investment Coefficients				
Flipper 1	0.019 (0.004)			
Flipper 2	0.056 (0.013)	0.027 (0.012)		
Flipper 3	0.064 (0.036)	0.042 (0.021)	0.039 (0.028)	
Flipper 4	-0.007 (0.073)	0.044 (0.036)	0.059 (0.035)	0.042 (0.021)

area as a whole. Not only might this lead to an understatement of market appreciation as a source of flipper returns, but it might lead to an overstatement of the premium that they receive at sale, essentially misinterpreting the reason that the flipper outperforms the LA market as a whole – interpreting it as a premium at sale rather than faster than average neighborhood-level appreciation.

To address this concern, we divide the Los Angeles metropolitan area into forty submarkets based on California state assembly lower voting district. We use voting districts because they are both large enough to have a reasonable number of flipper observations in each district, and small enough to characterize a meaningful submarket. A map of these voting districts is given in Figure 3.5. It is then straightforward to estimate an extended version of eqnReg1a that allows the coefficients on the time dummies to vary by submarket - i.e., to estimate separate price indices for each submarket. The results of this analysis are presented in Table 3.9 which has the same format as Table 3.6.

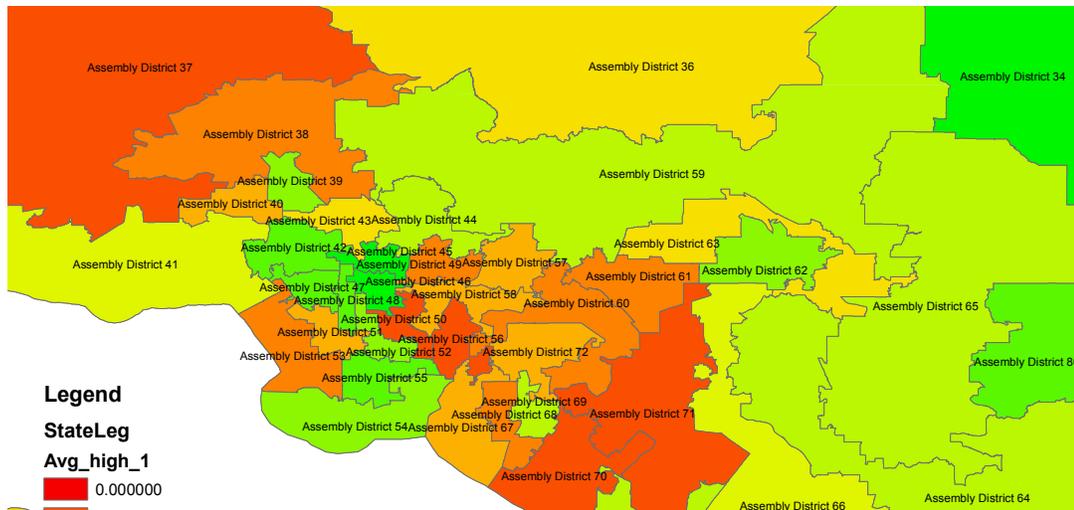


FIGURE 3.5: The map of voting districts used as submarkets in our analysis.

The results presented in Table 3.9 strengthen our qualitative conclusions for both experienced and inexperienced flippers, as inexperienced flippers target submarkets

Table 3.9: Source of returns by flipper type at neighborhood level: The table shows the sources of returns by flipper type. This table has the same format as Table 3.6 but is based on a specification that estimates a separate price index for of 40 submarkets of the LA metro area defined by state assembly lower voting districts.

	Nominal Rate of Return	Buyer Discount	Seller Premium	Market Growth	Quarters Held	N
Flipper 1	0.240	-0.037	0.047	0.156	4.01	25,114
Flipper 2	0.310	-0.079	0.058	0.115	3.25	5,672
Flipper 3	0.407	-0.125	0.071	0.094	2.86	2,313
Flipper 4	0.640	-0.209	0.082	0.056	2.17	2,596

experiencing faster than average appreciation and experience flippers actual target submarkets that are appreciating slower than average. Overall, the estimated total returns, as well as the source of returns remains similar in Table 3.9 to the case where we treat the Los Angeles metropolitan area as one market.

3.7 Submarket Level Results

In this section we provide corroborating evidence from the submarket-level that flippers operate as both speculators and middlemen. In addition, we present suggestive evidence that speculative flippers are associated with greater price instability at the submarket level.¹⁷

For exposition sake, for the rest of the paper we group flippers into two categories based on experience. Experienced flippers, or middlemen, are those who engage in 4 or more flips over our sample (in the language above, these are flippers 2-4's) and inexperienced flippers, or speculators, are those who flip two to three times during our sample (flipper 1's). While these definitions are somewhat arbitrary, and there is certainly not a perfect relationship between experience and whether flippers operate as middlemen or speculators, this appears to divide flippers into two categories of intermediaries that are following very distinct strategies for earning returns in the

¹⁷ This finding echoes that of Greenwood and Nagel (2009) who present evidence that inexperienced, speculative mutual fund managers are associated with price instability in that market.

market. The results presented below are robust to altering the thresholds of this dichotomy.

3.7.1 Which Submarkets are Targeted?

For this section of the paper, we continue to use the forty submarkets based on California state assembly lower voting district introduced in the previous section. We relate flipper activity to submarket price appreciation, calculated using a repeat sales index for each neighborhood. In all cases, appreciation is measured over a year. Table 3.10 associates flipping purchases with submarket-level price appreciation.¹⁸ The regressions include submarket level and quarter fixed effects. The specifications differ in the lag between flipping activity and price changes. The first two columns show that the greater holdings by inexperienced, speculative flippers are associated with above average rates of price appreciation over the following two years, while the last three columns show that these short term gains are followed by below average returns (mean reversion) over the following three years.

Experienced flippers (largely middlemen), on the other hand, operate in areas where prices are not rising as quickly as the rest of the metropolitan area. As discussed in Section 2, since middlemen earn their returns by finding “good” deals when purchasing and selling for high prices relative to the market rate, it is not surprising to find them operating in hot and cold submarket just as they operate in hot and cold portions of the housing cycle.

Given the strong association between speculative activity and the amplification of local housing booms and busts, it is useful to examine whether speculators are

¹⁸ We obtain similar results if lagged flipper holdings are used instead. We present evidence of this below where holdings are constructed as follows:

$$h_{itn} = h_{it-1n} + b_{itn} - s_{itn} \tag{3.4}$$

where h_{itn} are the holdings of all flippers of type i in quarter t in submarket n , b_{itn} are the purchases by flippers in submarket n in the quarter, and s_{itn} are the sales by flippers in same neighborhood and quarter.

Table 3.10: Effect of flipper activity on neighborhood prices: Standard errors in parentheses are clustered by submarket. All specifications include submarket and quarter fixed effects.

	(1)	(2)	(3)	(4)	(5)
Flipper Houses Bought:					
Inexperienced	1.466 (0.180)	1.539 (0.250)	0.451 (0.246)	-0.803 (0.268)	-0.674 (0.307)
Experienced	-0.304 (0.166)	-0.335 (0.208)	-0.069 (0.184)	0.142 (0.197)	0.889 (0.312)
Lag Time (quarters)	4	4	4	4	4
Dependent Variable	(t+1)-t	(t+2)-(t+1)	(t+3)-(t+2)	(t+4)-(t+3)	(t+5)-(t+4)
N	2,840	2,680	2,520	2,360	2,200
R ²	0.922	0.923	0.920	0.921	0.922

Table 3.11: Effect of earlier appreciation on neighborhood prices: Standard errors in parentheses are clustered by submarket. All specifications include submarket and quarter fixed effects.

	(1)	(2)	(3)	(4)	(5)
Lagged Appreciation	0.551 (0.063)	0.035 (0.055)	-0.294 (0.069)	-0.448 (0.076)	-0.369 (0.109)
Lag Time (quarters)	4	4	4	4	4
Dependent Variable	(t+1)-t	(t+2)-(t+1)	(t+3)-(t+2)	(t+4)-(t+3)	(t+5)-(t+4)
N	2,820	2,660	2,500	2,340	2,180
R ²	0.891	0.907	0.907	0.905	0.906

Table 3.12: Effect of flipper activity and earlier appreciation on neighborhood prices: Standard errors in parentheses are clustered by submarket. All specifications include submarket and quarter fixed effects.

	(1)	(2)	(3)	(4)	(5)
Lagged Appreciation	0.583 (0.068)	0.023 (0.063)	-0.335 (0.079)	-0.480 (0.084)	-0.327 (0.115)
Flipper Houses Bought:					
Inexperienced	1.466 (0.170)	1.513 (0.266)	0.449 (0.262)	-0.776 (0.222)	-0.727 (0.279)
Experienced	0.377 (0.183)	-0.308 (0.248)	-0.498 (0.208)	-0.391 (0.190)	0.585 (0.320)
Lag Time (quarters)	4	4	4	4	4
Dependent Variable	(t+1)-t	(t+2)-(t+1)	(t+3)-(t+2)	(t+4)-(t+3)	(t+5)-(t+4)
N	2,820	2,660	2,500	2,340	2,180
R ²	0.892	0.910	0.907	0.905	0.909

responding to recent submarket level price changes or might base their decisions on additional information. To explore this possibility, we begin by reconfirming, at the neighborhood level, the previous finding in the literature that lagged appreciation strongly predicts future price appreciation. We follow the technique suggested in Case and Shiller (1989) for dealing with measurement error in the price index by (i) splitting our sample in two, (ii) generating estimates of appreciation independently for each sample, and (iii) instrumenting for lagged appreciation in one sample with the estimate of lagged appreciation from the other sample.

