Essays on the Institutional Design of Health Care Markets

Essays on the Institutional Design of Health Care Markets

A dissertation presented

by

Adrienne Hope Sabety

to

Interfaculty Initiative in Health Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Health Policy

Harvard University

Cambridge, Massachusetts

May 2020

Dissertation Advisor:
David Cutler

Author:
Adrienne Hope Sabety

Essays on the Institutional Design of Health Care Markets

Abstract

The institutional design of health care markets directly impacts patients' access, quality, and affordability of health care. Chapter one quantifies how patient-primary care provider (PCP) relationships affect patients' access to and quality of care. Chapter two designs a care coordination intervention targeting undocumented immigrants' access to and quality of care. Chapter three asks how plan choice affects consumers' quality, access, and affordability of health insurance.

Chapter one shows that relationships determine where patients demand care and positively affect patients' health. In cases where a PCP exogenously exits the market, patients do not form new PCP relationships, decreasing their use of primary care long-term. Instead, patients switch to specialists they know for less-extensive preventive care, which increases patients' probability of death by about 50 deaths per 100,000 individuals. Emergency department and inpatient admissions also increase for one year after a PCP's exit, increasing patients' spending by $4,640 and Medicare spending by $16,052 per exiting PCP. I establish the importance of relationship-specific capital in explaining effects, information that grows over time. Travel-cost revealed preference estimates suggest that patients are willing to pay $400-$500 to maintain a PCP relationship, providing a lower bound on patients' monetary valuation of relationship-specific capital.

Chapter two uses a randomized evaluation to measure the impact of coordinating undocumented immigrants' health care on utilization and health status. This work is joint with Jonathan Gruber, Rishi Sood, and Jin Yung Bae. We show that coordinating undocumented immigrants' health care increases utilization of primary care, self-reported

access to care, and preventive care, leading to a 7% decrease in individuals' long-run probability of death in back-of-the-envelope calculations. The intervention also causes high risk individuals to decrease their use of the emergency department by 2%, leading to $19,000 in government savings on net.

Chapter three, joint with Victoria Marone, focuses on choice over coverage level, "vertical choice." There is limited evidence on how vertical choice affects welfare despite it being a widespread feature of U.S. health insurance markets. The socially efficient level of coverage for a given consumer optimally trades off the value of risk protection and the social cost from moral hazard. Providing choice does not necessarily lead consumers to select their efficient coverage level. We show that in regulated competitive health insurance markets, vertical choice should be offered only if consumers with a higher willingness to pay for insurance have a higher efficient coverage level. We test for this condition empirically using administrative data from a large employer representing 45,000 households. We estimate a model of consumer demand for health insurance and health care utilization that incorporates heterogeneity in health, risk aversion, and moral hazard. Our estimates imply substantial heterogeneity in efficient coverage levels, but we do not find that households with higher efficient coverage levels have a higher willingness to pay. It is therefore optimal to offer only a single coverage level. Relative to a status quo with vertical choice, offering only the optimal single level of coverage increases welfare by $302 per household per year. This policy shift also leads to a more even distribution of health-related spending (premiums plus out-of-pocket costs) in the population, suggesting equity and efficiency improvements.

Contents

List of Tables

List of Figures

Acknowledgments

I am extremely grateful for those who made entering a PhD program possible. Tanya Downing, my economics teacher at Cuesta Community College, was the first to recognize my affinity for economics by giving me her yearly economics prize. Her belief in my ability encouraged me to major in economics when I transferred to UC Berkeley. At Berkeley, my classmate Joseph Root gave me confidence to ask questions during class. Christina Romer noticed my questions and, in response, pushed me to seriously consider entering a PhD program.

My dissertation committee gave me invaluable research as well as career advice throughout the PhD. David Cutler kept me focused and on track, teaching me to follow my curiosity and reading my introduction countless times. Tim Layton kept me sane with life and career advice and always being on my side. Claudia kept me striving to be the best researcher I could be. Jon Gruber gave me unfiltered advice and mentorship.

At a more basic level, my peers taught me how to be an economist while also supporting me emotionally. Lest anyone thinks a PhD is easy, these past six years were some of the hardest yet. Colin Gray and Ariella Spitzer were my life and research coaches and my biweekly swims with Savannah Bergquist were soul cleansing. I learned an incredible amount from my NBER reading group—Colin Gray, Ariella Kahn-Lang Spitzer, Ryan Hill, and Jonathan Roth. Of the six years, the last year was the hardest. I couldn't have done it without my rock, my safe place, my everything Gabriel Unger.

I also benefited greatly from the advice, mentorship, and friendship of countless others, such as Alex Bartik, David Card, Moya Chin, Edward Glaeser, Nir Hak, Lawrence Katz, Victoria Marone, Thomas McGuire, Michael McWilliams, Hannah Neprash, Dev Patel, Mark Shepard, Niharika Singh, Melanie Wasserman, and Annetta Zhou. Thank you for supporting me on this journey.

Chapter 1

The Value of Relationship-Specific Capital in Health Care[1]

[1]This work was supported in part by the National Science Foundation Graduate Research Fellowship Program (Grant No. DGE1144152), the Alfred P. Sloan Foundation Pre-Doctoral Fellowship on the Economics of an Aging Workforce award from the NBER, and a Thomas Parry Research Fellowship award from the Integrated Benefits Institute. This project was also supported by grant number U19HS024072 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the author and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. I am grateful for the extensive support and guidance of my advisers: David Cutler, Claudia Goldin, and Timothy Layton. Marcella Alsan, Michael Barnett, Alex Bartik, Savannah Bergquist, Samantha Burn, David Card, Michael Chernew, Moya Chin, Edward Glaeser, Colin Gray, Jonathan Gruber, Nir Hak, Ryan Hill, Robert Huckman, Anupam Jena, Ariella Kahn-Lang Spitzer, Lawrence Katz, Victoria Marone, Thomas McGuire, Michael McWilliams, Hannah Neprash, Dev Patel, Jonathan Roth, Mark Shepard, Niharika Singh, Gabriel Unger, Scott Walker, Melanie Wasserman, and Annetta Zhou as well as numerous seminar and conference participants provided unrivaled support, advice, and suggestions. It truly takes a village.

Dissertation Advisor:
David Cutler

Author:
Adrienne Hope Sabety

Essays on the Institutional Design of Health Care Markets

Abstract

Do the relationships patients have with their doctors matter? I show that relationships determine where patients demand care and positively affect patients' health. In cases where a primary care provider (PCP) exogenously exits the market, patients do not form new PCP relationships, decreasing their use of primary care long-term. Instead, patients switch to specialists they know for less-extensive preventive care, which increases patients' probability of death by about 50 deaths per 100,000 individuals. Emergency department and inpatient admissions also increase for one year after a PCP's exit, increasing patients' spending by $4,640 and Medicare spending by $16,052 per exiting PCP. I establish the importance of relationship-specific capital in explaining effects, information that grows over time. Travel-cost revealed preference estimates suggest that patients are willing to pay at least $400-$500 per year to maintain a PCP relationship, providing a lower bound on patients' monetary valuation of relationship-specific capital.

1.1 Introduction

"It is more important to know what sort of person has a disease than to know what sort of disease a person has."–Hippocrates

Do the relationships patients have with their doctors matter? Many transactions, especially within the service sector, take place in the context of longstanding relationships, where relationships appear to be sticky and valuable to the parties in them. Examples range from client-lawyer and borrower-lender to patient-doctor relationships. Do relationships lead to the provision of a higher quality good when produced between two parties who know each other? Do relationships contain information that grows over time, the concept of relationship-specific capital (Jovanovic, 1979a)?

Relationships may be especially valuable in health care, where a bad outcome can be death. In particular, primary care providers (PCPs) are patients' first line of defense, providing preventive care, coordinating care across different providers, and referring patients to downstream care (Starfield, 1994). In spite of this, researchers still lack evidence for whether and to what extent PCP relationships matter to patients. On the one hand, PCPs provide a standardized good in a market characterized by extensive use of information technology and firms with far-ranging scope and scale.[2] Therefore, convenience, price, and quality may drive patient demand, for instance, in the form of retail clinics, price compare tools, and insurance networks. On the other hand, the market for primary care may be a function of patients' relationship with their PCP, and not simply the place, price, or quality of primary care in isolation.[3] For example, PCPs' knowledge of patients may grow over time, leading to higher-quality primary care when produced within the context of a relationship.

In this paper, I test if patients' demand for health care services as well as health

[2]For instance, Goldin and Katz (2016) show that patients view pharmacists as interchangeable, mainly in response to technological change, growth of pharmacy employment in retail chains and hospitals, and the related decline of independent pharmacies.

[3]A small number of papers have approached this question in specific contexts, like that between black patients and doctors (Alsan, Garrick and Graziani, 2019) or the relationship a obstetrician has with her patients (Johnson, Rehavi, Chan and Carusi, 2016).

outcomes are a function of their relationships with PCPs. I do this by implementing a quasi-experimental research design and estimating how Medicare patients are causally impacted by the retirement or relocation of a longstanding PCP. I identify the causal effect in a difference-in-differences research design, using a control group of PCPs who did not exit but are otherwise similar to exiting PCPs. For identification, I assume that PCPs depart for idiosyncratic reasons unrelated to the health of their patients. Outcomes of treated and control patients trend similarly 2-1 years before the exit, supporting this assumption.

I document how patients are impacted by the loss of a PCP, showing that relationships determine where patients demand care and are moderately important for patients' health. In response to the loss of a PCP, patients decrease their use of primary care by 17% for at least four years after the exit. Patients do not form new relationships. Rather, they receive preventive care from specialists with whom they already had a relationship. Receiving preventive care from a specialist, care that is generally outside their scope of practice, leads to less extensive preventive care, translating to an increase of about 50 deaths per 100,000 individuals. Emergency department (ED) and inpatient admissions also increase for one year after a PCP's exit, which increases patients' spending by $4,640 and Medicare spending by $16,052 per exiting PCP.

I interpret these results as clear evidence that patients value their relationship with PCPs and that relationships positively affect patients' production of health. I formalize the mechanism behind these results in a simple conceptual framework showing that relationship-specific capital grows over time, a concept akin to firm-specific capital first proposed by Jovanovic (1979a). I empirically test for relationship-specific capital by using plausibly exogenous variation in how clinics are managed, which affects patients likelihood of knowing other PCPs at the clinic and thereby their external stock of relationship-specific capital. Patients ability to rely on existing PCP relationships is lessened when the clinic closes when a PCP exits and higher in clinics managed as a team.

I show that patients who have less relationship-specific capital with other PCPs are more likely to not adhere to recommended preventive care, substitute to specialists for

primary care, and experience an adverse event. I show that adverse events are largest among patients who lose a clinic when a PCP exits, slightly smaller in open clinics where PCPs care for patients one-on-one, and small and insignificant in open clinics managed as a team.[4] However, even in team clinics, where PCPs are most interchangeable, patients significantly decrease their use of primary care and increase their use of specialists from the baseline level. This implies that relationship-specific capital is nontransferrable—that patients are negatively impacted by the loss of a relationship regardless of their external stock of capital.

To provide a lower bound on patients' monetary valuation of relationship-specific capital, I focus on cases where patients can choose to follow a PCP who moves to a new clinic or continue to seek care with a replacement PCP at the original clinic. When a PCP moves to a new clinic within 300 miles of the original clinic, 38% of patients follow the PCP.[5] This suggests that patients value the relationship around $400-$500 per year in back-of-the-envelope calculations using patients willingness-to-drive to follow their PCP.

To ensure my interpretation of the empirical results is valid, I explore four other explanations for why patients find the relationship valuable, showing that alternative explanations do not match the evidence. First, patients may decrease their use of primary care because they are unable to find a replacement PCP that is a similarly good match (Jovanovic, 1979b). I show that primary care use decreases similarly among patients with more specific needs, i.e. high risk, disabled, dual eligible, minority as well as female patients with female PCPs. Therefore match quality along these dimensions does not appear to be a main mechanism. Second, replacement PCPs may be hard to find. I do not find support for this by showing that patients in thinner and thicker markets decrease their use of primary care similarly. Third, the loss of a PCP may overwhelm staying PCPs at the main clinic by increasing their workload. I find that staying PCPs *do* take on more patients after a PCP's exit, but that the outcomes of staying PCPs' patients are not affected. Further, patients in

[4] Adverse events include non-preventive ED visits, inpatient admissions, or death.

[5] The number of patients moving with PCPs is especially large considering that 39 states enforce non-compete agreements, which legally prohibit PCPs from taking their patients with them when moving practices (Hausman and Lavetti, 2016).

smaller clinics are only slightly more likely to substitute away from the main clinic, while the increased workload for PCPs who work in small clinics is relatively larger. This suggests that the increased workload does not affect staying PCPs' ability to care for patients.[6] Fourth, I show that replacement PCPs do not refer to specialists at a higher rate than departing PCPs. Therefore, practice pattern differences between replacement and leaving PCPs do not explain why patients decrease their use of primary care and increase their use of specialty care after a PCP's exit.

The paper proceeds as follows. Section 2 outlines the institutional setting. Section 3 presents a simple conceptual framework that motivates the two main mechanisms. Section 4 describes how the data is constructed. In Section 5, I describe my empirical strategy and identification assumptions. Section 6 presents the main results. Section 7 explores the specific mechanisms behind the aggregate results. Section 8 concludes.

1.2 Institutional Setting: Primary Care

Primary care is widely considered a basic input into patients' health. PCPs specialize in administering primary care, such as immunizations, preventive screens, and medications. PCPs are also patients' first point of contact to the health care system and are tasked with coordinating patient care across specialists (Starfield, 1994).

The role of a PCP is especially relevant for Medicare patients who are in worse health and are more likely to have multiple providers, as compared to younger patients. For instance, the average Medicare patient sees 1.5 PCPs (median 1) over 6 different visits and

[6]These results contribute to work showing that outside, replacement workers are imperfect substitutes for incumbent workers (e.g. Jäger (2017); Stole and Zwiebel (1996)). I extend this literature in two ways. First, I illustrate that, while the workload of incumbent workers increases in response to the loss of a co-worker, this does not observably affect incumbent workers' productivity. Second, I show that, even if a firm can perfectly substitute between workers, clients themselves may not view workers as substitutes because of the existence of the relationship. This is of particular relevance to transactions that contain large amounts of asymmetric information.