The results presented in Table 3.11, which control for quarter dummies and submarket fixed effects, establish positive short-term persistence and long-term mean reversion in house price appreciation at the submarket level. In this way, a sharp uptick in lagged appreciation is a clear predictor of above-average short-term returns at the submarket level. To our knowledge, the fact that, even at the neighborhood-level, lagged appreciation strongly predicts future price appreciation is undocumented elsewhere in the literature.

Table 3.12, includes controls for both lagged appreciation and speculative activity. The results reveal that both factors continue to be strongly associated with increased price volatility at the submarket level. The positive correlation between these two measures, and the resulting diminished magnitudes of the coefficients on flipper activity in Table 3.12 versus Table 3.10 imply that speculators are engaging to some extent in positive feedback trading.

There are at least three explanations for a remaining positive relationship between speculative activity and submarket appreciation (positive in the short-run, negative in the intermediate run) after controlling for lagged appreciation. First, the relationship between flipper activity and lagged appreciation may be non-linear, with flipper activity triggered only after expected appreciation reaches a minimum threshold. Second, flippers might be responding to other information (unobserved

in our dataset) that predicts above-average returns over short horizons and mean reversion in the slightly longer run. Third, flipper activity might have a causal effect on appreciation rates, contributing to the amplification of the local housing price boom-bust cycle over the next several years. While distinguishing among these explanations is beyond the scope of this paper, our submarket level analysis allows us to conclude that speculator activity is strongly associated with the amplification of house price volatility at the submarket level.

3.8 How Well-Informed Are Speculators?

The evidence presented so far establishes a set of market intermediaries that act as speculators who are (i) inexperienced, (ii) enter the market as prices begin to rise, (iii) specifically target locations that subsequently experience rapid short-term price appreciation and (iv) earn their returns only by holding houses when prices rise since they are unable to (a) buy at great discounts or (b) sell their assets at a premium. This behavior is consistent with modern finance theory which calls into question the central tenet of the efficient markets hypothesis (e.g. Friedman (1953) and Fama (1955)) that rational investors should always “attack a bubble.” These more recent papers suggest instead that rational speculators may instead want to jump on the bandwagon with the noise traders, ride the bubble on the way up and sell out as the market nears or reaches its peak (DeLong et al. (1990)).

In Section 3.7, we established that the purchase behavior of speculators is strongly associated with neighborhood-level housing bubbles - i.e., neighborhoods that experience especially strong positive appreciation in the short term (1-2 years), but below average appreciation 3-5 years out.. What remained difficult to ascertain in that analysis was whether speculators actually have superior information, and are therefore timing both their purchases and sales optimally, or whether these inexperienced speculators are simply chasing trends themselves without any special access to su-

perior information. If they are operating according to the latter capacity, they may, in fact, be causing these short-term speculative bubbles, rather than using superior information to take advantage of bubbles that they can foresee to some extent.

While some might take the inexperience of the speculators in the data as *prima facie* evidence that they are not especially well informed, we offer a more formal analysis of their behavior in this section.¹⁹ In particular, we present evidence on the timing of purchases and sales by the speculators in our sample as the Los Angeles market neared, hit, and then went over its peak in 2006. For this analysis, we construct a sample of all homes purchased by individuals identified as flippers in our sample, regardless of whether they sold the homes within two years. We continue to divide flippers into speculators and middlemen based on experience, focusing here on the activities of speculators near the peak of the Los Angeles market.

Figure 3.6 plots the fraction of speculators' purchases from two years prior to time t that they continue to hold at time t . For homes purchased in the early and middle periods of the housing boom between 1999-2006, speculators generally unloaded about 55% of their holdings within two years, holding only 45% of their properties for more than two years. Looking two years after the market began to peak in 2005-2006, however, reveals that by late 2007, speculators were stuck holding a much higher fraction of the homes that they purchased as the market peaked. In fact, speculators were stuck holding over 60-70% of the homes purchased near the end of the boom period two years later. Given the rapid rates of price depreciation over this period, it is obvious that such speculators took substantial losses on these properties.²⁰

¹⁹ Other researchers (e.g. Greenwood and Nagel (2009)) have found that inexperienced traders engage in the type of trend chasing behavior exhibited here. There is also a multitude of evidence of this phenomenon from lab and retail investor survey settings. See, for example, Smith et al. (1988), Haruvy et al. (2007) or Vissing-Jorgensen (2003).

²⁰ Our findings that inexperienced speculators in this market hold properties too long as the market peaked in 2006 is really the mirror image of the findings reported in Brunnermeier and Nagel (2004),

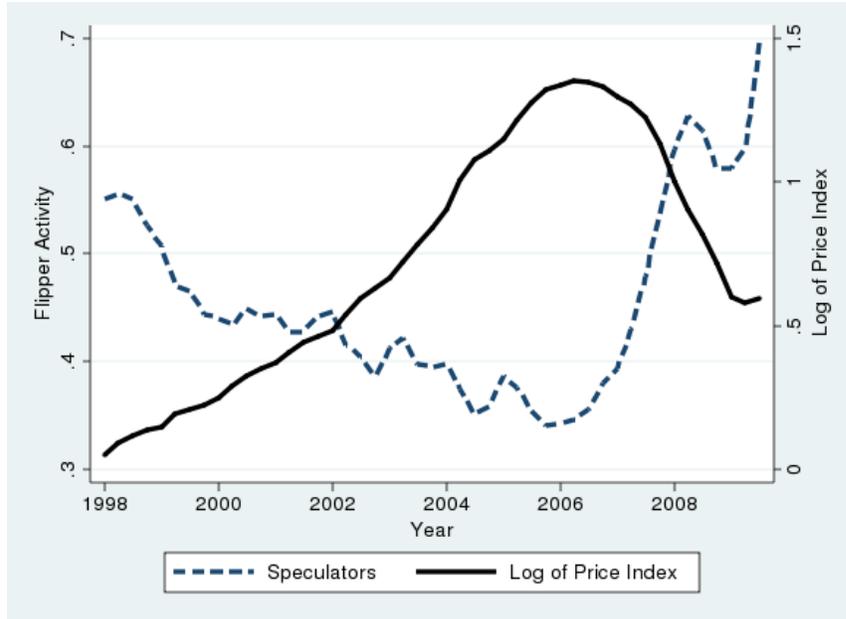


FIGURE 3.6: Houses held by speculators: The dashed line maps the fraction of houses purchased two years earlier that are still held by speculators for each quarter. The solid line gives the price index as estimated in the data.

Of course, if speculators had curtailed their activity in anticipation of the coming peak, getting stuck holding a relatively high fraction of their peak purchases two years later would not necessarily be that devastating – i.e., if the number of homes affected was small. As it turns out, as shown in Figure 3.7, speculators continued to purchase at a high rate all the way up to the peak in the market. During the housing boom, speculator activity had slowly increased from about 2% of all purchases in the late 1990s to consistently above 3% of purchases between 2004 and 2006. While speculator activity began to fall in 2006 it was not until the middle of 2007 that activity dropped to below 2% again.

To illustrate the clear sense that the fevered speculative activity went on far too long, Figure 3.7, also plots a measure of predicted 2-year appreciation rates based on a model of predicted appreciation similar to those reported in Table 3.11, but estimated

which demonstrated that more experienced traders, hedge fund managers, divested their holdings of assets with inflated prices just prior to prices falling in tech bubble.

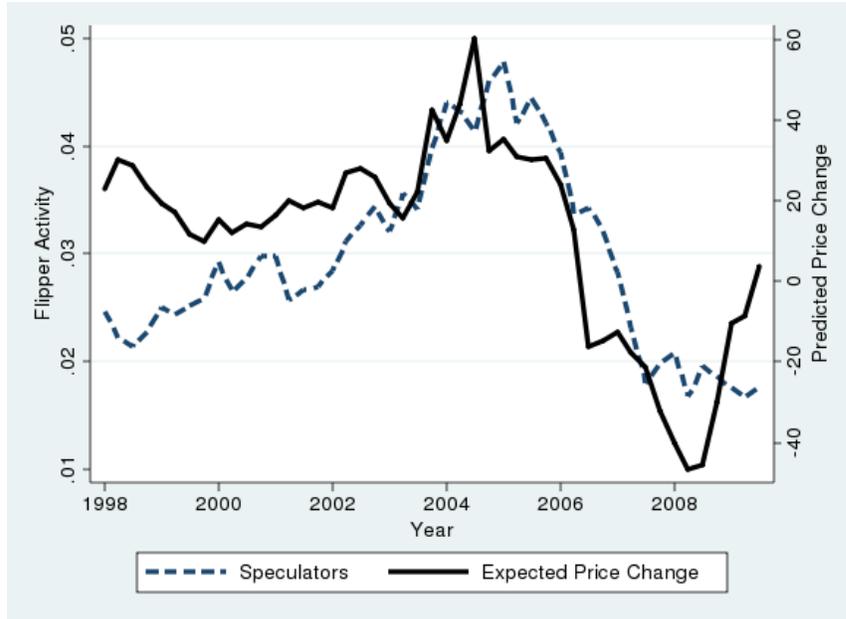


FIGURE 3.7: Speculator activity based on expected appreciation: The dashed line maps the percent of purchases made by speculators over time. The solid line gives the two year expected house appreciation based on lagged appreciation using the Federal Housing Finance Agency’s price index for the Los Angeles-Long Beach-Glendale MSA.

at the metropolitan level using FHFA price indices for over 400 metropolitan areas from 1975 to 2010. While predicted 2-year appreciation began to drop off quickly in early 2005, speculative purchases at rates above 3% continued until early 2006.

Looking at speculative behavior in the third quarter of 2006 provides a clear indication of this point. By the third quarter of 2006, the predicted rate of return over the next two years (based on lagged appreciation) had fallen from a high of 60% to negative 15%. Yet speculators continued to account for 3.2% of the purchases in the market that quarter. Over the next two years, speculators were only able to unload 42% of these purchases and thus took substantial losses on properties that well-informed agents would never have purchased in the first place.

Taken together, the evidence presented here regarding the timing of speculative activity near, at, and following the peak, strongly suggests that a large share of the

speculators operating in the Los Angeles market were not especially well-informed and, in fact, were likely simply chasing trends themselves, much like ordinary homeowners.

3.9 Conclusion

Making use of a large transactions database and a novel research design, this paper provides the first comprehensive study of intermediaries (middlemen and speculators) in the housing market: identifying (i) their activity, (ii) the sources of their returns, and (iii) the extent to which their activity is associated with local price dynamics. Our analysis for Los Angeles establishes that middlemen and speculators follow distinct strategies for when and where to buy and generate returns from almost completely distinct sources. Middlemen hold properties for very short periods of time and earn most of their return by buying houses relatively cheaply; they operate throughout booms and busts in the market and target all types of locations. By contrast, speculators earn almost their entire return through timing the market, operate only during boom times, and target submarkets with the highest expected price appreciation. Entry by speculative flippers is strongly associated with the short-term amplification of local housing price bubbles. And, given their inexperience flipping homes and apparent difficulty anticipating and reacting to the peak of the most recent housing boom, it seems likely that many of the speculators identified in the data may not be particularly well informed about market conditions.

This paper makes several important contributions to the literatures on housing and financial markets. Most directly, it expands our understanding of the microstructure of the housing market: establishing a number of new empirical facts about the activity of middlemen and speculators in the market that generally conform to the roles prescribed in economic theory. More generally, our ability to identify speculators in the data and analyze their strategies and impact on the market is relatively

rare in the wider empirical finance literature. While not completely conclusive, our findings suggest that (i) many of the speculators that entered the market during the recent housing boom may not be especially well informed about market fundamentals and (ii) speculators in this market have destabilizing effects on prices, leading to an amplification of local housing price cycles. While this increased volatility has important economic consequences, any policy remedies need also account for the welfare-enhancing role that flippers play by providing liquidity as middlemen and in improving the physical stock of housing in older neighborhoods.

Turnaways and Patient Composition: An Empirical Analysis of Hospitals

4.1 Introduction

Hospitals provide health care to millions of citizens and their work in the health industry comprises a major portion of total spending. While hospitals cannot continually lose money and stay in business, they also should not be treated as an ordinary firm that simply attempts to maximize lifetime profits. One factor influencing their behavior is that they consist of several ownership types including local government owned, nonprofit, and for-profit and they may derive utility from factors beyond profits, such as the number of patients treated. Additionally, the industry is regulated differently than most consumer industries. As a result, there have been many studies aiming to better understand hospital behavior.

My analysis adds to this literature by focusing on how a hospital's patient composition changes based on whether it is near capacity and may therefore have to turn patients away. This question is informative for two reasons. First, if the set of patients changes as a hospital becomes busier, it may inform policymakers about

what types of patients are most negatively affected by hospital congestion. Secondly, my analysis will inform researchers as to whether the patients turned away from a hospital are randomly drawn from the distribution of all patients seeking care at the hospital, or if facilities are more likely to turn away some types of patients than others.

To address this question, I introduce a theoretical model that predicts that as a hospital becomes busier, it is more likely to turn away less profitable patients whereas it will continue to admit more profitable patients until it has no more available beds. To test this model empirically, I analyze the flow of inpatients to hospitals and using a queuing model, estimate the probability in each period that a patient seeking admission would be turned away from the hospital if turnaways occurred when the hospital was at full capacity. To more specifically look at the hospital's patient composition, I use information on the patient's condition and race, payment type, and diagnosis related group (DRG).

I then run regressions that estimate the effect of a hospital's turnaway probability on patient composition and find that as a hospital's turnaway probability becomes larger, its share of mental health, black, and uninsured patients decreases. Each of these findings is consistent with my theoretical model and suggests that these patients are more likely to be turned away than other groups that are more profitable to a hospital. Additionally, I look at whether this behavior differs based on hospital ownership type and find that it is relatively constant across different ownership groups.

The topic of hospital turnaways and patient composition has policy implications as it is important to understand what types of patients are most affected by turnaways. Likewise, this analysis provides valuable information about hospital behavior as it demonstrates that a hospital's composition of patients changes with congestion suggesting that hospitals do not treat all patients similarly when dealing with

capacity constraints.

The rest of the paper is outlined as follows. Section 4.2 discusses the background literature in more detail. Section 4.3 introduces a theoretical model used to generate predictions of hospital behavior. Section 4.4 describes the data. Section 4.5 outlines the empirical model and the results are given in Section 4.6. Section 4.7 discusses potential extensions to this research. Section 4.8 concludes.

4.2 Background Literature

There has been a long line of literature characterizing hospital behavior. One of the major questions studied is how a hospital's ownership type affects its decisions and whether nonprofit and for-profit hospitals behave similarly. Sloan (2000) provides a detailed review of the literature comparing hospital behavior based on ownership type. Gowrisankaran et al. (2010) model hospital utility as a function of both profits and patients served where the relative value of each may vary depending on ownership type. With respect to patient composition, Horwitz (2005) finds that for-profit hospitals focus on providing profitable services whereas government hospitals disproportionately provide services that are less profitable to the hospital. Norton and Staiger (1994) conclude that when a for-profit and nonprofit hospital operate in the same market, they serve the same rate of uninsured patients; however for-profit hospitals tend to serve areas with above average insurance coverage.

A number of papers have looked at the relative profits of various types of patients based on payment source and diagnosis related group. Horwitz (2005) categorizes different types of hospital services as "relatively profitable," "relatively unprofitable," or "variable." In her categorization, she claims that psychiatric services are relatively unprofitable whereas cardiac services are profitable. I will make the same assumption in my analysis. The claim that cardiac procedures are profitable is consistent with the recent *New York Times* story on the for-profit hospital chain HCA, which performed

unnecessary cardiac procedures on patients to increase profits (Abelson and Creswell (2012)). Newhouse (1989) finds that patients with less profitable DRG's are more likely to be transferred between hospitals and to be admitted to a "hospital of last resort." As discussed in Johnson and Sataline (2010) and Decker (2012), Medicaid reimbursement is less profitable for hospitals than Medicare patients. And each group is less financially beneficial to hospitals than privately insured patients. The least profitable set of patients, however, are the uninsured who may only be able to pay a small fraction of their costs. When focusing on the least profitable patients, I will therefore use the uninsured.

Additionally, there is a literature focusing on hospital turnaways. Shonick (1970) describes a queuing model which estimates the probability that a patient is turned away because the facility has no empty beds. This methodology is applied in numerous papers including Joskow (1980) and state health planning boards. There have also been a number of theoretical papers looking at the predicted turnaways based on projected occupancies using simulations including Bagust et al. (1999). Hsia et al. (2012) finds hospitals that turn away a larger fraction of patients also serve larger minority populations suggesting that there may be distributional concerns associated with hospital turnaways.

This paper aims to build off of the existing literature on hospital behavior by studying whether a hospital's composition of patients changes as it becomes busier. While there is already a rich literature analyzing whether hospital ownership type impacts the type of patients treated, no work has focused on the question of if the set of patients treated by a hospital changes as it becomes busier. If this composition changes, it may provide evidence that hospitals do not turn patients away randomly and instead turn away less profitable patients while continuing to admit those that are more profitable.

4.3 Theoretical Model

My theoretical model is not intended to fully describe hospital behavior. I instead outline how, with some assumptions about hospital behavior, the composition of patients at a hospital may change based on the hospital's turnaway probability. Throughout the model, the strategic agent is the set of hospitals. I will focus on one hospital, h , as I abstract from how h 's actions affect its competitors. The model also includes a set of patients, but they simply show up to a hospital in hopes of receiving treatment and I do not model their utility.

The model assumes that there are an infinite number of periods where in each period t , a patient shows up to hospital h with some known probability. There are two types of patients, where high types yield a utility of u_H for each period they stay at the hospital and low types yield a per period utility of u_L with $u_H > u_L > 0$.¹ When a patient arrives, the hospital knows his type and can choose whether to admit him or turn him away. The hospital has a fixed number of beds b , and if all of its beds are being used, the hospital cannot accept any additional patients. Because it is capacity constrained, the hospital must evaluate the opportunity cost from utilizing a bed and weigh it against the utility not using the bed and potentially receiving a patient in future periods. At most, one patient arrives at the hospital in each period and the probability of a high type is p_H and that of a low type is p_L .