3.8 specialists (median 3) over 9 different visits per year.[7,8] PCP-patient relationships also tend to be long-lasting: the average Medicare patient over age 75, has known their PCP for 4.7 years (median 6).

As a result, the loss of a PCP may significantly change how Medicare patients' interact with primary care as well as the medical system more broadly. The loss of a PCP may create a discontinuity in care, which may adversely affect patients' health (e.g. Saultz and Albedaiwi (2005); David and Kim (2018); Agha, Frandsen and Rebitzer (2017)). Further, Medicare patients face similar out-of-pocket costs across settings and can see any provider at any time (no networks). Therefore, patients may benefit from specialists taking on routine care, like signing existing chronic medication prescriptions and flu vaccines.

The relevance of PCPs as an integral part of the health care delivery system relates to concerns about the growing number of specialists relative to PCPs in the physician workforce. As an increasing number of primary care physicians retire, new medical graduates overwhelmingly choose to specialize instead of going into primary care (Sabety, 2020; Whitcomb and Cohen, 2004)).[9] Adding to pressures facing the PCP workforce, new delivery models incentivize PCPs to take responsibility for the continuum of patient care, requiring PCPs to invest more in patients (e.g. Sessums, Basu and Landon (2019); McWilliams, Landon, Rathi and Chernew (2019); Sinaiko, Landrum, Meyers, Alidina, Maeng, Friedberg, Kern, Edwards, Flieger, Houck, Peele, Reid, McGraves-Lloyd, Finison and Rosenthal (2017)). As such, there has been near-consensus that more PCPs are needed (Chernew, Sabik, Chandra and Newhouse, 2009).

The delivery of primary care is changing in response to these increased demands. Tasks typically done by PCPs are shifting to other stakeholders. It is increasingly the

[7]In contrast, Pham, Schrag, O'Malley, Wu and Bach (2007) uses slightly different Medicare data and sample restrictions and finds that patients see two PCPs and five specialists on average.

[8]PCPs are within a 30 mile radius of each other for 96% of patients that have more than one PCP. The remaining 4% may have multiple residences and have a PCP in each location. For instance, if an individual spends 8 months of the year in Michigan and 4 months of the year in Florida, they may have a PCP in each location.

[9]I make the distinction between primary care physicians and PCPs, which are defined to include nurse practitioners, physician assistants, and primary care physicians who practice in the primary care setting.

case that primary care, urgent care, retail, and specialty *clinics* are patients' first points of contact (Friedberg, Hussey and Schneider, 2010). Further, many clinics have created roles for dedicated care coordinators, with at most a nursing background, who specialize in managing patient care across settings (Bayard, Calianno and Mee, 1997). Thus, with a robust network of care surrounding established patients, the role of the PCP-patient relationship, and under what conditions it should be prioritized, remains unclear.

1.3 Conceptual Framework

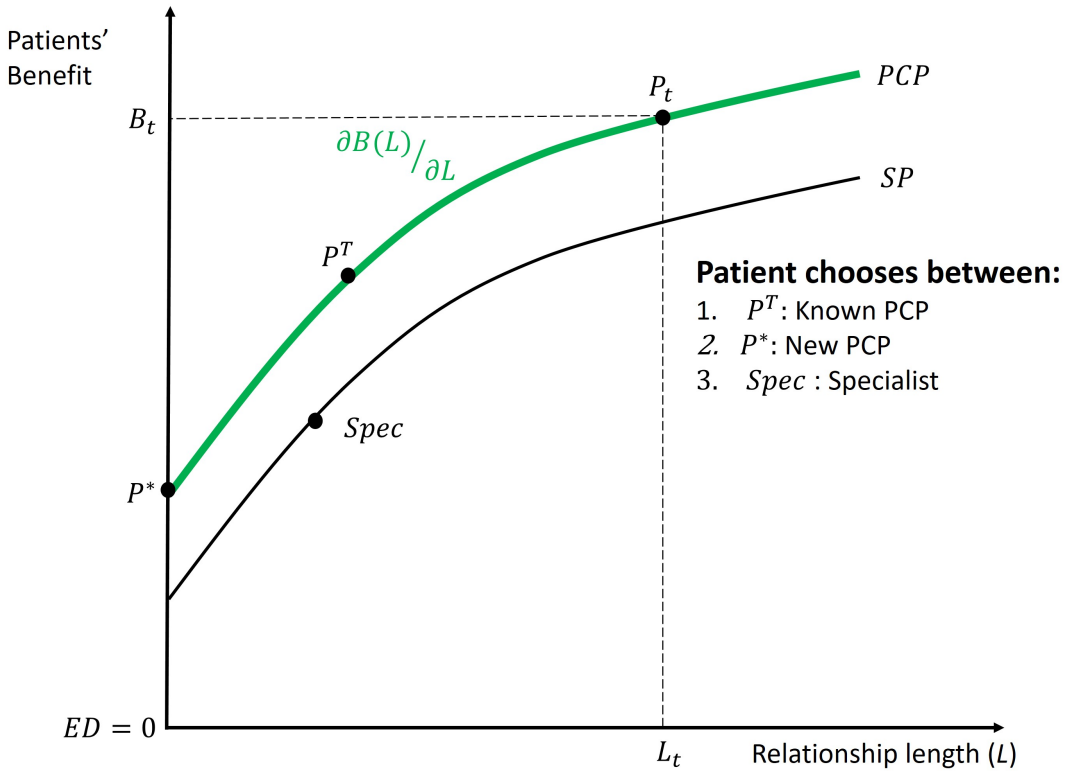
The following conceptual framework motivates why primary care markets may be a function of patients' relationship with their PCP, and not simply the place, price, or quality of primary care in isolation. It illustrates that patients receive more benefit from relationships over time, the concept of relationship-specific capital, and that this information is specific to a particular relationship.

The framework's intuition is displayed graphically in Figure 1.1, which shows the benefit a patient receives visiting a PCP, specialist, and the ED over the length of a relationship (L). I assume there is no variation in quality among different PCPs and specialists. Therefore, all PCP options are captured by PCP and all specialist options are captured by SP, where each unique relationship has its own curve.[10]

I begin by considering the benefit patients receive consuming care across options. By revealed preference, patients' benefit maximizing choice is to visit their initial PCP in $t = 0$ (P_t on Figure 1.1). Point P^T is the benefit patients receive visiting a PCP they know, where $L_T < L_t$. Point P^* is the benefit patients receive visiting a new PCP at $L = 0$. $Spec$ is the benefit a patient receives from visiting a specialist they have known for $L = L_{Spec}$. The specialist curve, $Spec$, is below the PCP curve, PCP, because providing primary care is outside of a specialist's scope of practice. I assume that the PCP and $Spec$ curves have the same slope—if patients receive benefit from having a relationship with a PCP, they are

[10]This framework can easily be generalized to the case where providers have different levels of quality by allowing the curves to shift based on a providers average quality level.

Figure 1.1: *Benefit Received by Patients from Providers*



Notes: Figure 1.1 contains the length of the relationship (L) on the x-axis and the benefit a patient receives consuming care with a PCP and specialist on the y-axis. The graph illustrates how patients develope relationship-specific capital with PCPs and specialists over time, $\frac{\partial U(L)}{\partial L}$. Patients start at P_t with their original PCP where the length of PCP-patient relationship is L_t and receive benefit B_t. After a PCP's departure, patients can choose between P^T, P^*, and $Spec$. P^T is the benefit a patient receives visiting a PCP they have a relationship with, where $L_t > L_T$. P^* is the benefit a patient receives visiting a new PCP, where $L = 0$. $Spec$ is the benefit received visiting a specialist who the patient has known for L_{spec}.

assumed to also benefit from having a relationship with a specialist. ED is the benefit a patient receives visiting the ED, which is normalized to zero and does not change with the length of the relationship L.

Where a patient chooses to receive care after a PCP's exit depends on the slope of the PCP curve, $\frac{\partial B(Length)}{\partial Length}$. By observing how patients are impacted by the loss of a PCP, I am able

to empirically determine the sign of the slope of PCP. For instance, if relationship-specific capital exists and grows over time, $\frac{\partial B(Length)}{\partial Length} > 0$, as shown in Figure 1.1.

The theoretical framework deliveres three testable predictions concerning relationship-specific capital:

1. Relationship-specific capital exists and grows over time ($\frac{\partial B(Length)}{\partial Length} > 0$). After the loss of a longstanding PCP, patients will visit a PCP they have a relationship with (P^T) over establishing a new PCP relationship (P^*). Patients will only visit a specialist over establishing a new PCP relationship (P^*) if patients have an existing relationship with a specialist ($Spec$).

2. Relationship-specific capital contains health-specific information. Patients who do not have known PCPs, outside of their relationship with the exiting PCP, should be most likely to experience an adverse event, such as visiting the ED for urgent conditions, an inpatient admission, and death.

3. Relationship-specific capital is nontransferrable. Patients should be affected by the loss of a long-term relationship regardless of their ability to rely on other relationships.

A challenge in testing these predictions is finding variation in patients' likelihood of having outside PCP relationships that is not correlated with patients' health. To overcome this, I leverage heterogeneity across three different clinic environments: clinics that close when a PCP exits, open clinics where PCPs see patients one-on-one, and open clinics that practice as a team. If relationship-specific capital drives results, effects should monotonically decrease as patients are better able to rely on a replacement PCP they know.

1.4 Data Construction

1.4.1 Data Sources

Analyses use a 20% sample of Medicare patients from 2008-2017. The data contains about 11 million patients and one million providers. The Medicare data captures all patient

encounters with the health system paid by Medicare. Claims start when patients become eligible for Medicare (typically age 65) and end when patients die or enroll in Medicare Advantage (MA). All billed care delivered to patients is observed, such as outpatient care, inpatient care, and prescription fills. The data also includes a rich set of patient demographics, such as sex, age, race, zip code of residence, and whether or not the patient is also covered by Medicaid.

Identifying clinics, also known as the doctors' office, is challenging. Medicare data does not contain a clinic identifier, so I construct my own by combining the tax identification number (TIN) and nine-digit zip code (ZIP) associated with the claim. The drawback to this definition is that different clinics located within the same facility may be considered the same. I therefore exclude clinics with over 100 providers because larger "clinics" are more likely to include multiple clinics.[11] This restriction also increases the probability that the treatment group has a common support. Larger clinics are mechanically more likely to have departures, so large control clinics without departure are rare. For example, almost all clinics with over 100 PCPs experience a departure within a three year window. I also exclude non-US clinics because the institutional context likely differs.

As for health providers, Medicare data is ideal for answering questions surrounding the health care workforce. It contains a nearly nationally representative sample of clinics and providers: 93% of American PCPs accept Medicare (Boccuti, Fields, Casillas and Hamel, 2015). Further, all providers are uniquely identified by their National Provider Identifier (NPI), a universal index used to submit billing claims.

Using the NPI, I supplement the Medicare claims with information on providers from four other data sources: the National Plan and Provider Enumeration System, Doximity, Medicare's MD-PPAS, and Physician Compare. I identify providers' specialty, sub-specialty, sex, age, type of training, and whether NPIs belonged to individuals or organizations using these additional data sources.[12]

[11] This excludes large organizations like the Cleveland Clinic, Kaiser, and Intermountain.

[12] Analyses drop NPIs associated with an organization because it is unclear if a provider left during the period of interest. Table A.1 illustrates the algorithm used for each year of data. Combining years, I resolve

1.4.2 Primary Variable Construction

My identification strategy relies on being able to accurately define when a PCP leaves a clinic. I define a **departure** to occur when a PCP fully disappears from the data or relocates. Full disappearances occur at an average age of 59 (median 61), suggesting that these exits are likely retirements. Relocations occur when a PCP moves to a new clinic where the first three digits of the zip code differ (e.g. out of state).[13] I also identify cases when a PCP's departure is followed by the main clinic closing, which occurs in 27% of cases. (See Appendix A.1.1 for a more detailed definition of "clinic closure.")

To compare treatment effects by the **length of the PCP-patient relationship**, I supplement years 2008-2017 in the main analyses with years 2002-2007 to identify relationship length.[14] The length of the relationship is measured from the first time I observe the patient seeing the provider until 36 months before a PCP's departure, where any visit within a year is a point of contact.[15] I drop patients who are less than age 75 as of $t = -36$ when I compare treatment effects by the length of the PCP-patient relationship. I do this to circumvent left censoring due to only observing patients when they become Medicare eligible at age 65. The restricted sample of patients is on average 81 years old, or 10 years older than patients in the main sample.

Whether a patient sees one or more than one PCP is confounded with patients' health status: sicker and older patients are more likely to see multiple providers than healthier and younger patients. Therefore, defining whether a clinic is managed to practice as a **team or individually** relies on a definition that is plausibly orthogonal to patient observables. I

mismatches by taking the modal value across years.

[13]This occurs when the first three digits of the new clinic's ZIP are different from the original ZIP.

[14]I use 2008-2017 for main analyses because Medicare data uses UPINs, instead of NPIs, to represent PCPs from 2002-2007. I derive a UPIN to NPI crosswalk to connect NPIs used in my main analyses to their corresponding UPIN in pre-2008 years.

[15]Patients are categorized as having a long relationship regardless of the frequency of the interaction between the first and last point of contact, potentially biasing estimates towards zero. For instance, if a patient interacted with a PCP once 10 years before visiting them again from $-36 \leq t \leq -24$, I would categorize their relationship as lasting 10 years. This is a more tenuous link than a patient who has seen their PCP yearly for 10 years.

define management practices at the clinic level using all patients who visit a clinic within the first 12 months the clinic is observed in the data. I then restrict to patients who had three evaluation and management (E&M) visits over this 12 month period and categorize whether the three visits were with the same PCP. I then take the average rate over the clinic of how many patients exclusively saw one PCP, calling clinics above the average "individual clinics" and those below the average "team clinics." I find that 59% of clinics were on individual models, where 58% of patients were seen by the same PCP for their first three E&M visits.