In each period, the hospital discharges one patient with probability p_d .² This probability is not affected by how long each patient has already been at the hospital, and the facility does not know the patient's length of stay when choosing whether to admit the individual. There are therefore two sources of uncertainty to the hospital

¹ The difference in the utility stems from their profitability. High types are more profitable for a hospital than low types.

² This probability can change based on the number of patients currently at the hospital and the model's predictions will not change.

- whether a patient will show up in each period (and the patient's type), and how long each of the patients admitted to the hospital will stay. Because the probability of a patient arrival is constant across periods, and admitted patients have the same probability of being discharged independent of how many periods they have been at the hospital, a hospital with b beds can have its expected lifetime utility described based on how many of each patient type are currently at the hospital.

The timing of each period in the model is outlined below.

- (1) The period begins.
- (2) A patient currently at the hospital is discharged with probability p_d .
- (3) A high type arrives with probability p_H and a low type arrives with probability p_L (where both types cannot arrive in the same period)
- (4) If a patient arrives and the hospital is not currently full, it decides whether to admit him.
- (5) The period ends.

In each period, the hospital only makes one potential decision - whether to admit the patient or not, conditional on a patient arriving. In this model, the state space is very simple as it only includes two variables, the number of high and low types currently admitted to the hospital, v_H and v_L . Given that per-period hospital utility is time-independent, I assume that the hospital plays a Markov strategy where its decision on whether to admit high and low types depends only on its current set of patients. The hospital's value function for its expected lifetime utility is written below.

$$V(v_H, v_L) = \alpha_H v_H + \alpha_L v_L + \beta \sum_i (\rho_i V'_i(v_H, v_L)) \quad (4.1)$$

The first part of equation (4.1) represents the utility that the hospital will receive in the current period from having v_H high and v_L low types. This is equal to the total utility derived from the set of patients currently being treated. The second part represents the expected lifetime utility beginning in the following period, conditional on the hospital's admission decision in the current period. The probabilities associated with the transition between the current state and that in the next period (the ρ_i values) depend on hospital's Markov strategy. This set of probabilities will allow for the possibility that the hospital subtracts a patient of either type, that the total number of patients and the composition of patients is unchanged, the total number of patients is unchanged but it adds or subtracts one high type, or that the hospital adds a patient of either type. Each of these probabilities will be affected by the state vector and the hospital's Markov strategy. For example, if the state vector indicates that the hospital is at full capacity, the probability that a hospital adds a patient of either type is zero. Likewise, the chance that a hospital adds a low type is zero when the hospital's strategy is to only admit high types.

I now introduce two propositions describing results predicted from the model. I then explain the logic supporting each of these claims.

Proposition 1. *If the hospital is not currently at full capacity and a high type arrives, the hospital will always admit the individual. In other words, $V(v_H+1, v_L) > V(v_H, v_L)$ for all v_H and v_L .*

The rationale behind this claim is that the maximum utility a hospital can receive in a given period is $\alpha_H b$ if all b beds are filled with high types. As a result, a hospital always wants to add more high types and will never turn one away when it has availability.

Proposition 2. *Whether a hospital accepts a low type depends on several factors including the relative values of v_H and v_L , the relative probabilities of each patient*

type showing up (p_H and p_L), and how many empty beds the hospital currently has ($b - v_L - v_H$).

As the value of v_H increases relative to v_L , a hospital becomes less likely to accept low types. This is intuitive as more valuable high types makes accepting a low type more costly in comparison to waiting and keeping the bed open for a high type in future periods. This is demonstrated by two examples, one where v_H equals v_L and the second where v_L equals 0. When the per period utility from each type is equal, the hospital will treat both groups as equal and therefore admits all low types when it has open beds as it does with high types. However, when the utility from low types is zero, hospitals receive no additional utility from adding a low type. However, it may prevent the hospital from adding a high type in future periods which would provide positive utility so the hospital will never admit low types.

As the probability of a high type showing up increases relative to a low type, the likelihood that the hospital admits low types again decreases. The rationale is that as this ratio skews towards high types, the opportunity cost of filling the best increases because it is more likely to be filled by a high type in the future. The value from filling it with a low type is unchanged, however and this means that a hospital is more likely to turn away the low type and leave it empty.

Finally, the number of empty beds is important in the hospital's decision because it impacts the likelihood that a hospital will be capacity constrained and unable to take any additional patients. For example, if a hospital is entirely empty and a low type shows up, the hospital is likely to admit the patient because it is very unlikely that the hospital will completely fill up before that patient is discharged from the hospital. On the other hand, if a hospital only has one bed remaining, admitting a low type in the current period makes it much more likely that a high type will be turned away in future periods because the hospital is full. As a result, admitting

a low type when the hospital is nearly full is far more likely to inhibit the hospital from adding a high type than when the hospital has lots of empty beds.

If hospitals behave in a manner consistent with my theoretical model, they will turn away more low types when they are nearly full than when they have many empty beds. I will test this prediction empirically with hospital discharge records to determine if less profitable patients make up a smaller fraction of total patients as a hospital becomes busier.

4.3.1 Generalizing the Model

While my theoretical model is relatively simple and makes several assumptions about patients and hospital behavior that are restrictive, many of these can be generalized without changing the model's core predictions. For example, I currently allow for two types of patients, but instead the model could include more types, or a continuous spectrum of patient types and the prediction would be unchanged. Likewise, I use discrete periods where a hospital may add or discharge patients, but this could be adjusted to occur in continuous time.

4.4 Data

My analysis uses two primary datasets. The first dataset consists of hospital inpatient discharge records for all non-federal hospitals over a fifteen year period. The second dataset is the American Hospital Association (AHA) annual survey data which includes hospital level characteristics for all U.S. hospitals. I merge these data using a hospital identifier.

4.4.1 Inpatient Data

I use inpatient discharge data from 1995 to 2009 for Arizona, New Jersey, New York, and Washington state.³ This set of states and years was employed because of its availability through the National Bureau of Economic Research (NBER). The data is collected from each hospital by the Agency for Healthcare Research and Quality (AHRQ) through the Healthcare Cost and Utilization Project (HCUP) program. This data is superior to national inpatient data for my analysis because it includes all inpatient visits and hospitals whereas national datasets such as the Medicare claims data and the National Inpatient Sample (NIS) only include a subset of patients and/or hospitals.

The inpatient data includes a number of patient characteristics that will be useful for my analysis. It includes variables for the patient's length of stay, month of admission, racial status, diagnosis related group, whether it was an emergency or elective procedure, and the patient's payment type.⁴ The inpatient data also includes an identifier for which hospital the patient attends for treatment. I will use the information on patient length of stay and which hospital each individual attends (as well as the facility's capacity which is given in the hospital data) to estimate each hospital's turnaway probability in each period. The patient variables such as admission type are used to generate estimates at the hospital level for the fraction of patients admitted for emergencies and elective procedures.

4.4.2 Hospital Data

The hospital data is collected by the American Hospital Association annually which allows me to track individual hospitals over time. The data includes detailed in-

³ The data begins in 1997 in New Jersey and in 1999 in Washington.

⁴ Ideally, the data would include the patient's date of admission so that I could more precisely estimate how many patients were at a hospital on any given day, but due to confidentiality concerns, the data only includes the month of admission.

formation about the hospital including its location, ownership type, and capacity, which can be tracked over time.

4.4.3 Final Dataset

I merge the two datasets using the hospital identifier and collapse the data so that the unit of observation is no longer the inpatient visit and instead the hospital-month combination.⁵ This allows me to focus on how the hospital's composition of patients changes based on its busyness in each period. Summary statistics of the final dataset are given in Table 4.1 where the values represent the fraction of inpatients admitted to the hospital meeting the given the criteria.

As the summary statistics show, nonprofits compose three quarters of my sample with local government hospitals making up an additional fifteen percent and for-profit hospitals comprising the remaining ten percent. While the composition of patients and patient characteristics is similar in some ways across hospital ownership types, it also varies dramatically in many dimensions.

About half of all admissions in my sample are for emergencies, but this value varies from 51 percent of nonprofit admissions to 35 percent of patients at for-profit hospitals. Instead, for-profit hospitals see a larger share of elective admissions at 41 percent whereas nonprofit hospitals only have 26 percent of their admissions for elective procedures. The fraction of patients admitted for child births is relatively constant across hospital ownership types at approximately 8 percent.

There is also a big difference in the payment types based on hospital ownership. The summary statistics are consistent with the notion that for-profit hospitals serve wealthier patients whereas government owned hospitals serve as safety net hospitals treating those with less financial means. This is demonstrated by the fact that

⁵ While my inpatient data allows me to observe the flow of patients each month, the hospital data is yearly so I only observe hospital capacity changes annually. As a result, I assume that hospital capacity is constant across each month of the year.

Table 4.1: Summary statistics: All values represent the mean fraction of patients with the given characteristic or condition. Values in parentheses represent the standard deviation of each mean. Each observation in N represents a hospital-month combination. Columns differentiate patient composition by hospital ownership type.

	All Hospitals	Gov't Owned	Nonprofit	For-Profit
<hr/> Admission Type <hr/>				
Emergencies	0.49 (0.25)	0.45 (0.31)	0.51 (0.23)	0.35 (0.32)
Urgent	0.16 (0.19)	0.22 (0.26)	0.15 (0.17)	0.17 (0.24)
Elective	0.27 (0.23)	0.24 (0.25)	0.26 (0.19)	0.41 (0.37)
Birth	0.08 (0.07)	0.09 (0.08)	0.08 (0.07)	0.07 (0.10)
<hr/> Payment Type <hr/>				
Medicare	0.39 (0.17)	0.37 (0.21)	0.39 (0.15)	0.44 (0.23)
Medicaid	0.18 (0.16)	0.28 (0.19)	0.16 (0.15)	0.16 (0.17)
Private Insurance	0.34 (0.16)	0.24 (0.17)	0.36 (0.15)	0.34 (0.20)
Self-Pay	0.06 (0.08)	0.09 (0.12)	0.05 (0.07)	0.02 (0.04)
<hr/> Other <hr/>				
Black	0.12 (0.17)	0.24 (0.24)	0.11 (0.16)	0.06 (0.10)
Minority	0.30 (0.28)	0.51 (0.36)	0.28 (0.27)	0.25 (0.20)
Cardiac	0.16 (0.10)	0.12 (0.08)	0.16 (0.09)	0.15 (0.16)
Mental Health	0.07 (0.13)	0.08 (0.15)	0.07 (0.12)	0.05 (0.17)
Die	0.03 (0.06)	0.03 (0.04)	0.03 (0.06)	0.03 (0.06)
N	75,611	11,602	56,649	6,159

for-profit hospitals have higher shares of Medicare and privately insured patients (tallying 78 percent of admissions versus 61 percent) and government owned hospitals have nearly four times as many uninsured patients as for-profit hospitals (9 percent versus 2 percent), which is consistent with the findings of Norton and Staiger (1994).

This notion of cherry picking patients is also evident when looking at racial status instead of insurance type as government owned hospitals serve a larger fraction of black and minority patients than the other types of hospitals.⁶ As discussed earlier, some types of admissions such as cardiac patients are typically profitable for a hospital, whereas others such as mental health patients are not. For-profit hospitals serve more cardiac patients than government owned hospitals (15 percent versus 12 percent) and fewer mental health patients (5 percent versus 8 percent). These statistics support the argument that for-profit hospitals serve a more profitable set of patients when categorizing profitability using DRGs. This finding is consistent with Horwitz (2005) which concludes that for-profit hospitals provides more services that are typically profitable to the hospital.

4.5 Empirical Model

My empirical model consists of two parts. In the first, I estimate the hospitals turnaway probability in a given period using a queuing model. In the second, I use this turnaway estimate to analyze whether the composition of inpatient hospital visits changes based on the hospital's crowdedness with a fixed effects regression framework.

⁶ I categorize any patient who is identified as "black," "hispanic," "Asian or Pacific Islander," "Native American," or "other" as a minority where the option of "white" is also given.

4.5.1 Estimation of Turnaway Probabilities

The inpatient data detailing each hospital visit does not include the specific date of admission or discharge for the patient. Instead, it includes the months in which the patient was admitted and released, as well as the length of stay in days. As a result, I do not observe how many patients are in a hospital at a given time. I therefore use a queuing model developed in Shonick (1970) and used by Joskow (1980) and state health planning agencies to estimate the probability that a hospital will turn away a patient based on its average occupancy during the period. This model estimates turnaways under the assumption that a hospital accepts a patient whenever it has an empty bed available.

Each month will represent a period as this is the shortest time unit in my data. An advantage to using the month as the period of observation as opposed to the year is that it provides more precise estimates of turnaways as the patient arrival rate and average length of stay may vary seasonally meaning a hospital could have a higher turnaway probability in one month, and a lower one the next.

I begin by assuming that patients arrive at hospital h in period t with a Poisson distribution of λ_{ht} and their length of stay has a negative exponential distribution with a mean of $1/\mu_{ht}$. With these assumptions, hospital h has a probability P_{htj} that j of its total b beds are full at any moment in period t .

$$P_{htj} = \frac{(\lambda_{ht}/\mu_{ht})^j/j!}{\sum_{k=0}^b (\lambda_{kt}/\mu_{kt})^k/k!} \quad (4.2)$$

The distribution of P_{htj} is not Poisson because it is truncated at the hospital's capacity limit, b . However, as this value becomes large relative to the mean number of patients, the Poisson distribution becomes a good approximation for P_{htj} . Likewise, the mean and variance of a Poisson are equal and it can be approximated with

a normal distribution as it becomes large. Using these assumptions, I rewrite the probability that hospital h is full as the probability that all b of h 's beds are currently occupied. This is done in equation 4.3 where the average daily census for the time period of interest is written as ADC and calculated using the patient arrival rate and mean length of stay in each month.

$$P_{ht} = 1 - \Phi\left(\frac{b_{ht} - ADC_{ht}}{\sqrt{ADC_{ht}}}\right) \quad (4.3)$$

P_{ht} measures the turnaway probability for hospital h in period t . It is possible that this measure could be biased in either direction. The probability may undercount turnaways if the patient arrival rate or mean length of stay is not random over the sample period.⁷ If hospital congestion is expressed in different ways such as waiting lists for elective patients, or patients do not show up with a Poisson distribution, this model may overstate the number of turnaways. Fortunately, if the bias is consistently in one direction, it is unlikely to change the interpretation of my analysis. If my estimated turnaway probability overstates (or understates) turnaways by 10 percent, my analysis will still be informative in how the composition of patients at a hospital is impacted by its busyness.

This model estimates turnaways under the assumption that all patients are similar and hospitals only turn away patients when they have no remaining beds available. I use this estimated probability to look at how the hospital's patient composition changes when turnaways increase thus implying that hospitals may not turn patients away solely based on bed availability. If this assumption is true, my estimated turnaway probability will instead act as a measure of hospital congestion where the larger the turnaway probability, the greater the congestion issue. It serves the same purpose, however, as increased congestion will mean that more patients need to be

⁷ For example, Joskow (1980) finds that patients arrive more frequently on weekdays than on weekends.

turned away.

4.5.2 Regression Framework

To specifically study how a hospital's turnaway probability impacts its composition of patients, I use a simple regression framework that includes hospital and time fixed effects. I include several specifications that look at different aspects of inpatient composition. In equation (4.4) below, I give the functional form used for each.

$$F_{ht} = \alpha_h + \alpha_t + \alpha_f P_{ht} + \epsilon_{ht} \quad (4.4)$$

F_{ht} represents fraction of all inpatient visits to hospital h in period t that fit the criteria of interest. For example, F_{ht} may refer to the fraction of visits that are classified as emergencies. The remaining variables in the equation are unchanged across specifications. α_h is a hospital fixed effect that controls for the overall patient composition at the hospital and α_t is a time fixed effect to control for seasonal differences or time trends in the overall makeup of patient compositions in my sample. ϵ_{ht} is an error term that is unobserved by the econometrician.

α_f is the parameter that governs the effect of the estimated turnaway probability on patient composition. Ultimately, this is the variable of interest because if turnaways do not affect the overall composition of patients at a hospital, the value of α_h will be zero. The fixed effects are necessary for α_f to have this interpretation as they control for variation across hospitals and over time. As a result, a positive and significant α_f coefficient can be interpreted as meaning that as a hospital's turnaway probability increases, the fraction of patients of type F_{ht} increases as well. This is different than the result in Hsia et al. (2012) which looks specifically at turnaways and minority patients, but focuses on the differences across hospitals and finds that hospitals serving larger minority populations also turn away more individuals. That analysis emphasizes differences in hospital crowdedness across neighborhoods and

cities whereas I focus on individual hospital behavior with respect to turnaways.

4.6 Results

I present my results in three parts. The first looks at the effect of turnaways on the composition of inpatient admission types such as emergencies and elective procedures. As with each set of results, I estimate parameters for the entire set of hospitals as well as for each hospital ownership type - government owned, nonprofit, and for-profit. The second part analyzes the effect of turnaways on the composition of patient payment types. Finally, I look at some additional groups of patients broken out by race and diagnosis related group.

In each set of regressions, the format used is that given in equation (4.4) and each coefficient and robust standard error (clustered by hospital) represents a separate regression where the independent variable has changed, but the dependent variable of interest remains the estimated turnaway probability. This coefficient on this variable is estimated separately for each hospital and month using equation (4.3).

4.6.1 Admission Type

I analyze the admission type to gauge whether turnaways affect patients with emergencies differently than those undergoing elective procedures. If the results suggest that the groups were impacted differently, it may make the interpretation of any significant results with respect to payment types or DRGs more challenging. The rationale is that if hospitals only turn away emergency patients, but they do so randomly, these individuals may not be representative of all inpatients in other dimensions. For example, they may be less likely to be insured. If I then find that the uninsured see their hospital composition drop as the facility turns more patients away, the hospital may just be turning away more emergency patients. If I do not find evidence that patient admission type is affected by turnaways, the argument

that hospitals turn patients away non-randomly becomes stronger.

Table 4.2: Regressions on admission type: In each regression, the independent variable is the fraction of inpatients in the month who are of the admission type given in the row. Dependent variables include hospital and time fixed effects, and the hospital’s turnaway probability. The coefficient on the turnaway probability is given in the table and the robust standard error which is clustered by hospital is given in parentheses. * indicates that the regression is statistically significant at the 10 percent level and ** indicates that it is significant at the 5 percent level. Each coefficient and clustered standard error therefore represents a separate regression.