I also categorize the **local density of PCPs** to define the thickness of the local PCP market. The local density of PCPs is defined as the number of PCPs filing billing claims within a 30 mile radius of each focal clinic ZIP divided by the population.[16]

I use three different patient level definitions: **risk**, **disability**, and **racial** status. First, I define a patient's risk level using their calculated Elixhauser Index, which creates an index based on commorbities and pre-existing conditions that are predictive of death by scanning over a patient's ICD-9 diagnosis codes. I use diagnosis codes three years before the departure to circumvent the potential endogeneity of patient outcomes to treatment.[17] High risk patients are defined to be patients with the top quartile of Elixhauser scores within a PCP's pool of patients. Second, disabled individuals include patients with spinal cord injuries, blindness, mobility impairments, muscular dystrophy, chronic pain fatigue/fibromyalgia, spina bifida, multiple sclerosis, and cystic fibrosis. The three largest groups are chronic pain fatigue/fibromyalgia, mobility impairments, and visual impairments. Third, I classify a patient's racial status based on the race variable in the Master Beneficiary Summary File.

Main outcomes include utilization of clinic based services and quality of care. Utilization includes primary care visits, specialist visits, and urgent care visits.[18] An increase or decrease in these outcomes has an ambiguous effect on patient health. I classify if visits are by new or

[16] A 30 mile radius was chosen because patients travel 17 miles on average to their assigned clinic, which was defined to be patients' modal clinic. The distance between five digit zip codes were great-circle distances calculated using the Haversine formula based on internal points in the geographic area. The data set was obtained from the NBER at `https://www.nber.org/data/zip-code-distance-database.html`.

[17] I then aggregate these individual scores, which I use as my measure of riskiness.

[18] I use the outpatient files to identify urgent care visits.

existing patients using evaluation and management (E&M) billing codes. E&M visits include annual exams, wellness visits, physician exams, and consultations.[19] A major advantage to using E&M codes is that patients are only considered new if they have not seen that physician or another physician of the same specialty in the clinic within the previous three years. This allows me to more cleanly identify the demand side (patients seeking out new relationships) opposed to supply side changes (patients being transferred to replacement PCPs within the clinic).

Of all primary care visits, about 75% are for E&M visits and, of these, 96% are for existing patients and 4% are for new patients. Comparatively, E&M visits are much less common in the specialty setting. Specialists bill for E&M visits 24% of the time and, of these, 12% are for new patients and 84% are for existing patients. Therefore, E&M codes billed by specialists may not capture the complete picture, whereas those billed by PCPs likely do.

Quality of care metrics include adverse events, preventive care, and medications. Adverse events include death as well as ED visits for non-preventable conditions and inpatient admissions.[20] Preventive care includes influenza (flu) vaccines, annual exams, and preventive screens.[21] Preventive screens include: mammography screens, colorectal cancer screens, cholesterol screens, and diabetes screens.[22,23] The medication category separately quantifies

[19]New patients are patients who did not receive any professional services from the physician (or non-physician) or another physician of the same specialty in the same group practice within the previous 3 years. Existing patients are individuals who received care from the physician (or non-physician) or another physician of the same specialty in the same group practice within the previous three years. CMS' coding rules can be found here: https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/eval-mgmt-serv-guide-ICN006764.pdf

[20]An algorithm by Billings, Parikh and Mijanovich (2000), updated by Johnston, Allen, Melanson and Pitts (2017), classifies ED visits.

[21]Obtaining a yearly flu vaccine is considered a key input into patient health, especially among the elderly. In spite of this, many patients do not receive a yearly flu vaccine. For instance, the Centers for Disease Control and Prevention reported that 59.6% of adults over age 65 received a flu vaccine during the 2017-2018 flu season (CDC, 2018).

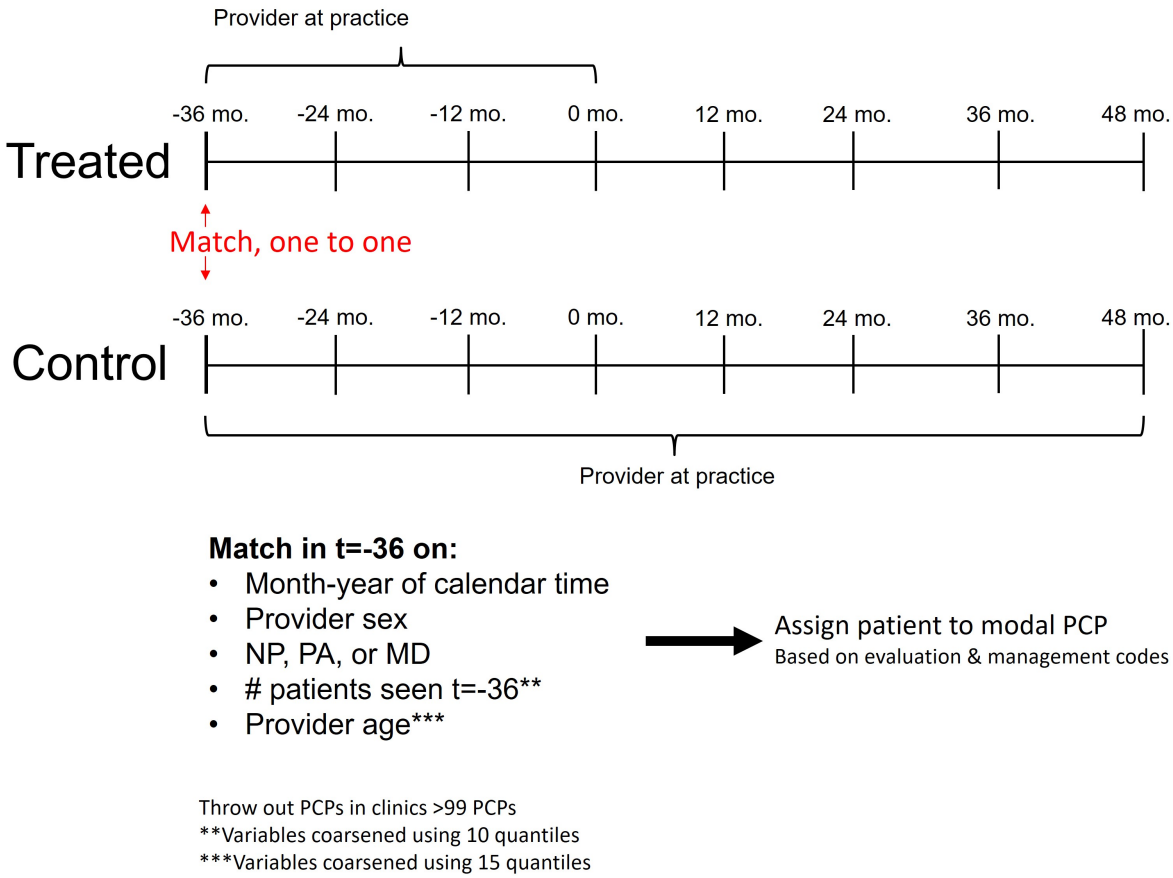
[22]The rate of mammograms was categorized as the total number of mammograms within a PCP's pool of patients divided by the number of *women* in that PCP's patient pool.

[23]Preventative care was identified from the carrier file's Health care Common Procedure Coding System codes based on a crosswalk used by Centers for Medicare and Medicaid Services to categorize quality scores for Accountable Care Organizations (ACO) in the domain of preventative health. https://www.cms.gov/Medicare/Prevention/PrevntionGenInfo/

the total number of medications as well as only chronic medications (about 40% of all medication prescriptions). See Appendix A.1.1 for details.

1.4.3 Sample Restrictions

Figure 1.2: *Data Restriction and Matching Strategy*



Notes: "Tx" are departing PCPs and "Ct" are staying PCPs. **Indicates that the variable was coarsened using 10 quantiles. ***Indicates that the variable was coarsened using four quantiles. I restrict to departing PCPs who were at the clinic at least three years prior to exit and control PCPs who exist in the data for at least four years, 3 years before exit and 1 year. Different subsamples place different restrictions on the length of time staying PCPs have to be observed at the clinic. A sample is also used that matches clinics, which is described in Appendix A.2.

`medicare-preventive-services/MPS-QuickReferenceChart-1.html`

Figure 1.2 illustrates the creation of the PCP sample. The sample was constructed to avoid three main issues. First, I use a 20% sample of Medicare data. It is therefore possible to miscategorize a PCP as departing if the PCP's patients were not sampled. To get around this, I restrict to PCPs who saw more than 30 total patients from $-36 \leq t < -24$.[24] Once I condition on seeing 30 patients, treatment PCPs see on average 130 patients (median 88) and control PCPs see on average 156 patients (median 115) from $-36 \leq t < -24$. This sample yields a departure rate of 12% from 2011 to 2015 (Table A.2). See Appendix A.1.2 for more details on departures.

Second, compositional changes in the number and types of patients seen around a PCP's departure may affect identification. For instance, PCPs may transfer sicker patients to replacement PCPs before healthier patients (or vice versa). To address this, I use a sample of PCPs practicing three years before the departure. I then assign patients to PCPs 2-3 years before the departure to circumvent the endogeneity of visits around the departure. To create a parallel control sample, I restrict to control PCPs in the data for four years, including three years before and one year after the departure. Third, I construct relative time for the control group by matching control and treatment PCPs in month $t = -36$, or 36 months before a PCP's exit.[25]

1.5 Empirical Strategy

I implement a difference-in-differences (DD) design where I match 10,437 PCPs who left a physician group in a given month-year to a comparison group of 10,437 PCPs with similar lagged characteristics, that did not retire or relocate. I then analyze the effect of the departure using 627,647 patients associated with either control or treatment PCPs.

[24] 30 was chosen because treatment PCPs see 67 total patients on average (median 26) and control PCPs see 32 total patients on average (median 0) over my *unrestricted* primary care sample. Note, the reason treatment and control PCPs have such different patient loads over the *unrestricted* primary care sample is because there are a lot of PCPs in my sample that see very few patients.

[25] Control PCPs practicing in clinics experiencing a departure do not enter the control sample.

1.5.1 PCP Matching Procedure to Select Comparison Group

Departing and staying PCPs may be different along at least three dimensions. First, Figure A.1a shows that departing PCPs are 0.4 years older than staying PCPs on average. Older PCPs may be different from younger PCPs: they may have older patients or have specific practice styles. Second, the rate of departure as well as practice styles may differ by whether the PCP is a nurse practitioner (NP), physician assistant (PA), or medical doctor/doctor of osteopathy (MD) as well as the gender of the PCP. Third, the physician workforce is in the midst of a burn out crisis and a growing number of PCPs are leaving clinical practice (Sabety, 2020).

To adjust for these differences, I match observably similar treatment and control PCPs one-to-one, three years before the departure. I match exactly on month-year of calendar time, which enables me to derive a relative time measure for control PCPs and their associated patients. I also match exactly on PCP sex and the type of PCP (i.e. NP, PA, or MD) as well as four coarsened bins of PCP age in $t = -36$ and 10 coarsened bins of the number of patients seen. This matching procedure follows the standard in the literature.[26]

I intentionally match on time invariant covariates to address the potential for mean reversion. The one exception to this is the number of patients seen in $t = -36$, which is a lagged covariate. Matching on the volume of patients is important because PCPs who see large volumes of patients may be more likely to leave practice than PCPs who see fewer patients, which in turn could affect patient outcomes.

I then assign patients to PCPs based on their modal number of E&M visits 3-2 years before the departure. I use E&M visits for the assignment in order to isolate visits made to a patient's PCP as opposed to secondary staff.[27,28] A similar procedure was used to define

[26]I use coarsened exact matching (CEM) following recent literature (e.g. Jäger, 2017; Jaravel, Petkova and Bell, 2018; Sarsons, 2017; Azoulay, Zivin and Wang, 2010). Recent work argues that CEM is more transparent and interpretable than other approaches, such as propensity score matching (e.g. King and Nielsen, 2016; Iacus, King and Porro, 2012). In determining the number and coarseness of matching covariates I minimize the number of covariates while maintaining balance to maximize the number of successful matches.

[27]Section 1.4.2 provides additional details on E&M visits.

[28]This allocation method follows the standard in the literature (Pollack, Hussey, Rudin, Fox, Lai and

the clinic level sample, which is detailed in Appendix A.2.

1.5.2 Summary Statistics

Table 1.1 describes summary statistics for PCPs and patients by treatment status. The goal of the matching process is to create a balanced comparison group. The difference-in-differences design then absorbs average levels of outcome variables between control and treatment groups, relying on a common-trends assumption (see Section 1.5.3). The illustration of summary statistics confirms that the matching procedure created a balanced comparison group for the difference-in-differences design. Further, it provides context when interpreting treatment effects.

Match Rate. The first section of the table illustrates the match rate and resultant sample size. Of treated PCPs that meet the sample restrictions, 90% are matched, or 10,437 control and treatment PCPs. This translates to 298,943 treated patients and 328,704 control patients, or 57,914,522 PCP-patient-time observations. Table A.3 illustrates that the strongest restriction is exactly matching on whether or not the PCP was a NP, PA, or MD.

PCP Matching Covariates. The second panel shows that the matching covariates defining the coarsened bins are fairly balanced. Treated patients are 0.4 years older than control PCPs, which does not appear to be an economically meaningful difference. Both treated and control PCPs see 12.2 patients in $t = -36$. In addition to PCP age and caseload, which are displayed in Table 1.1, the type of the PCP (NP, PA, or MD), PCP gender, and month-year of calendar time are exactly matched on.

Characteristics of Patients by Treatment Status. The third section of the table shows that treated and control patient characteristics are broadly similar, although there are some differences. Treated patients are more likely to be white and live in rural areas. They are also slightly healthier, being less likely to have end stage renal disease and be enrolled in Medicaid. Treated patients also see slightly more PCPs (1.6 v. 1.1) and specialists (3.4 v. 2.7).