Admission Type	All	Gov’t Owned	Nonprofit	For-Profit
Emergency	0.005 (0.014)	0.049 (0.054)	-0.000 (0.016)	0.002 (0.10)
Urgent	-0.017 (0.015)	-0.087* (0.051)	0.006 (0.012)	-0.083 (0.051)
Elective	0.009 (0.014)	0.028 (0.025)	-0.010 (0.10)	0.082 (0.052)
Birth	0.003 (0.003)	0.012 (0.007)	0.003 (0.004)	0.004 (0.002)

The results of these regressions are given in Table 4.2. The first column provides the regression coefficient on turnaways for all hospitals, and the next three columns restrict the analysis to hospitals of the specified ownership type. Overall, the results suggest that turnaways have little effect on the composition of admission types as the specification for all hospitals does not find the coefficient to be statistically significant for any of the admission types. Likewise, the results appear similar when looking at specific hospital ownership types as the only statistically significant result is that the fraction of urgent admissions for government owned hospitals decreases as they become busier.

These results suggest that patients of all admission types are similarly affected by turnaways. If turnaways only impacted emergency admissions through ambulance diversions, the hospital would likely see its fraction of elective procedures increase as it became busier because the hospital would turn away emergency patients in favor of elective ones thus increasing the fraction of elective patients admitted. While this

is not proof that the composition of patients seeking admission to a hospital changes based on congestion, it supports the argument that the distribution of potential patients is constant regardless of how busy the hospital is.

4.6.2 Payment Type

I look at how a hospital’s turnaway probability affects its composition of patients by payment type as certain groups may be more desirable to hospitals than others as discussed earlier. For example, patients who pay with private insurance are generally profitable for a hospital whereas patients who are uninsured are not.

Table 4.3: Regressions on payment type: In each regression, the independent variable is the fraction of inpatients in the month who are of the payment type given in the row. Dependent variables include hospital and time fixed effects, and the hospital’s turnaway probability. The coefficient on the turnaway probability is given in the table and the robust standard error which is clustered by hospital is given in parentheses. Each coefficient and clustered standard error therefore represents a separate regression.

Payment Type	All	Gov’t Owned	Nonprofit	For-Profit
Medicare	0.003 (0.006)	-0.025* (0.013)	0.011* (0.006)	-0.023 (0.019)
Medicaid	0.004 (0.007)	0.011 (0.028)	0.012 (0.007)	-0.021 (0.020)
Private Insurance	-0.008 (0.009)	0.031* (0.016)	-0.011 (0.010)	-0.012 (0.030)
Self-Pay	-0.007* (0.003)	-0.006 (0.017)	-0.007* (0.004)	-0.003 (0.003)

Table 4.3 provides the regression results which follow the same form as the earlier results in pooling all hospitals together as well as breaking them out by ownership group. As with the admission type parameters, many are not statistically different than zero. For all hospitals, however, the fraction of self-pay patients decreases when the hospital becomes more crowded. This is consistent with the notion that as a hospital becomes busier, it is less likely to admit a low type patient and instead may keep the bed open for a more profitable patient who has insurance.

There are also some interesting results when focusing on specific hospital ownership types. For example, as the turnaway probability increases, government owned hospitals admit a smaller fraction of Medicare patients and more individuals with private insurance. Nonprofits, on the other hand, admit a larger fraction of Medicare patients and fewer uninsured. In each case, the patient profile becomes more profitable for the hospital as it gets busier because privately insured patients are more profitable than Medicare patients, who are more profitable than the uninsured.

4.6.3 Race and Diagnosis Related Group

Hsia et al. (2012) study whether hospital turnaways disproportionately affect hospitals serving larger minority populations. I include breakdowns of race to address a related, but different question. My analysis looks at the how the racial composition of a hospital changes as it becomes busier. This question relates to my other results as racial status may serve a proxy for profitability for hospitals as minorities are less likely to be insured, and may be less profitable as a result.

Additionally, I look at the fraction of patients who fall within the cardiac and mental health diagnosis related groups as cardiac procedures are generally very profitable for hospitals whereas mental health patients are not. As a result, if hospitals do not turn patients away randomly, they may be more likely to accept cardiac patients at the expense of those with mental health diagnoses as they get busy. Finally, I include a measure for the fraction of patients who died while at the hospital to gauge whether the hospital's congestion affects the quality of care it provided.

Beginning with the entire pool of hospitals, my results suggest that as a facility becomes busier, its fraction of black patients and mental health patients decreases. Both of these results are consistent with the notion that a hospital is more likely to turn away patients that are less profitable as each of these sets of patients would be categorized as low types in my theoretical model. Additionally, while the result with

Table 4.4: Regressions on race and diagnosis related group: In each regression, the independent variable is the fraction of inpatients in the month who are of the racial status or diagnosis related group given in the row. Dependent variables include hospital and time fixed effects, and the hospital’s turnaway probability. The coefficient on the turnaway probability is given in the table and the robust standard error which is clustered by hospital is given in parentheses. Each coefficient and clustered standard error therefore represents a separate regression.

Other	All	Gov’t Owned	Nonprofit	For-Profit
Black	-0.012** (0.006)	-0.022 (0.016)	-0.012 (0.007)	-0.012* (0.006)
Minority	-0.020 (0.013)	0.008 (0.010)	-0.012 (0.010)	-0.058 (0.058)
Cardiac	0.006 (0.008)	-0.005 (0.009)	-0.001 (0.005)	0.047 (0.038)
Mental Health	-0.006** (0.003)	-0.021** (0.006)	-0.005 (0.003)	-0.000 (0.003)
Die	0.003 (0.002)	0.013 (0.011)	0.001 (0.001)	0.003 (0.12)

respect to black patients is different than that found in Hsia et al. (2012), they both suggest that black patients are disproportionately affected by hospital turnaways.⁸

When analyzing hospitals by ownership type, the results suggest that government hospitals admit a smaller fraction of mental health patients as their turnaway probability increases. Additionally, for-profit hospitals take in fewer black patients when they become busier. While the coefficients for minority patients and cardiac patients are not statistically significant for any of my specifications, they are also each in the direction that would be consistent with hospitals being more likely to turn away less profitable patients. Finally, the coefficients on death are positive for all hospital ownership types, but not statistically significant. If a hospital’s quality of care suffers when it becomes busier, we would expect increased turnaways to lead to a higher death rate.

My results are consistent with the predictions with my theoretical model which

⁸ I also look at the fraction of black patients and hospital turnaways using Hsia’s framework and find that hospitals with higher turnaway probabilities serve more minority patients.

assumes that hospitals do not randomly turn patients away and instead are more likely to turn away patients that are less profitable. This result occurs when differentiating patients based on payment type as uninsured patients see their share of admissions decrease as the hospital becomes busier. This finding also holds with black patients and mental health patients.

4.7 Extensions

While this analysis provides some insight into how patient composition changes at the hospital level when turnaway probabilities change, there is room for additional analysis to further shed light on the questions addressed in this study. For example, at the moment, I use four states and fifteen years of data. In future analysis, I hope to include more states and years to have a nationally representative sample. With additional observations, I will be able to more precisely estimate the effects of turnaways and hospital congestion on all types of patients.

While my analysis focuses on the effect of hospital congestion on patient composition, there may be other mechanisms through which a hospital's congestion affects patients. As I briefly mention in Table 4.4, I consider whether the fraction of patients who die in the hospital is impacted by the total number of turnaways. The framework I use is simple, but may not account for changes in the condition of patients when a hospital becomes more crowded. Likewise, I may not fully capture how quality is affected by crowdedness through other dimensions besides mortality. A more thorough analysis on this topic studying the issue of quality is needed.

Additionally, when a hospital becomes busy its behavior may change in ways that I am not currently measuring. For example, when a hospital is at capacity and a new patient shows up, it may discharge another patient in order to admit the new one. If such behavior occurs, we would expect patient length of stays to decrease as the hospital becomes busier. Likewise, if hospitals are busier, they may change how they

treat patients and choose less invasive procedures that will allow them to discharge the individuals more quickly. Because the inpatient data has both diagnosis and procedure codes for each inpatient visit, I plan to measure the intensity of treatment conditional on the diagnosis to determine whether it changes based on the hospital's turnaway probability.

While I categorize inpatients along several dimensions that may demonstrate whether hospitals' patient compositions change in ways consistent with utility maximization when they become busier, the richness of the inpatient data means that this question can be addressed in many ways. Future research will look at other DRGs or used procedure codes as well to better measure the relative profitability of patients.

In my summary statistics I briefly discuss whether cherry picking seems to occur based how patient compositions change by hospital ownership type. My data would allow for a more detailed analysis of this question by looking at additional factors such as the distance that patients travel to visit a hospital. Accounting for markets and the set of hospitals near a patient may help determine whether for-profit hospitals treat a different set of patients than others because of their location, as suggested in Norton and Staiger (1994), or if this disparity stems from them targeting specific types of patients.

A natural extension of this analysis would use different categorizations of hospitals besides ownership type. For example, do hospitals in urban markets that have several local competitors behave differently than rural hospitals that are the only local option for patients? Additionally, does hospital size or inclusion in a multi-hospital network impact behavior? I plan to address many of these questions in future research.