Schneider, 2016). If a patient is assigned to a PCP that is not in the matched sample, that patient is not included in analyses. If I assign away all of a PCP's patients, I not only drop that PCP from my sample, but I also drop the matched pair.

Table 1.1: *Balance Table for Exiting PCPs and Their Patients*

	Treatment	Control	P-Value
% Match	90%	29%	
No. of PCPs	10,437	10,437	
No. of Patients	298,943	328,704	
No. of Observations	28,957,261	28,957,261	
PCP Matching Covariates, 3 Years Before PCP Exit			
PCP Age (yr)	53.8	53.4	0.005
Caseload per PCP/Month	12.2	12.2	1.00
Patient Covariates That Were Not Matched On, 3-2 Years Before PCP Exit			
Patient Demographics			
Patient Age (yr)	71.1	71.5	$p< 0.001$
White (%)	85.6	81.7	$p< 0.001$
Female (%)	37.0	36.9	0.94
Urban (%)	79.2	83.8	$p< 0.001$
Patient Clinical Characteristics			
Elixhauser Risk Score	2.5	2.6	$p< 0.001$
End Stage Renal Disease (%)	0.91	1.1	$p< 0.001$
Also Enrolled in Medicaid (%)	19.4	20.7	$p< 0.001$
Average Annual Rate per Patient			
No. of Primary Care Visits	6.1	6.2	$p< 0.001$
No. of PCPs Seen	1.6	1.1	$p< 0.001$
No. of Specialty Care Visits	9.7	9.9	$p< 0.001$
No. of Specialists Seen	3.4	2.7	$p< 0.001$
No. of Emergency Department Visits	0.70	0.65	$p< 0.001$
No. of Urgent Care Visits	0.018	0.014	$p< 0.001$
No. of Inpatient Department Visits	0.39	0.39	0.33
Annual Spending ($)	8433.98	8827.97	$p< 0.001$
Prob. of Death (%)	0.0032	0.0036	$p< 0.001$

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' clinic locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending includes provider charges from the carrier file, inpatient charges, and outpatient charges.

This suggests that the loss of a PCP may impact treated patients slightly less than control patients, had they also lost a PCP.

1.5.3 Estimating Equations and Identification

I estimate the causal impact of a PCP's departure on patient outcomes using an event study, difference-in-differences design. In particular, equations are of the form:

$$y_{jt} = \rho_{m(j)} + \sum_{\tau=-24}^{12} \beta_\tau \times \mathbf{1}(t=\tau) + \sum_{\tau=-24}^{12} \beta_\tau^{Treated} \times \mathbf{1}(t=\tau) \times Treated_j + \epsilon_{jt} \qquad (1.1)$$

where y_{jt} denotes the average outcomes over PCP j's pool of patients in relative time t. $\mathbf{1}(period_t)$ includes relative time t fixed effects. $\rho_{m(j)}$ are PCP fixed effects, which absorb average differences across PCPs.[29] I cluster the standard errors at the pre-departure PCP-match level to account for idiosyncratic factors that are specific to a matched pair. This assumes that each matched pair's errors are uncorrelated with other matched pairs. As outlined in Section 1.4.3, I restrict the data to support the plausibility of this assumption. Identification is based on comparing outcomes within a PCP's group of patients to the matched control PCP's pool of patients, relative to the omitted group.

For *event study graphs* that follow patients one year post-departure, t is the month-year relative to the departure at $t = 0$. The coefficient of interest, $\beta_t^{Treated}$, captures the effect of a departure in month-year t and is normalized to zero in $t = -24$.

Regressions are at the relative year level. All estimates are relative to $24 \leq t < -12$.[30,31] This allows for anticipation up to 12 months before a PCP's exit. Visually inspecting event study graphs supports this assumption. The main results use a sample that follows patients

[29] As a robustness check, I also estimate main results with PCP-match fixed effects, which leads to virtually identical standard errors. I additionally estimate the main specification clustering at the firm level, yielding similar results.

[30] I start at the patient level and sum over patient i's y_{ijt}s within a PCP's pool of patients. I then normalize this outcome by the number of patients in each PCP's pool, summing over the monthly PCP level averages of each outcome to get the yearly rate. To obtain estimates representative of the original patient population, frequency weights are used in all regressions. Regressions are run at the PCP-year level.

[31] Specifications using more than one post year include a dummy for each year relative to treatment, interacted with treatment status.

for one year post-departure unless otherwise specified.

1.6 Aggregate Impact of the Loss of a PCP on Patients

The loss of a PCP has a significant impact on patient care. I start by describing how patients' visits to main PCPs and clinics change after a PCP's departure. I then show how patients are impacted by the loss of a PCP on aggregate. I explore the specific mechanisms behind these results in Section 1.7.

Figure 1.3a and 1.3b illustrate the identifying variation. Graphs plot the average number of visits to assigned PCPs and clinics over relative time t for one year post-departure. For instance, if the average number of primary care visits per month is 0.3, patients visit their main PCP slightly more than three times a year. Blue triangles represent control patients and red crosses represent treated patients. Patients are assigned to PCPs from $-36 \leq t < -24$, $t = -24$ marks the start of the treatment period, and $t = 0$ marks the last month exiting PCPs see patients.[32] Treatment and control patients do not visit their main PCP at the same rate because treated patients see slightly more PCPs than control patients (1.6 vs. 1.1 PCPs), seeing the main PCP slightly less (Table 1.1).

Figure 1.3a shows that after a PCP exits in $t = 0$, patients no longer see that PCP for primary care. Treated and control patients see assigned PCPs at the same rate, as illustrated by the curves moving in parallel. Both curves slope downward due to mean reversion and patients dying over time. Curves begin to separate around eight months before the departure implying that exiting PCPs see fewer patients leading up to the departure. As a result, I estimate event studies relative to $t = -24$ and regressions relative to $-24 \leq t < -12$ to allow for anticipation up to -12 months before the departure. Figure 1.3a shows that visits decrease from 0.24 visits at $t = -1$, to 0.17 visits at $t = 0$, to 0 at $t = 1$ in response to a PCP's exit. Visits are not zero in $t = 0$ because PCPs stop seeing patients at various times

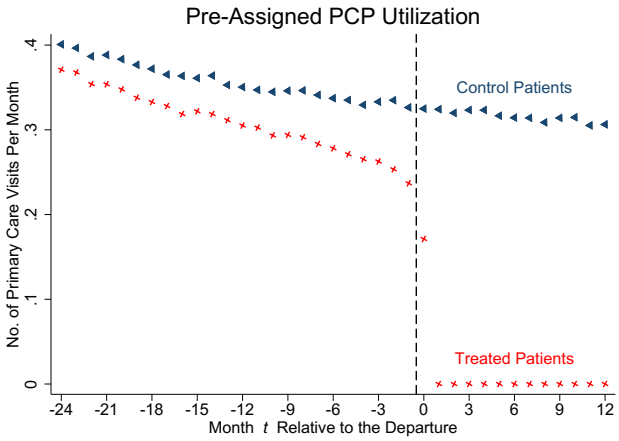
[32] Treatment and control PCPs are matched in $t = -36$ and patients are assigned to the PCP (and clinic) that provided the majority of their primary care from $-36 \leq t < -24$. The first month in the post-departure period is $t = 0$

during the month.

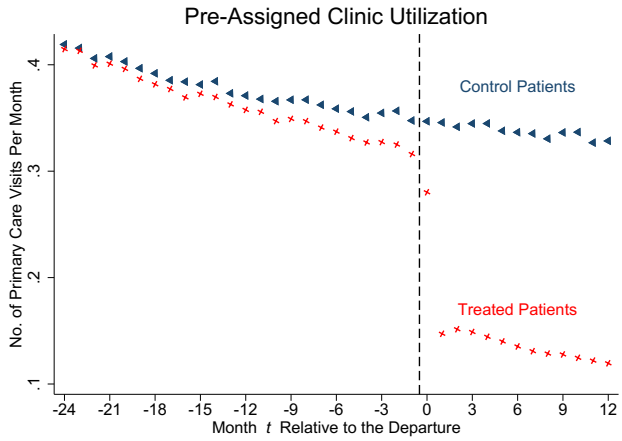
Figure 1.3b graphs the number of visits a patient makes to the main clinic over time. The graph is similar to Figure 1.3a except slightly closer in terms of levels. In response to a PCP's departure, patients' decrease their rate of visiting the main clinic 47%, or from 0.32 visits in $t = -1$ to 0.15 visits in $t = 1$. For more context, Figure 1.3c graphs the number of PCPs *per clinic* two years before and after a PCP's departure. It shows that clinics slowly replace PCPs and that the replacement rate is not one-to-one.[33]

[33]This graph is based on a sample that matches clinics (instead of PCPs), which is used for clinic level analyses in Section 1.8.3. See Appendix A.2 for more details on the clinic level sample.

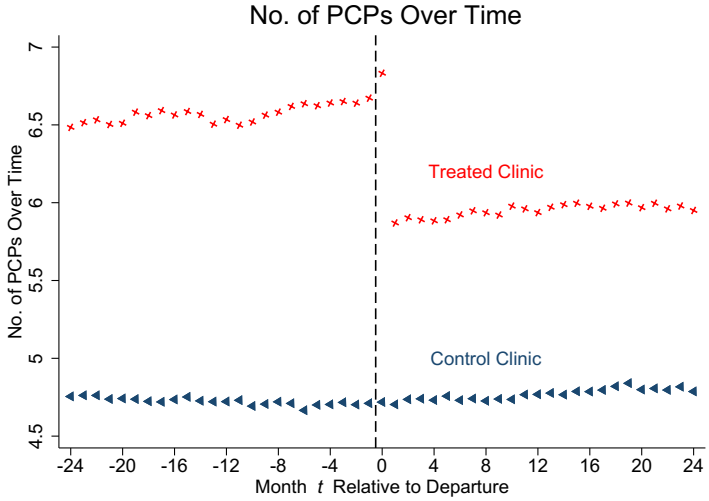
Figure 1.3: *Number of Visits Patients Make to the Assigned PCP and Clinic*



(a) *Number of Primary Care Visits by Patients to Assigned PCPs*



(b) *Number of Primary Care Visits by Patients to Assigned Clinics*



(c) *Total Number of PCPs at the Clinic Over Time*

Notes: Figure 1.3a and Figure 1.3b plot the number of primary care visits by patients to assigned (a) PCPs and (b) clinics over relative time t. Graphs show that the loss of a PCP is a shock to patients. The y-axis starts in relative time -24 and ends in relative time 12, or 1 year post-departure. The underlying sample matches leaving to staying PCPs in $t = -36$. Patients are assigned to PCPs and clinics from $-36 \leq t < -24$ based on where the majority of their primary care was provided. Primary care visits among patients of staying PCPs are reflected in blue triangles and patients of exiting PCPs are reflected in red crosses. Figure 1.3c uses a sample that matches *clinics*, instead of PCPs, which is described in Appendix A.2. The graph shows the total number of PCPs observed filing Medicare claims in relative time t among clinics that stayed open post-departure. The y-axis starts in relative time -24 and ends in relative time 24, or 2 years post-departure. It shows that exiting PCPs are not replaced one-for-one.

1.6.1 How the Loss of a PCP Affects Patient Outcomes

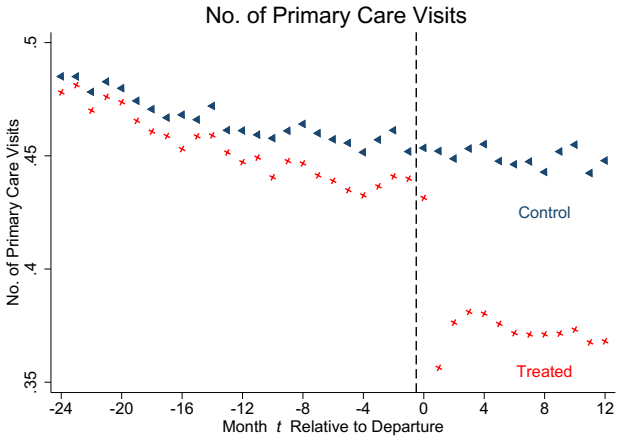
The raw data in Figure 1.4 shows how patients are impacted by the loss of a PCP. The x-axis contains time relative to the departure in $t = 0$ measured in months. Red crosses represent treated patients and blue triangles represent control patients. Figure 1.4a shows the number of primary care visits per month in a sub-sample that follows patients one year post-departure. Figure 1.4b shows the probability a patient forms a new relationship as a cumulative hazard rate. Figure 1.4c graphs the probability a patient forms a new relationship. Figure 1.4b and Figure 1.4c use a sub-sample that follows patients four years post-departure.

Figure 1.5 plots coefficients $\beta_t^{Treated}$ from equation 1.1 at the quarter level. Coefficients are only identified up to a constant term, so $t = quarter = -8$ is normalized to zero. Primary care visits (black line, with triangle points) and specialty visits (blue line, with circular points) are dependent variables in Figure 1.5a. The number of chronic medications prescribed by PCPs (black line, with triangle points) and specialists (blue line, with circular points) are dependent variables in Figure 1.5b. Figure 1.5c plots the number of ED visits and Figure 1.5d plots the number of ED visits for primary care treatable conditions. Lending credibility to the research design, there is no significant pre-trending from $-8 \leq t < -4$ for any of the outcome variables.

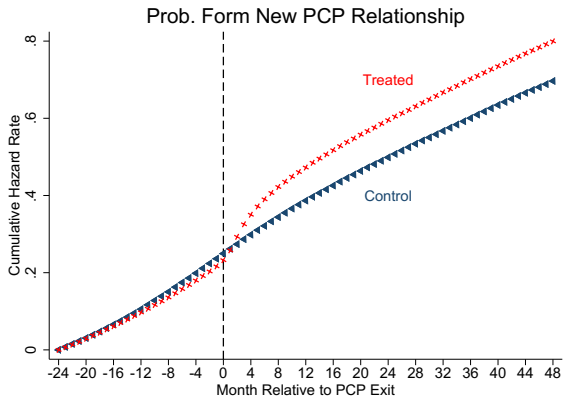
The raw data in Figure 1.4a and the event plots in Figure 1.5a show a sharp, discontinuous, and long-term decrease of -0.92 primary care visits per year when a PCP exits (16.9% decrease). Figure 1.4a shows that PCP visits dip about 0.02 visits below the long-term rate the month after a PCP's departure ($t = 1$). This suggests that patients do not immediately rematch to replacement PCPs.

The -0.92 long-term decline in primary care visits is due to decreases in the number of patients visiting PCPs at least once (18.6% extensive margin decrease) as well as the number of visits per patient (12.5% intensive margin increase) (1.2). Changes are driven by patients shifting to PCPs they know over starting new PCP relationships. Of the PCP visits made by patients, 90% are to PCPs that patients had a pre-existing relationship with the first

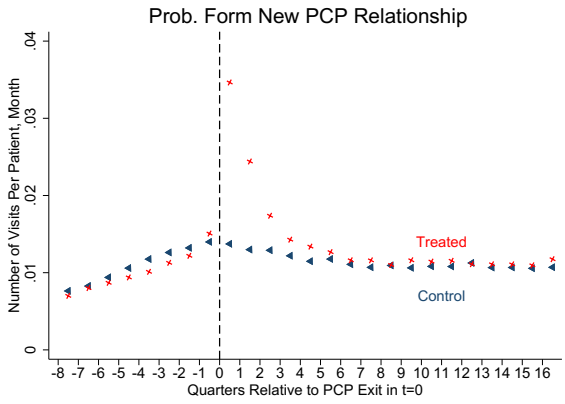
Figure 1.4: *Effects of a PCP Leaving a Clinic on Patients' Utilization of Care*



(a) *Raw Avg. No. PCP Visits*



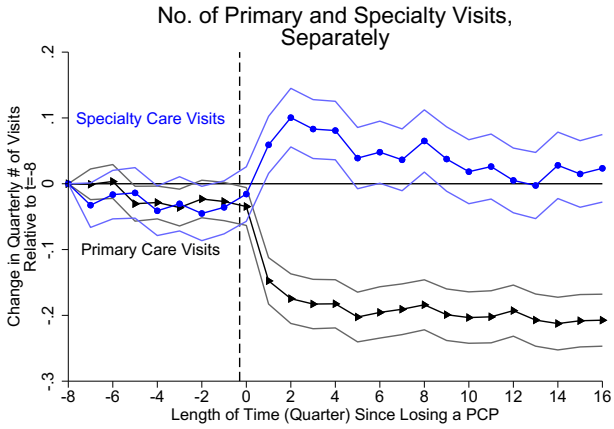
(b) *Cumulative Hazard Rate of Patients Forming New Relationships*



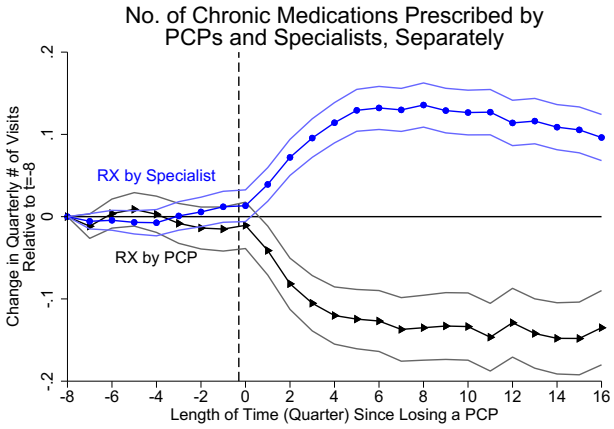
(c) *Raw Avg. No. New Relationships Formed*

Notes: Figure 1.4 shows the effect of a PCP's departure in the raw data. The x-axis contains time relative to the departure in $t = 0$, where relative time is measured in months. Figure 1.4a shows the raw number of primary care visits by month in a sub-sample that follows patients one year post-departure. Figure 1.4b and Figure 1.4c use a sub-sample that follows patients four years post-departure. Figure 1.4b shows the probability a patient forms a new relationship as a cumulative hazard rate. Figure 1.4c graphs the probability a patient forms a new relationship by relative time t. Treated patients are represented by red crosses, whereas control patients are represented by blue triangles.

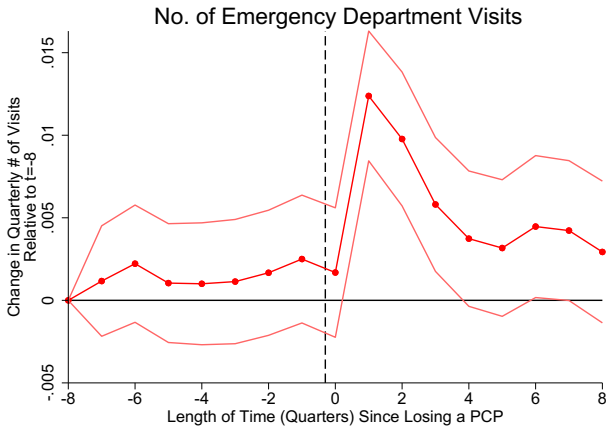
Figure 1.5: *Effects of a PCP Leaving a Clinic on Patients' Utilization of Care*



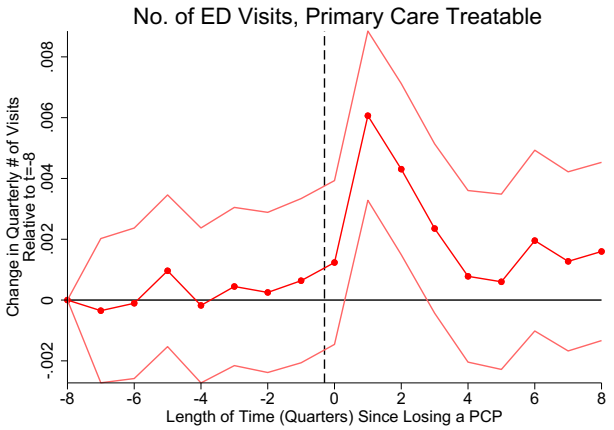
(a) *Number of Visits to PCPs and Specialists*



(b) *Number of Chronic Medications by Type of Prescriber*



(c) *Number of Emergency Department (ED) Visits*



(d) *Number of ED Visits for PCP Treatable Conditions*

Notes: Event study graphs plot each coefficient from the difference-in-differences specification outlined in Section 1.5.3. Regressions are at the PCP-quarter level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. "No." indicates that the outcome is the quarterly number. Plots use different data samples illustrated by the length of the x-axis. The first and second figure use a sample that follows patients 4 years post-departure, whereas the third and fourth follow patients 2 years post-departure. "ED" represents emergency department visits. Pooled, yearly estimates are in Table 1.2 and Table 1.3.

year post-departure.[34] Patients who lose a PCP are only 10 pp more likely to form a new relationship than patients who do not lose a PCP over the four years post-departure (Figure 1.4b and Figure 1.4c).[35,36]

Partially offsetting the decrease in PCP visits, the blue line in Figure 1.5a shows a 0.53 increase in specialty visits, which is immediate and sustained (5.6% increase). Increases are primarily driven by the intensive margin. The number of patients with at least one specialist visit increases by 1.3%, whereas the number of visits among patients with at least one visit increases by 9.4% (Table 1.2). This is driven by patients shifting to specialists they are familiar with: 80% of specialist visits are to specialists that patients have an existing relationship with (Table A.16).[37]

When patients decrease their use of PCPs and increase their use of specialists, this affects where patients receive preventive care. Specialists administer significantly *more* flu vaccines, annual exams, preventive screens, total prescriptions, and chronic medication prescriptions after a patient's PCP exits, whereas PCPs administer significantly *fewer* (Table 1.3). The number of preventive screens and flu vaccines administered in retail settings also significantly increase post-departure (Table A.20). On aggregate, this translates into patients receiving 6.4% fewer flu vaccines, 25.0% fewer annual exams, and 2.2% fewer preventive screens the first year post-departure (Table 1.2). Although Table 1.2 shows that all prescriptions as well as chronic prescriptions do not change on aggregate, patients' prescription regimens change in potentially beneficial ways. Prescription process measures show that there is an increase in new medication prescriptions as well as an increase in patients switching prescriptions within the same medication class post-departure (Table A.20). Further, the number of opioid prescription fills significantly decrease by 2.7%.

[34]Table A.16 shows estimates for E&M visits. 90% = (3.1 existing E&M - 0.73 point estimate) / (3.2 all E&M -0.57 point estimate).

[35]See Table A.17 for results that follow patients four years post-departure.

[36]Main effects are maintained when practice size is added as an additional matching covariate (Table A.18).

[37]Table A.16 shows estimates for E&M visits. 80% = (2.5 existing E&M + 0.16 point estimate) / (3.1 all E&M + 0.23 point estimate).

In addition to patients substituting to specialists for primary care long-term, urgent care visits, ED visits, inpatient admissions, and death significantly increase. Urgent care visits increase by 0.0025 visits per year (6.8%), ED visits increase by 0.033 visits per year (4.1%), and inpatient admissions increase by 0.011 visits per year (2.5%) (Table 1.2 and Figure 1.5). These increases are isolated to the first year post-departure (Table A.17). Figure 1.5 shows that patients increase their use of the ED for primary care treatable conditions by 0.016, or about 50% of the increase in total ED visits. ED visits for not preventable conditions also significantly increase by 0.011 visits per year. As a falsification test, I show that injuries are not significantly affected (Table A.16). As for death, Table 1.2 shows that there is an increased probability of death of 4.3%. Lending credibility to this finding, in specifications controlling for whether the clinic practiced as a team and individual, the probability of death is 3.4%, but only marginally significant at the 10% level A.19.

These effects cause total spending to increase $143.70 the first year after a PCP's exit (Table 1.2).[38] In a 100% sample of Medicare patients, where the average PCP sees 29 unique patients, this translates to $4,640.00 in increased out-of-pocket costs and $20,691.50 in increased total costs per exiting PCP. Medicare's cost is $16,051.50 ($20,691.50 - $4,640.00). This should be considered a lower bound for the costs associated with a PCP's exit because it only includes Medicare patients.

In total, relationships determine where patients' demand care. After a PCP leaves, patients shift to specialists that they had a pre-existing relationship with for primary care long-term. To provide further support for this, I show that patients substitute to specialists who act closest to PCPs—e.g. nephrologists, cardiologists, and gastroenterologists—opposed to specialties that deliver short-term, condition specific care, such as surgeons (Table A.20).

Patients health is adversely affected by a PCP's exit. Patients have an increased probability of death and increase their use of the ED and inpatient setting, with about half of the increase in ED use being driven by urgent conditions. The decrease in flu vaccines and

[38]Most of this increase is due to increased ED and inpatient use: Table A.16 shows that ED and inpatient charges increase by $124.70, or 87% of the total increase.

preventive screens also has a meaningful impact on patients' probability of death long-term. Relying on estimates from the clinical literature, the decrease in preventive care leads to an increase of about 50 deaths per 100,000 patients (Appendix B.1).

Table 1.2: *Treatment Effect of a PCP Leaving a Clinic*

Type	Mean	Impact
Utilization of Clinic Based Services		
No. of Specialist and Primary Care Visits	14.8	-0.38***
		(0.044)
		-2.6%
No. of Primary Care Visits	5.3	-0.92***
		(0.030)
		-17.2%
No. of Patients Visting PCP at Least Once	0.87	-0.11***
		(0.0024)
		-12.5%
No. of PCP Visits, Intensive Margin	6.2	-1.2***
		(0.097)
		-18.7%
No. of Specialist Visits	9.5	0.53***
		(0.033)
		5.6%
No. of Patients Visting SP at Least Once	0.86	0.012***
		(0.0011)
		1.4%
No. of SP Visits, Intensive Margin	11.1	1.2***
		(0.15)
		11.2%
No. of Urgent Care Visits	0.012	0.00084
		(0.00057)
		7.1%
Tot. Spending	8444.3	141.4**
		(65.3)
		1.7%
Tot. Out of Pocket	1258.6	15.5
		(10.4)
		1.2%
Treated PCP Sample Size	12497	
Control PCP Sample Size	12497	

Type	Mean	Impact
Medications		
No. of Filled Prescriptions	18.8	0.12*
		(0.062)
		0.63%
No. of Chronic Med RX Fills	7.0	0.075***
		(0.024)
		1.073%
Preventive Care		
Tot. Amount of Preventive Care	2.2	-0.086***
		(0.015)
		-3.9%
Prob. of a Flu Vaccine	0.49	-0.030***
		(0.0022)
		-6.1%
No. of Annual Exams	0.100	-0.019***
		(0.0029)
		-19.6%
No. of Preventive Screens	1.7	-0.039***
		(0.014)
		-2.3%
Aggregate Markers for Poor Care		
No. of Emergency Department Visits	0.74	0.030***
		(0.0048)
		4.0%
No. of ED Visits, Primary Care Treatable	0.36	0.011***
		(0.0031)
		2.9%
No. of ED Visits, Not Preventable	0.11	0.0061***
		(0.0017)
		5.6%
No. of Inpatient Visits	0.36	0.0085**
		(0.0031)
		2.4%
Prob. of Death	0.045	0.0019**
		(0.00083)
		4.1%

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability. "No." and "Tot." indicate that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. Preventive screens include mammography screens, colorectal cancer screens, cholesterol screens, and diabetes screens (see Table A.20 for breakouts). See Section A.1.1 for how medications were defined. See Section 1.4.2 for other variable definitions.