Finally, right now I do not study whether congestion is an issue at the hospital level, or if it may be a local market issue where all hospitals in an area may become congested at the same time due to fluctuations in demand. Such a finding would

likely increase any welfare concerns as it would imply that if a less profitable patient was turned away from a local hospital, he may face similar issues at other local hospitals thus requiring him to travel further to seek treatment.

4.8 Conclusion

My analysis finds evidence that the composition of inpatients changes as a hospital becomes busier and starts turning patients away. These results suggest that patients whose fraction of admissions decreases are less profitable to the hospital such as mental health patients, uninsured patients, and black patients. While this result does not prove that hospitals are discriminatory in which patients they turn away, it is consistent with such a story, which is outlined in my theoretical model.

More generally, this research has important policy implications for a couple of reasons. First, the U.S. population is aging and living longer. While advances in medical technology has moved some procedures out of hospitals to outpatient facilities and shortened the necessary length of stay for others, there is still likely to be a growing demand for hospital services in the future. As a result, hospitals are likely to become busier making the issues outlined in this paper even more important going forward. It is pivotal for policy makers to understand how different types of patients are affected when capacity becomes constrained and this study intends to help inform them on this topic. This paper also contributes to a growing literature that examines hospital behavior and attempts to determine whether they act as typical firms, or behave in ways more consistent with other nonprofit institutions.

Conclusions

My dissertation studies three questions in applied microeconomics with important policy implications. I was drawn to such topics because I wanted to ensure that my results were not only relevant to other economists studying similar questions, but rather that they would be of interest to policymakers and the public more broadly. For these reasons, my analysis focus on markets that are important to average consumers as evidenced through total spending - health and housing.

The second chapter of my dissertation studies the effect of a state level regulation on hospital capacity decisions, and ultimately on patient welfare. To conduct this analysis, I use an array of econometric techniques commonly employed in the industrial organization and health literatures. My results suggest that the regulation decreases total hospital investment in capacity and disproportionately affects smaller and for-profit hospitals. I estimate the effect on patient welfare and conclude that patients are made worse off by the regulation as they are more likely to be turned away from their preferred hospital.

The third chapter analyzes the real estate market in Los Angeles with a focus on house flippers and how they earn their returns. I find that flippers can be dif-

ferentiated based on their experience level and more experienced flippers earn their returns by targeting houses being sold at below market value. Inexperienced flippers do not look for underpriced homes and instead attempt to time the market and earn returns by holding houses during periods of high market appreciation. I exploit the panel nature of the data to control for unobserved investment to the home made by flippers so that the estimated returns are not caused by physical improvements to the homes. Additionally, I consider whether flippers contributed to the housing market bubble and find evidence that inexperienced flippers accentuated the bubble whereas experienced flippers did not.

The fourth chapter researches how hospital composition changes as a health facility becomes more busy. I build a theoretical model to predict that a hospital will turn away less profitable patients while still accepting those that are more profitable when it becomes congested. I test this prediction using inpatient data and a fixed effects regression framework and find that as a hospital becomes busier, its fraction of less profitable patients decreases suggesting that they are more likely to be turned away. This is consistent with my theoretical model and implies that such patients are disproportionately affected by hospital turnaways.

I am hopeful that my findings in each of these chapters will help to inform policymakers about the healthcare and housing industries. My results suggest that state level regulations governing hospitals hurt patient welfare meaning state level policymakers should consider discontinuing the regulation. I also find that hospitals appear to turn away less profitable patients at a higher rate than their more profitable counterparts. Given that the state level regulation impacts hospital turnaways, this result implies that the CON regulation also introduces distributional effects to consider. Finally, I study the housing industry and characterize how real estate flippers behave while also finding evidence that those with less experience likely contributed to the housing bubble. As a result, measures aimed to lessen the number of flips

from inexperienced flippers would help prevent another housing bubble from arising in the future. Such policies would need to specifically target the flippers acting as speculators, however, as those who serve as middlemen may help improve the market by injecting it with liquidity.

Appendix A

Estimation of Hospital Turnaway Probabilities and Substitution Patterns

This appendix outlines how I calculate the substitution patterns between hospitals, and the two other variables used in the estimation of first choice shares, s_t and s_s . s_t represents the share of patients who wanted to attend the hospital, but do not because they were turned away. s_s is equal to the share of patients who are admitted to the hospital after being turned away from their first choice hospital. My methodology follows from the nested logit equations outlined in Berry (1994). The values for s_t and s_s are calculated assuming that I have already estimated the regression coefficients in equation (2.5), and that these coefficients represent the true patient parameters. Using these coefficients, I calculate the predicted share of patients for each hospital from each ZIP code, \bar{s}_h .

Equation (A.1) estimates the within-nest share which is equal to the fraction of patients in ZIP code z who choose hospital h , conditional on choosing a hospital of h 's ownership type, g . This is equivalent to a standard multinomial logit equation for all patients and hospitals with ownership type g . Equation (A.2) predicts the

nest shares which are equal to the fraction of patients in z who choose a hospital of ownership type g . In both of these equations, δ_h is the predicted utility for hospital h and σ is the nesting parameter's correlation that were calculated from the parameter estimates generated in equation (2.5). The predicted share, \bar{s}_h , is equal to the hospital's within-nest share multiplied by the nest's share as given in equation A.3.

$$\bar{s}_{hg}(\delta_h, \sigma) = \frac{e^{\delta_h/(1-\sigma)}}{\sum_{h \in G} e^{\delta_h/(1-\sigma)}} \quad (\text{A.1})$$

$$\bar{s}_g(\delta_h, \sigma) = \frac{[\sum_{h \in G} e^{\delta_h/(1-\sigma)}]^{(1-\sigma)}}{\sum_G [[\sum_{h \in G} e^{\delta_h/(1-\sigma)}]^{(1-\sigma)}]} \quad (\text{A.2})$$

$$\bar{s}_h = \bar{s}_{h/g} \bar{s}_g \quad (\text{A.3})$$

To estimate the substitution patterns for hospital h , I decrease the utility associated with hospital h , δ_h , by a small amount, ϵ , while holding the utility for all other hospitals constant. I compare how the predicted share of patients changes based on this decrease in δ_h and determine how the fraction of patients who no longer select h reallocate among the remaining hospitals by reestimating equations (A.1) and (A.2). I repeat this calculation for each hospital to get a complete set of substitution patterns.

I estimate the share of patients turned away from each hospital using the formula for turnaway probability given in equation (2.2). Multiplying the turnaway probability by the total number of patients whose first choice is h gives an estimate for the number of turned away patients, s_t . I reallocate these patients to other hospitals using the substitution patterns for hospital h .¹ After reallocating all turned away

¹ For example, if the substitution patterns estimate that one quarter of patients who leave hospital h_1 now select hospital h_2 , then one quarter of the patients turned away from h_1 will be admitted to h_2 .

patients, I sum the total number of additional patients each hospital receives because they were turned away from their first choice hospital to generate s_s . Having solved for these two variables, I now estimate the first choice shares, s_f .

Bibliography

- Abelson, R. and Creswell, J. (2012), “Hospital Chain Inquiry Cited Unnecessary Cardiac Work,” *New York Times*.
- Bagust, A., Place, M., and Posnett, J. W. (1999), “Dynamics of Bed Use in Accommodating Emergency Admissions: Stochastic Simulation Model,” *British Medical Journal*, 319.
- Bajari, P., Benkard, C. L., and Levin, J. (2007), “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, 75.
- Barberis, N. and Thaler, R. (2003), “A Survey of Behavioral Finance,” in *Handbook of the Economics of Finance*, eds. G. Constantinides, M. Harris, and R. Stulz, vol. 1.
- Benkard, C. L., Bodoh-Creed, A., and Lazarev, J. (2010), “Simulating the Dynamic Effects of Horizontal Mergers: U.S. Airlines,” .
- Berman, L. E., Geissler, C., and Toder, E. J. (2008), “How Big Are Individual Income Tax Expenditures and Who Benefits From Them?” *American Economic Review Papers and Proceedings*, 98.
- Berry, S. T. (1994), “Estimating Discrete-Choice Models of Product Differentiation,” *The RAND Journal of Economics*, 25.
- Brunnermeier, M. K. and Nagel, S. (2004), “Hedge Funds and the Technology Bubble,” *Journal of Finance*, 59.
- Burda, D. (1991), “CONspiracies to Crush Competition: Hospitals Using CON Laws to Thwart Rival’s Projects,” *Modern Healthcare*, 27.
- Cardell, N. S. (1997), “Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity,” *Econometric Theory*, 13.
- Case, K. E. and Shiller, R. J. (1987), “Prices of Single-Family Homes Since 1970: New Indexes for Four Cities,” *New England Economic Review*, pp. 45–56.