Table 1.3: *Shift of Care from Primary Care to Specialty Setting*

PCP Administered			Specialist Administered		
Medications					
No. of Filled Prescriptions from PCP	17.5	-1.3*** (0.078) -7.3%	No. of Filled Prescriptions from Specialists	8.7	1.2*** (0.057) 14.1%
No. of Chronic Med RX Fills from PCP	7.2	-0.46*** (0.033) -6.5%	No. of Chronic Med RX Fills from Specialists	2.5	0.50*** (0.023) 19.7%
Preventive Care					
Tot. Amount of Preventive Care by PCP	1.7	-0.24*** (0.014) -14.4%	Tot. Amount of Preventive Care by SP	0.37	0.14*** (0.0055) 38.1%
Prob. of a Flu Vaccine by PCP	0.27	-0.061*** (0.0026) -22.4%	Prob. of a Flu Vaccine by Specialist	0.040	0.019*** (0.0011) 48.5%
No. of Preventive Screens by PCP	1.3	-0.14*** (0.013) -10.9%	No. of Preventive Screens by Specialist	0.33	0.10*** (0.0054) 31.7%
No. of Annual Exams by PCP	0.12	-0.028*** (0.0040) -24.5%	No. of Annual Exams by Specialist	0.0050	0.010*** (0.00061) 207.4%

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability. "No." and "Tot." indicate that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. The chronic medication category includes Selective Serotonin Reuptake Inhibitors (SSRIs), antihypertensives, antidiabetics, and statins.

1.6.2 Robustness Checks

In addition to analyzing pre-trends, I test for differential attrition into Medicare Advantage (MA). I do not observe patients when they switch to MA, so this could create an issue if MA switches are different than patients who do not switch. This test addresses at least two concerns. First, if a PCP stops taking Traditional Medicare (TM) patients, this would be categorized as an exit. In response, patients may switch to MA in order to continue seeing their PCP. Second, patients may switch to MA to access additional services or providers in response to the loss of a PCP. Alleviating these concerns, Table A.16 shows that patients do not systematically switch into MA in response to a PCP's exit. This then motivates dropping patients that switch to MA from the main sample to reduce noise.

I also test if clinic level changes cause PCPs to exit. This would imply that effects

attributed to a PCP's departure may instead be due to clinic level changes. to test for this, Figure A.2 plots the number of PCPs exiting over time. It shows that PCP exits occurring after the main PCP departs do not systematically happen at the treatment threshold, but rather the line trends smoothly downward two years post-departure. This implies that there are not systematic changes occurring at the clinic that are driving effects.[39]

1.7 Why and How Much is the Relationship Valued?

1.7.1 The Importance of Relationship-Specific Capital

Evidence presented in Section 1.6.1 suggests that patients may be impacted by the loss of a PCP because they lose relationship-specific capital. I use plausibly exogeneous varaition in patients' relationships with other PCPs and specialists to test if this is the case. When a PCP exits, all treated patients lose the relationship-specific capital they had with the exiting PCP. If the information is specific to a particular relationship, patients who have additional PCP and specialist relationships, outside of their relationship with the exiting PCP, should be able to maintain a higher stock of capital than patients with fewer outside relationships.

To do this, I leverage heterogeneity in the clinic environment determined by management practices. I compare treatment effects among patients who belong to clinics that close in response to a PCP's exit, patients who belong to open clinics where PCPs care for patients one-on-one, and patients belonging to open clinics where patients are cared for by a team of PCPs. Patients are least able to rely on a PCP they know when a clinic closes in response to a PCP's exit and most able to substitute to PCPs they know in clinics that are managed as a team. Patients are therefore monotonically less likely to have to form a new PCP relationship across options.

PCPs and patients of clinics that closed versus stayed open are observably quite similar,

[39]Three other tests are still in progress. First, I focus on a group of PCPs who were not pre-replaced. Second, I focus on PCPs who do not wind down their patient load pre-departure. Third, I isolate PCPs whose patients didn't leave their PCP and go elsewhere pre-departure. I will ensure that effects are maintained across subgroups.

with the exception of clinic size. Table A.7 shows that clinics that close when a PCP departs have on average 1.5 PCPs at baseline, whereas those that stay open have on average 12.7 PCPs. Given size differences, it could be the case that smaller clinics that remain open post-departure are a better counterfactual than the aggregate open clinic category. Regardless, results are robust to size (Table 1.7, Table A.33, and Table A.31).

Table A.8 shows balance across individual and team control clinics. PCPs working in team clinics are more likely to be female. Team clinics are also twice as likely to have a NP or PA on staff and have 3 times more PCPs on average than individual clinics (12.6 v. 4.4). Patients in individual clinics are less likely to be white, more likely to be enrolled in Medicaid, have more primary care visits, and have a similar number of specialty visits 3-2 years before a PCP's exit compared to patients in team clinics. Patients in individual clinics also have smaller primary care and specialty networks than patients in team clinics. Patients in individual clinics visit 1.1 PCPs and 2.6 specialists on average, whereas patients in team clinics visit 1.8 PCPs and 3.8 specialists on average 3-2 years before exit.[40,41]

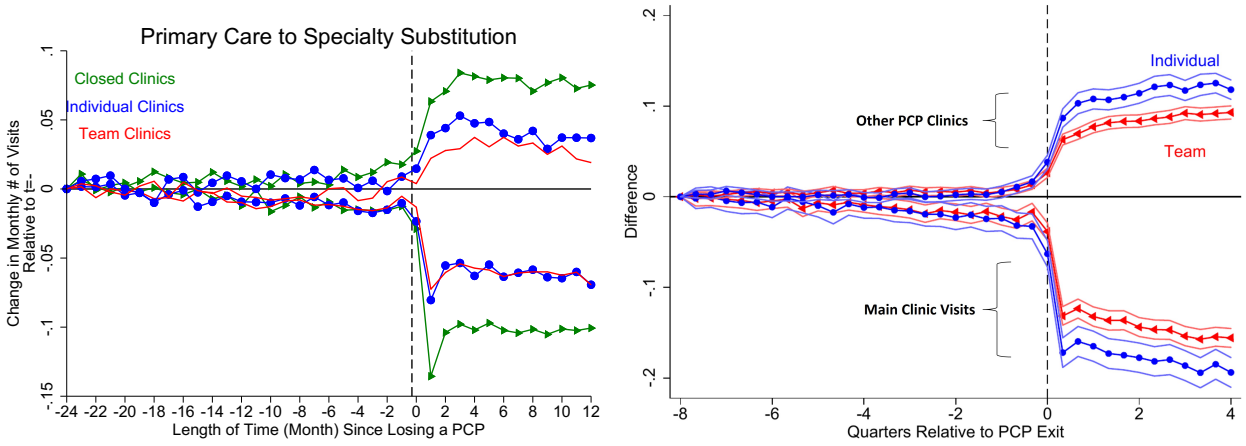
Patients' use of primary and specialty care across clinic environments is shown in Figure 1.6, which plots estimated coefficients $\beta_t^{Treated}$. Following equation 1.1, t is estimated at the month level. Coefficients are only identified up to a constant term, so the value for $t = month = -24$ is normalized to zero. Figure 1.6a overlays primary care and specialty visits. It includes clinics that close (green line), stay open and practice individually (blue line), and stay open and practice as a team (red line). The next panel, Figure 1.6b, breaks aggregate primary care visits into primary care visits at the main clinic and visits at other PCP clinics.

To guide the interpretation of the empirical results, I revisit the conceptual framework in Section 1.3. The theoretical framework delivers three testable predictions concerning

[40]Given differences, one may be concerned that control PCPs are not an appropriate control group. Table A.23 shows results when matching on whether the clinic was on an individual and team model, showing extremely similar results.

[41]Main results use an individual and team clinic definition that uses an above and below average threshold. (See Section 1.4.2 for more details on how the groups were defined.) As a robustness check, Table A.24 uses the 25th, 50th, and 75th percentile as cut-offs and shows that results are qualitatively the same.

Figure 1.6: *Effects of a PCP Leaving a Clinic on Patients' Utilization of Care*



(a) *Use of primary and specialty care across clinic environments*

(b) *Location of PCP visit by team and individual clinics*

Notes: Event study graphs plot estimated coefficients $\beta_t^{Treated}$ from equation 1.1, where t is estimated at the month level and patients are followed one year post-departure. Regressions are at the PCP-month level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Coefficients are only identified up to a constant term, so the value for $t = month = -24$ is normalized to zero. Figure 1.6a overlays primary care against specialty visits. It includes clinics that close (green line), stay open and practice individually (blue line), and stay open and practice as a team (red line). Standard errors are not included for clarity. The next two panels take the decrease in primary care visits from the main figure and focus on the differences between individual and team clinics. Figure 1.6b breaks aggregate primary care visits into primary care visits at the main clinic and visits at other PCP clinics. Table 1.4 shows point estimates.

relationship-specific capital: relationship-specific capital exists and grows over time, contains health-specific information, and is nontransferrable.

First, if relationship-specific capital exists and grows over time, patients will visit PCPs they have a relationship with over establishing a new PCP relationship after the loss of a longstanding PCP. Patients will only visit a specialist over establishing a new PCP relationship if patients have an existing relationship with a specialist. In support of this prediction, Figure 1.6a shows that as patients are less likely to know a PCP in the main clinic, patients are more likely to substitute to specialists. Patients decrease their use of primary care by -1.3 visits per year when clinics close, -0.76 visits per year in open individual clinics, and -0.69 visits per year in open team clinics. Patients increase their use of specialists by 0.82 visits when the clinic closes, 0.42 visits in open individual clinics, and 0.23 visits in

open team clinics.

Figure 1.6b focuses on the aggregate -0.76 and -0.69 primary care visit decline in open individual and team clinics. Patients in individual clinics are more likely to leave main clinics and visit an alternative primary care clinic than patients in team clinics. Visits to the main clinic decrease by 2.1 visits for individual clinic patients and 1.7 visits for team clinic patients. Visits to other primary care clinics increase by 1.4 visits among individual clinic patients and 0.98 visits for team clinic patients.

Second, if relationship-specific capital contains health-specific information, patients who do not have known PCPs, outside of their relationship with the exiting PCP, should be most likely to experience an adverse event. I define adverse events to include visiting the ED for urgent conditions, an inpatient admission, and death. Table 1.4 supports this prediction by showing that adverse events monotonically increase across clinic environments.[42]

Third, if relationship-specific capital is nontransferrable, all patients, no matter their ability to switch to known PCPs and specialists, should be affected by the loss of a long-standing PCP. In support of this, even patients in team clinics, where PCPs are the most interchangeable, decrease their use of PCPs and increase their use of specialists in response to the loss of a PCP.

The empirical results rule out three other information models shown in Figure A.7. Model A considers an information model where specialist-specific information grows over time, but PCP-specific information does not. In this case, patients' utility maximizing point is to consume primary care with the specialist they have a relationship with, at point $Spec$, opposed to a pre-existing PCP, no matter the length of a patient's relationship with the pre-existing PCP. This counters the empirical results in Figure 1.6, that shows that patients are more likely to visit a PCP they have a relationship with over a specialist. Model B shows a model where PCP-specific information grows over time, but specialist-specific information does not. In this case, patients with and without relationships with pre-existing PCPs will be equally likely to switch to a specialist after a PCP's exit, which counters the empirical results.

[42]Table A.23 shows results after matching on whether the clinic practiced on a panel or individual model.

Model C considers a model where neither PCP-specific nor specialist-specific information grows over time. Predictions are the same as Model B for similar reasons, a prediction that is not borne out by the data.

Table 1.4: *Treatment Effect of a PCP Leaving a Clinic by Patient Starting a New Relationship*

	Clinic Also Closed		Clinic Stayed Open				
			Individual		Team		
Type	Mean	Impact	Mean	Impact	Mean	Impact	P-Value
No. of Specialist and Primary Care Visits	15.3	-0.42***	15.6	-0.30***	14.4	-0.38***	0.55
		(0.078)		(0.11)		(0.055)	
		-2.8%		-2.0%		-2.6%	
No. of Primary Care Visits	5.5	-1.3***	6.0	-0.81***	5.1	-0.68***	0.16
		(0.056)		(0.086)		(0.033)	
		-22.9%		-13.5%		-13.2%	
Prob. Form New PCP Relationship	0.044	0.12***	0.043	**0.038*** **	0.046	**0.031*** **	0.03
		(0.0018)		(0.0026)		(0.0013)	
		267.9%		88.2%		68.8%	
No. of Specialist Visits	9.7	0.85***	9.7	**0.50*** **	9.3	**0.30*** **	0.03
		(0.052)		(0.082)		(0.043)	
		8.7%		5.2%		3.2%	
Tot. Spending	8935.20	271.30***	8814.4	99.2	8156.4	64.9	0.84
		(97.50)		(148.6)		(84.5)	
		3.0%		1.1%		0.80%	
Medications							
No. of Filled Prescriptions	19.3	0.074	21.3	-0.065	18.1	0.23***	0.08
		(0.093)		(0.15)		(0.077)	
		0.38%		-0.31%		1.3%	
No. of Chronic Med RX Fills	7.1	0.13***	7.9	0.034	6.8	0.059**	0.71
		(0.036)		(0.061)		(0.030)	
		1.8%		0.43%		0.87%	
Preventive Care							
Tot. Amount of Preventive Care	2.2	-0.12***	2.3	-0.12***	2.2	-0.050***	0.07
		(0.022)		(0.033)		(0.020)	
		-5.6%		-5.1%		-2.3%	
Prob. of a Flu Vaccine	0.47	-0.038***	0.48	-0.029***	0.50	-0.025***	0.55
		(0.0039)		(0.0058)		(0.0030)	
		-8.0%		-6.1%		-5.1%	
Aggregate Markers for Poor Care							
No. of Emergency Department Visits	0.73	0.044***	0.79	0.036***	0.73	0.017***	0.13
		(0.0073)		(0.012)		(0.0062)	
		6.1%		4.6%		2.3%	
No. of ED Visits, Not Preventable	0.11	0.011***	0.11	0.0093***	0.11	0.0017	0.08
		(0.0026)		(0.0038)		(0.0023)	
		10.0%		8.4%		1.5%	
No. of Inpatient Visits	0.36	0.014***	0.37	0.0071	0.35	0.0051	0.81
		(0.0046)		(0.0072)		(0.0040)	
		3.8%		1.9%		1.4%	
Prob. of Death	0.046	0.00097	0.045	0.0022	0.044	0.0022**	0.98
		(0.0012)		(0.0020)		(0.0011)	
		2.1%		5.0%		4.9%	
Not Preventable ED, Inpatient Use, and Death	0.51	0.025***	0.53	0.019**	0.51	0.0089*	0.37
		(0.0061)		(0.0094)		(0.0054)	
		4.9%		3.5%		1.8%	
Treated PCP Sample Size	3523		2134		6840		
Control PCP Sample Size	12497		7334		5163		

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. Bolded estimates indicate that the groups are significantly different at the 5% level. "Prob." indicates that the outcome is the yearly probability. "No." and "Tot." indicate that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. See Table A.22 for additional outcomes. See Section 1.4.2 for more details on how heterogeneity and variables were defined.

1.7.2 Back-of-the-Envelope Willingness to Pay

The empirical results in Section 1.7.1 show that relationship-specific capital exists, grows over time, contains health-specific information, and is nontransferrable. To get traction on patients' valuation of relationship-specific capital, I quantify how far patients travel to maintain relationships with PCPs who move 300 miles away (these moves are not used in the main analyses). I show that 38% of a PCP's original pool of patients move with the PCP on average (Figure A.3). The number of patients moving with PCPs is especially large considering that 39 states enforce non-compete agreements, which legally prohibit PCPs from taking their patients with them when moving practices (Hausman and Lavetti, 2016).

Of those patients who travel a positive distance to move with PCPs, patients travel an average of 66 miles (median 57 miles) one-way to continue seeing their PCP, or the 99^{th} percentile of a control patient's travel time. Integrating over Figure A.3, patients are willing-to-drive an additional 208 miles per visit to follow their PCP on average. In terms of the total miles driven per year, Table A.21 shows that patients visit their assigned PCP 5.3 visits per year, which insignificantly decreases 0.44 visits per year among patients who move with their PCP, or 4.86 visits per year. As a result, patients are willing-to-drive an additional 1011 miles, or 19 additional hours, per year on average.[43]

To translate driving time into dollars, I use three pieces of information from the literature. First, a wage rate of \$13.81 per hour is typically used when estimating the cost of time (AARP Public Policy Institute, 2015). This wage rate is a lower bound on an elderly person's valuation of their time because, by choosing to be retired, individuals value leisure time at least as much as their wage rate by revealed preference. Second, 31% of individuals over age 50 have a caregiver, who likely accompanies an elder to an appointment, especially if traveling a long distance (Reinhard, Feinberg, Houser, Choula and Evans, 2019). Third, the literature typically calculates fuel costs at \$0.10 per mile.[44] This translates into an elderly individual valuing their PCP-relationship around \$300-\$400 per year.

[43] 1011 miles = 208 miles round-trip*4.86 visits per year

[44] This assumes a cost per gallon of \$3.25 and a vehicle that gets 30 miles per gallon.

1.8 Alternative Explanations

1.8.1 Are Patients Unable to Find a PCP that is a Good Match?

PCP-specific information may be match specific, where good matches are hard to find (Jovanovic, 1979b). If patients decrease their use of primary care because they cannot find a good match, effects should be magnified among patients with more specific needs. To test for this, I compare patients that are high risk, disabled, and of a minority race. Female patients may also prefer having a female PCP, so I compare female patients to male patients with the same female exiting PCP.

Tables A.12-A.15 give a sense of how these groups observably differ. Table A.12 shows that high risk patients are about three years older on average, are more likely to be in end stage renal disease, and enrolled in Medicaid than lower risk patients.[45] They also use the medical system at a higher rate and have larger PCP networks. The average high risk individual visits a PCP 8.1 times and sees 4.1 different PCPs per year, whereas the average not high risk individual visits a PCP 5.5 times and sees 1.4 PCPs per year.

Similarly, Table A.13 shows that disabled patients look fairly similar to high risk patients. They are higher risk and use the medical system at a higher rate than not disabled patients (7.4 vs. 5.6 primary care visits per year). Table A.14 compares minority and white patients. Minority patients are twice as likely to be also enrolled in Medicaid as well as in end stage renal disease. While minority patients use more ED care than their white counterparts, they also use less primary care, specialty care, and urgent care. They see 6.0 PCPs and 13.7 specialists in a given year, whereas white patients see 1.6 PCPs and 3.9 specialists. This suggests that minority patients may interact with the health system differently than white patients.

Female and male patients who have the same female PCP are also compared. If patients decrease their use of primary care because it is hard to find a good match, female patients who lose a female PCP should take longer to re-match if patients prefer female PCPs. Table

[45] See Section 1.4.2 for more details on how this group was defined.

A.15 compares balance between groups. Male patients are two years older and less likely to be in end stage renal disease than their female counterparts, but otherwise look fairly similar. Female patients also have larger primary care and specialty networks.

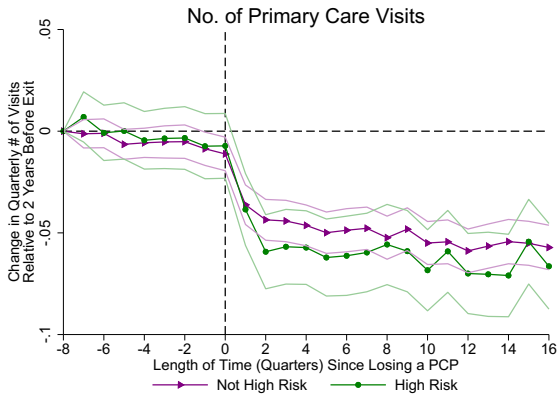
Figure 1.7 plots estimated coefficients $\beta_t^{Treated}$. Following equation 1.1, t is estimated at the quarter level. Coefficients are only identified up to a constant term, so the value for $t = quarter = -8$ is normalized to zero. Figure 1.7a compares high risk to low risk patients, Figure 1.7b compares disabled to not disabled patients, Figure 1.7f compares minority to white patients, and Figure 1.7d compares male to female patients with exiting female PCPs. Table 1.5 contains relevant point estimates, that are pooled over the entire year. All analyses control for PCP as well as clinic level factors that may be different between groups by comparing patients with the same exiting PCP.

Patients with more specific needs should be (a) less likely to re-match to a PCP and (b) more likely to substitute from primary to specialty care if it is harder for them to find a good match. Figure 1.7 shows that all patients decrease their use of primary care for at least four years after a PCP's exit, with no sign of recovering. This suggests that patients with more specific needs are not less able to re-match to a PCP. In terms of the rate of substitution from primary to specialty care, high risk patients decrease their use of primary care and increase their use of specialty care more in terms of the number of visits, but less in terms of percent changes. Disabled patients are less likely to substitute away from the primary care to specialty setting, which is counter to what we would expect if disabled patients had a harder time matching to a replacement PCP. Minority and white patients decrease their use of primary care similarly, but minority patients are more likely to switch to a specialist, which may suggest that minorities like their replacement PCP less than white patients.[46] When comparing male to female patients with female PCPs, both groups decrease their use of primary care similarly and males are more likely to increase their use of specialists. In sum, there is limited evidence to support match quality along these dimensions as a

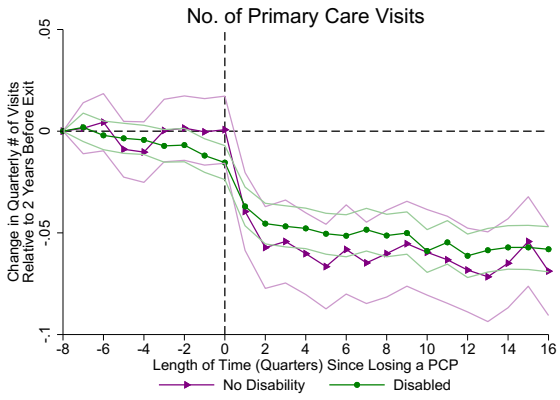
[46]Ideally, I would be able to compare minority to white patients with minority PCPs, but, unfortunately, the data limits my ability to determine the race of a PCP.

relevant mechanism.

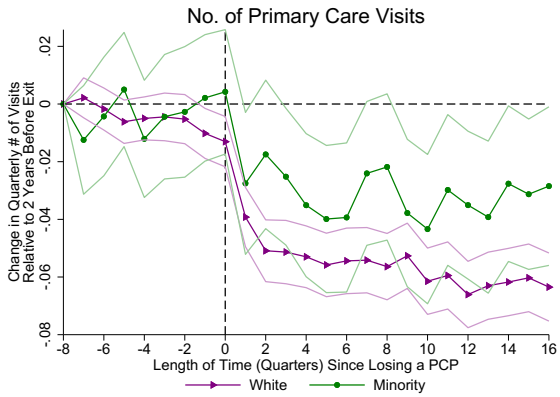
Figure 1.7: *Effects of a PCP Leaving on PCP Visits by Groups With More Specific Needs*



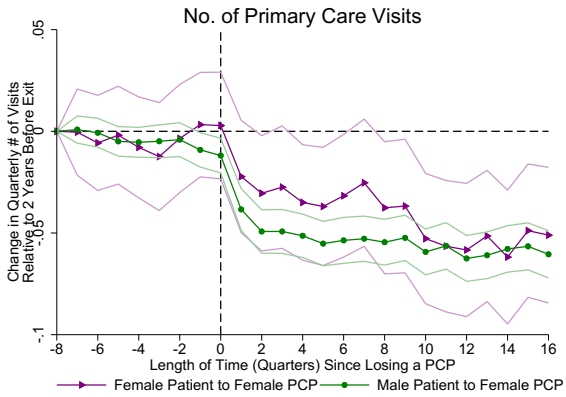
(a) *Primary care visits by patient risk status*



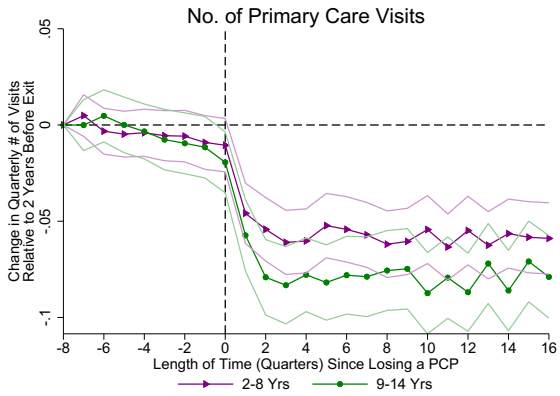
(b) *Primary care visits by disability status*



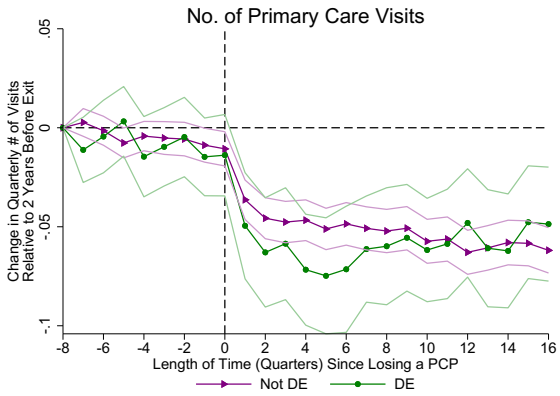
(c) *Primary care visits by patient race*



(d) *Primary care visits by sex of PCP and patient*



(e) *Primary care visits by length of PCP relationship*



(f) *Primary care visits by dual eligibles*

Notes: Event study graphs plot each coefficient from the difference-in-differences specification outlined in Section 1.5.3. "No." indicates that the outcome is the monthly number. Regressions are at the PCP-month level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Event studies rely on the a sample that follows patients 4 years post-departure. High risk patients are patients with the top quartile of Elixhauser risk scores. See Table 1.5 for pooled estimates across outcomes.

Table 1.5: *Treatment Effect of a PCP Leaving a Clinic by Patient Status*

	High Risk		Not High Risk		Func.		No Func.		Minority		White		Male-Female Match		Female-Female Match	
No. of Primary Care Visits	6.9	**-0.98*****	4.8	**-0.77*****	4.9	**-0.74*****	6.4	**-1.062*****	5.3	-0.87***	5.4	-0.86***	5.3	**-0.63*****	5.0	**-0.48*****
		(0.043)		(0.028)		(0.027)		(0.046)		(0.072)		(0.032)		(0.031)		(0.065)
		-14.1%		-16.0%		-15.1%		-16.6%		-16.5%		-15.8%		-11.8%		-9.7%
p-value			p< 0.001				p< 0.001				0.86				0.03	
No. of Specialist Visits	13.4	**0.55*****	8.1	**0.48*****	8.3	**0.45*****	12.5	**0.62*****	8.1	**0.59*****	9.7	**0.42*****	9.4	0.38***	9.5	0.19*
		(0.063)		(0.033)		(0.033)		(0.059)		(0.083)		(0.037)		(0.036)		(0.11)
		4.1%		6.0%		5.4%		5.0%		7.2%		4.3%		4.0%		2.0%
p-value			p< 0.001				0.004				0.05				0.09	
Rate Adverse Events	0.86	**0.032*****	0.38	**0.014*****	0.40	0.013***	0.78	0.030***	0.59	0.019	0.59	0.021***	0.50	0.015***	0.52	0.0055
		(0.0098)		(0.0040)		(0.0041)		(0.0087)		(0.016)		(0.0056)		(0.0045)		(0.016)
		3.8%		3.6%		3.2%		3.8%		3.2%		3.5%		2.9%		1.059%
p-value			0.009				0.06				0.91				0.56	
Prob. of Death	0.076	**0.0057****	0.032	**0.00077**	0.037	0.0017*	0.061	0.0017	0.030	-0.00076	0.045	0.0014	0.042	0.0020**	0.046	0.0047
		(0.0020)		(0.00081)		(0.00087)		(0.0017)		(0.0023)		(0.00094)		(0.00091)		(0.0032)
		7.4%		2.4%		4.6%		2.7%		-2.6%		3.1%		4.6%		10.4%
p-value			p< 0.001				0.98				0.38				0.39	
Treated PCP Sample Size	12191		12185		12285		11119		5154		9986		8907		2738	
Control PCP Sample Size	12191		12185		12285		11119		5154		9986		12427		3349	

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. Bolded estimates indicate that the groups are significantly different at the 5% level. "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Comparisons between groups are within PCP. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data only follows patients for one year post-departure. See Section 1.4.2 for more details on how heterogeneity and variables were defined. "Male-Female Match" represents male patients who had a female PCP and "Female-Female Match" represents female patients who had a female PCP.