- Case, K. E. and Shiller, R. J. (1989), "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, 79.
- Conover, C. J. and Sloan, F. A. (1998), "Does Removing Certificate-of-Need Regulations Lead to a Surge in Health Care Spending?" *Journal of Health Politics, Policy and Law*, 23.
- Decker, S. L. (2012), "In 2011 Nearly One-Third of Physicians Said They Would Not Accept New Medicaid Patients, But Rising Fees May Help," *Health Affairs*, 31.
- DeLong, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990), "Positive Feedback Investment Strategies and Destabilizing Rational Speculation," *Journal of Finance*, 45.
- Department of Justice and Federal Trade Commission (2004), "Improving Health Care: A Dose of Competition," .
- Dobson, A., DaVanzo, J., Doherty, J., and Tanamor, M. (2005), "A Study of Hospital Charge Setting Practices," *Lewin Group*, 05.
- Eichengreen, B., Tobin, J., and Wyplosz, C. (1995), "Two Cases for Sand in the Wheels of International Finance," *Economic Journal*, 105.
- Ericson, R. and Pakes, A. (1995), "Markov Perfect Industry Dynamics: A Framework for Empirical Work," *Journal of Health Politics, Policy and Law*, 62.
- Fama, E. F. (1955), "The Behavior of Stock-Market Prices," *Journal of Business*, 38.
- Finch, L. E. and Christianson, J. B. (1981), "Rural Hospital Costs: An Analysis with Policy Implications," *Public Health Reports*, 96.
- Friedman, M. (1953), *Essays in Positive Economics*, University of Chicago Press.
- Geissler, C. and Harris, B. H. (2008), "Taxpayer Eligibility for IRAs," *Tax Notes*, 118.
- Geissler, C. and Harris, B. H. (2009), "The Automatic 401(K): Revenue and Distributional Estimates," in *Automatic: Changing the Way America Saves*, eds. W. G. Gale, J. M. Iwry, D. John, and L. Walker, Brookings Institution Press.
- Glower, M., Haurin, D. R., and Hendershott, P. H. (2003), "Selling Time and Selling Price: The Impact of Seller Motivation," *Real Estate Economics*, 26.
- Goetzmann, W. and Peng, L. (2006), "Estimating House Price Indexes in the Presence of Seller Reservation Prices," *Review of Economics and Statistics*, 88.

- Gowrisankaran, G., Lucarelli, C., Schmidt-Dengler, P., and Town, R. (2010), "Government Policy and the Dynamics of Market Structure: Evidence from Critical Access Hospitals," .
- Greenwood, R. and Nagel, S. (2009), "Inexperienced Investors and Bubbles," *Journal of Financial Econometrics*, 93.
- Haruvy, E., Lahav, Y., and Noussair, C. N. (2007), "Traders' Expectations in Asset Markets: Experimental Evidence," *American Economic Review*, 97.
- Haughwout, A., Lee, D., Tracy, J., and van der Klaauw, W. (2011), "Real Estate Investors, the Leverage Cycle, and the Housing Market Crisis," .
- Himmelberg, C., Mayer, C., and Sinai, T. (2005), "Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions," *Journal of Economic Perspectives*, 19.
- Horwitz, J. R. (2005), "Making Profits and Providing Care: Comparing Nonprofit, For-Profit, and Government Hospitals," *Health Affairs*, 24.
- Hsia, R. Y.-J., Asch, S. M., Robert E. Weiss, D. Z., Liang, L.-J., Han, W., McCreath, H., and Sun, B. C. (2012), "California Hospitals Serving Large Minority Populations Were More Likely Than Others to Employ Ambulance Diversion," *Health Affairs*, 8.
- Johnson, A. and Sataline, S. (2010), "Health Insurers Toughen Stance in Disputes over Hospital Rates," *Wall Street Journal*.
- Joskow, P. L. (1980), "The Effects of Competition and Regulation on Hospital Bed Supply and the Reserve Quality of the Hospital," *Bell Journal of Economics*, 11.
- Kessler, D. P. and McClellan, M. B. (2000), "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics*, 15.
- Lane, S. G., Longstreth, E., and Nixon, V. (2001), "A Community Leader's Guide to Hospital Finance," .
- Lanning, J. A., Morrissey, M. A., and Ohsfeldt, R. L. (1991), "Endogenous Hospital Regulation and Its Effects on Hospital and Non-Hospital Expenditures," *Journal of Regulatory Economics*, 3.
- Mankiw, N. G. and Whinston, M. D. (1986), "Free Entry and Social Inefficiency," *RAND Journal of Economics*, 17.
- National Cancer Center (2013), "Health Service Areas (HSAs)," <http://seer/cancer/gov/seerstat/variables/countyattribs/hsa.htm>.

- National Conference of State Legislatures (2013), "Certificate of Need: State Health Laws and Programs," <http://www.ncsl.org/issues-research/health/con-certificate-of-need-state-laws.aspx>.
- New York State Department of Health (2012), "Certificate of Need," <http://www.health.ny.gov/facilities/cons/>.
- Newhouse, J. P. (1989), "Do Unprofitable Patients Face Access Problems?" *Health Care Financing Review*, 11.
- Newhouse, J. P. (1995), "Economists, Policy Entrepreneurs, and Health Care Reform," *Journal of Regulatory Economics*, 14.
- Norton, E. C. and Staiger, D. O. (1994), "How Hospital Ownership Affects Access to Care for the Uninsured," *The RAND Journal of Economics*, 25.
- Peterson, E. D., Coombs, L. P., DeLong, E. R., Haan, C. K., and Ferguson, B. (2004), "Procedural Volumn as a Marker of Quality for CABG Surgery," *Journal of the American Medical Association*, 291.
- Rivers, P. A., Fottler, M. D., and Frimpong, J. . (2010), "The Effects of Certificate of Need Regulation on Hospital Costs," *Journal of Health Care Finance*, 36.
- Ross, J. S., Ho, V., Wang, Y., Cha, S. S., Epstein, A. J., Samoudi, F. A., Nallamotheu, B. K., and Krumholz, H. M. (2007), "Certificate of Need Regulation and Cardiac Catheterization Appropriateness After Acute Myocardial Infarction," *Circulation*, 115.
- Ryan, S. P. (2011), "The Costs of Environmental Regulation in a Concentrated Industry," .
- Sagness, J. (2007), "Certificate of Need Laws: Analysis and Recommendations for the Commission on Rationalizing New Jersey's Health Care Resources," .
- Salkever, D. S. and Bice, T. W. (1976), "The Impact of Certificate-of-Need Controls on Hospital Investment," *Milbank Quarterly*, 54.
- Salkever, D. S. and Bice, T. W. (1979), *Hospital Certificate-of-Need Controls: Impact on Investments, Costs and Use*, American Enterprise Institute for Public Policy Research.
- Shiller, R. J. (2003), "From Efficient Markets Theory to Behavioral Finance," *Journal of Economic Perspectives*, 17.
- Shleifer, A. and Summers, L. H. (1990), "The Noise Trader Approach to Finance," *Journal of Economic Perspectives*, 4.

- Shonick, W. (1970), "A Stochastic Model for Occupancy-Related Variables in General-Acute Hospitals," *Journal of American Statistical Association*, 65.
- Sloan, F. A. (2000), "Not-For-Profit Ownership and Hospital Behavior," in *Handbook of Health Economics*, eds. A. Culyer and J. Newhouse, vol. 1.
- Sloan, F. A. and Steinwald, B. (1980), "Effects of Regulation on Hospital Costs and Input Use," *Journal of Law and Economics*, 32.
- Smith, V. L., Suchanek, G. L., and Williams, A. W. (1988), "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Markets," *Econometrica*, 56.
- Springer, T. M. (1996), "Single-Family Housing Transactions: Seller Motivations, Price, and Marketing Time," *Journal of Real Estate Finance and Economics*, 13.
- Summers, L. H. and Summers, V. P. (1988), "When Financial Markets Work Too Well: A Cautious Case for a Securities Transactions Tax," *Journal of Financial Services Research*, 3.
- Tobin, J. (1974), "The New Economics One Decade Older," in *The Eliot Janeway Lectures on Historical Economics in Honor of Joseph Schumpeter*, Princeton University Press.
- Tobin, J. (1978), "A Proposal for International Monetary Reform," *Eastern Economic Journal*, 4.
- Vaughn-Sarrazin, M. S., Hannan, E. L., Gormley, C. J., and Rosenthal, G. E. (2002), "Mortality in Medicare Beneficiaries Following Coronary Artery Bypass Graft Surgery in States With and Without Certificate of Need Regulation," *Journal of the American Medical Association*, 15.
- Vissing-Jorgensen, A. (2003), "Perspectives on Behavioral Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions," in *NBER Macroeconomics Annual*, MIT Press.
- Washington State Department of Health (2013), "Certificate of Need Overview," .
- World Health Organization (2013), "Countries: United States of America," <http://www.who.int/countries/usa/en/>.
- Yee, T., Stark, L. B., Bond, A. M., and Carrier, E. (2011), "Health Care Certificate-of-Need Laws: Policy or Politics?" *National Institute for Health Care Reform Research Brief*, 4.

Biography

Christopher Scott Geissler was born in Hartford, Connecticut on October 27, 1983. He earned his B.A. in Economics and Mathematics from Williams College in 2006, his Master's in Economics from Duke University in 2009, and his PhD in Economics from Duke University in 2013.

In between graduating from Williams College and attending Duke University, he worked for two years as a research assistant at the Brookings Institution in Washington D.C. focusing on tax policy, the federal budget, and retirement savings. While there, he published several articles and book chapters including Geissler and Harris (2008), Berman et al. (2008) and Geissler and Harris (2009). He will be working as an economist at ISO New England beginning in the summer of 2013 where he will study electricity markets in New England.