1.8.2 Are Patients Unable to Find a Replacement PCP?

Patients may substitute away from the primary care setting towards specialty care because they are unable to find a replacement PCP. If the availability of PCPs drives the decrease in primary care visits, patients in thinner markets should be less able to find a new PCP than patients in thicker markets. Patients in thinner markets should therefore decrease their use of primary care more than patients in thicker markets.

To understand the relevance of this mechanism, I compare the probability a patient forms a new PCP relationship as well as the number of PCP visits by the local density of PCPs. This is defined as the number of PCPs filing billing claims within a 30 mile radius of each focal clinic ZIP divided by the local population. I focus on clinics that remained open after a PCP's departure because clinics may be more likely to close in rural areas. Thick markets are defined to be above average density areas and thin markets are below average density areas.

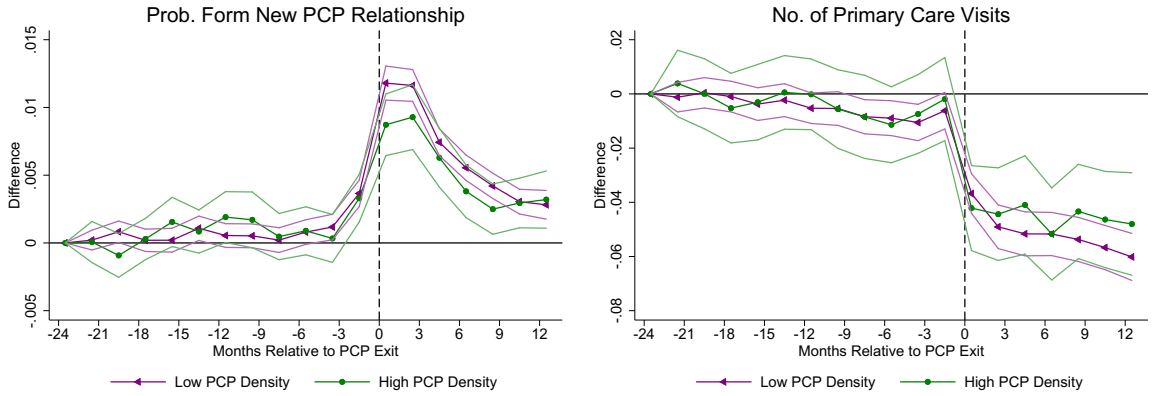
Table A.9 compares observables between high and low density areas 3-2 years before a PCP's departure. Groups are fairly similar except patients in thin markets are more likely to live in rural areas. Figure 1.8 plots estimated coefficients $\beta_t^{Treated}$ from equation 1.1, where t is estimated at the month level. Coefficients are only identified up to a constant term, so the value for $t = month = -24$ is normalized to zero. The first panel shows total primary care visits and the second panel the number of new patient visits.

Figure 1.8 shows that patients in thicker markets are *less* likely to establish a new PCP relationship than patients in thinner markets (2.4% vs. 3.3% of visits, $p = 0.001$). Further, primary care visits decrease by similar amounts in thin and thick markets. This provides evidence against the local availability of PCPs affecting patients' rate of using primary care.[47,48]

[47]Table A.27 shows additional estimates. It shows that urgent care and specialty visits do not increase significantly more in low density areas, compared to high density areas. This is likely due to the local density of PCPs being highly correlated with the availability of specialist as well as urgent care clinics.

[48]I varied this analysis in two ways. First, I compare patients who were and were not affiliated with a clinic that has multiple sites in Table A.28. When comparing patients who lost a clinic using this heterogeneity, it was not an important indicator of patients ability to re-match. However, when comparing patients in open

Figure 1.8: *Effects of a PCP Leaving on Patients' Utilization of Care by Local Density of PCPs*



Notes: Event study graphs plot each coefficient from the difference-in-differences specification outlined in Section 1.5.3. Estimates are relative to $month = -24$. "Prob." indicates that the outcome is the monthly probability and "No." indicates that the outcome is the monthly number. Regressions are at the PCP-month level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Event studies rely on the main sample shown in Table 1.1, which only follows patients for one year post-departure. Density is defined by the number of PCPs within a 30 mile radius divided by the population in that zip code. High PCP density areas are above average density and low PCP density areas are below average density. Table A.27 contains pooled point estimates.

1.8.3 Are Patients Unable to Access Care at Focal Clinics?

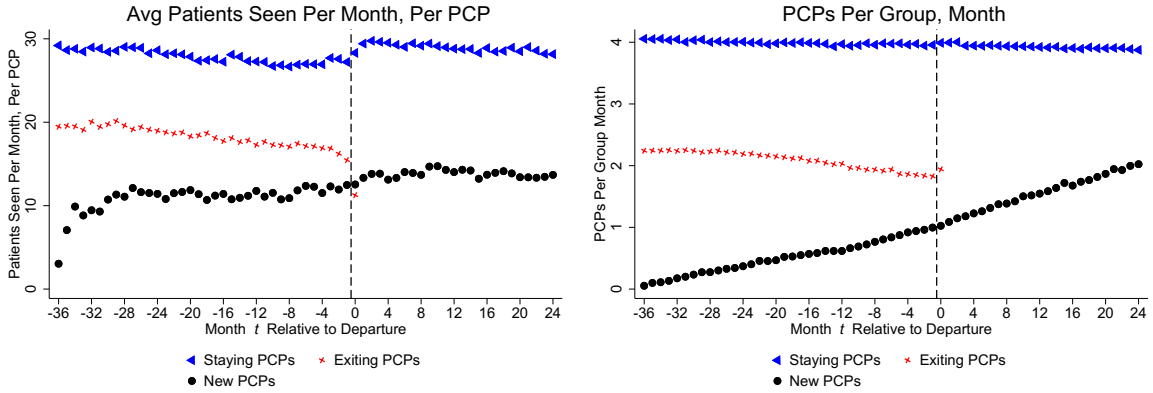
Patients may substitute to non-primary care settings because they are unable to access care at main clinics. I test this in two ways. First, I quantify the number of patients seen by staying PCPs after a PCP exits. I then show how these changes affect outcomes among staying PCPs' patients. If one assumes that staying PCPs treat their existing set of patients the same as the patients they inherit from exiting PCPs, this should indicate if directly affected patients are impacted vis-à-vis staying PCPs. Second, I compare treatment effects by the size of patients' home clinic. If clinics are constrained in their ability to care for patients, patients in smaller clinics should be more affected than patients in larger clinics.

Firm Disruption and Network Effects. A long literature has studied how firms compensate for the loss of a worker, starting with Slichter (1919). However, much less work

clinics by this metric, effects were more mixed. Second, I used the Center for Medicare and Medicaid Services rural/urban fee schedule to compare rural and urban areas. This is a coarser metric than the market thickness definition constructed in the data, which is based on local availability within 30 miles. However, Table A.29 shows that the results are qualitatively similar.

has examined how the loss of a worker affects the quality of a firm's outputs, especially in the health care context. I estimate spillovers by moving to the clinic level and matching clinics instead of PCPs (see Appendix A.2 for details on the sample). I then quantify how the loss of the main PCP affects staying PCPs and staying PCPs' patients at the main clinic. For instance, it may be the case that overburdened clinics are no longer able to maintain the same quality of care, causing patients to substitute away from their main primary care clinic. Figure 1.9a shows the number of patients seen per month per PCP, grouped by the

Figure 1.9: *Clinic Level Effects*
Patients Seen Per Month and Replacement Rate Over Time



(a) *Average Number of Patients Seen Per Month, Per PCP by Group*

(b) *Number of PCPs by Type of PCP Over Time*

Notes: Graphs rely on a sample that matches *clinics*, instead of PCPs. This sample does not include practices with one PCP. Additional details are described in Appendix A.2. Figure 1.9a shows the number of patients seen per month, per PCP by type of PCP. Blue triangles represent staying PCPs, a group that includes PCPs that existed at the clinic in $t = -36$. Red crosses represent exiting PCP, who exit from $0 \leq t \leq 24$. Black circles represent PCPs that are new to the clinic between $-35 \leq t \leq 24$. Because points in Figure 1.9a do not account for the size of the group, Figure 1.9b shows the number of PCPs per group over time (the denominator of the rate showed in Figure 1.9a).

type of PCP. The red crosses represent exiting PCPs, the blue triangles represent staying PCPs, and the black circles represent new PCPs. Exiting PCPs are defined to exit in $t = 0$. Staying PCPs are defined to be PCPs who were practicing at the clinic in $t \leq -36$, or 3 years before the main PCP's exit. New PCPs are any PCP that began practicing at the clinic from $-36 < t \leq 24$. Zeros are not included in the average, so to aid in the interpretation of Figure 1.9a, the total number of PCPs in each group is shown in Figure 1.9b (the denominator of

the average). The average clinic size is larger than that in the matched PCP sample. Solo clinics are not included in the clinic level sample because spillovers onto indirectly affected patients cannot be estimated in this case.

Figure 1.9a shows that staying PCPs are affected by a co-working PCPs departure. The first year post-departure, the average staying PCPs sees 44.1 more patients per year, which is sustained in year 2 (Table 1.6). The sustained increase in the number of patients seen by staying PCPs may be because clinics do not immediately replace leaving PCPs. Figure 1.9b shows that the rate of new PCPs being added to the clinic trends smoothly over time, opposed to suddenly increasing when a PCP exits.

I next quantify how increases in the number of patients seen by staying PCPs affects staying PCPs' patients. By assuming that staying PCPs treat their existing set of patients (indirectly affected patients) similarly to the patients they take-on post-departure (directly affected patients), this can tell us something about how directly affected patients are treated by staying PCPs.[49] Table 1.6 shows that outcomes among indirectly affected patients remain unchanged post-departure. This suggests that staying PCPs are able to compensate for the loss of a co-worker and maintain the same standard of care.

[49] "Indirectly affected" patients are patients of staying PCPs and do not ever have a claim billed by exiting PCPs.

Table 1.6: *Treatment Effect of PCP Unexpectedly Leaving Practice on Clinic Level Outcomes*

Type	Mean	Impact Year 1 Post Exit	Impact Year 2 Post Exit	P-Value
Firm Level Outcomes				
Avg Number of Pat Seen Per Month-PCP, Staying PCPs	214.9	**44.1***** (6.0) 20.5%	**38.1***** (6.5) 17.7%	0.02
Avg Number of Pat Seen Per Month-PCP, New PCPs	23.9	**35.7***** (6.1) 149.5%	**38.7***** (6.5) 162.1%	0.04
Count of New PCPs	0.056	0.028*** (0.0064) 51.2%	0.018*** (0.0061) 33.2%	0.10
Treated Clinic Sample Size Control Clinic Sample Size	1573 1573			
Indirectly Affected Patients' Outcomes				
Utilization of Clinic Based Services				
No. of Primary Care Visits	5.0	-0.050 (0.046) -1.0080%	-0.0012 (0.055) -0.025%	0.11
No. of Specialist Visits	9.8	-0.030 (0.060) -0.31%	0.091 (0.071) 0.92%	0.02
No. of Urgent Care Visits	0.015	-0.00061 (0.0012) -4.2%	0.00072 (0.0011) 5.0%	0.21
Aggregate Markers for Poor Care				
No. of ED and Inpatient Visits	0.89	0.0033 (0.0091) 0.37%	0.012 (0.0095) 1.3%	0.33
Treated Clinics Control Clinics	1558 1558			

Notes: This table displays results from a specification similar to the difference-in-differences specification outlined in Section 1.5.3, with one large difference: analyses rely on a data set that matches clinics, not PCPs, and follows patients two years post-departure. As a result of the clinic level match, clinics with only one PCP are not included. Regressions are at the clinic-year level, contain clinic fixed effects, and cluster at the clinic level. Indirect patients are patients who were never observed to visit a departing PCP. Regressions are at the clinic-year level, contain clinic fixed effects, and cluster at the clinic level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. See Section 1.4.2 for variable definitions and Appendix A.2 for more details on how the clinic level sample was created.

Effects by Size of Focal Clinic. The results above imply that a clinic is impacted by the loss of a PCP, but it does not affect staying PCPs' ability to care for patients. However, the above analysis quantifies aggregate effects, which may obscure significant heterogeneity by the size of a clinic. If a clinic's ability to provide care is impacted by the loss of a PCP, there are fewer PCPs to take on the increased workload in smaller clinics, so effects should be

larger among small clinic PCPs. I test this hypothesis in the matched PCP sample used in the main analyses and focus on clinics that remain open after a PCP's exit. I compare clinics with 1-3 PCPs versus 4-100 PCPs (3 PCPs is the median clinic size). Regression results follow the main specification outlined in equation 1.1.

I start by showing how small and large control clinics observably differ in Table A.11. The largest difference is that 66% of small clinics and 19% of large clinics practice on individual models. This suggests that results may be confounded by a clinic's management structure. Patients in larger clinics also see 1.8 PCPs and 3.6 specialists, whereas patients in smaller clinics see 1.1 PCPs and 2.8 specialists 3-2 years before a PCP's departure.

Table 1.7 shows that changes in aggregate primary care visits are not significantly different by clinic size.[50] However, patients are slightly more likely to shift away from the main clinic and increase their use of other primary care clinics and specialists. This could mean that smaller clinics may be more constrained post-departure, which affects patients' ability to access care at focal clinics.

Size is highly correlated with whether the clinic practices on an individual or team model. As a result, Table A.32 shows a 2x2 matrix with treatment effects by size and management structure. It shows that treatment effects are driven by differences in the clinic model, rather than size. For instance, the number of visits to patients' main clinics decreases by 1.7 visits in shared models with 1-3 PCPs and 1.7 visits in shared models with 4-100 PCPs. For individual clinics, the number of visits to the main clinic decreases by 2.4 visits in clinics with 1-3 PCPs and 2.1 visits in clinics with 4-100 PCPs, but the estimates are not significantly different from each other. In sum, differences by the size of the clinic are minimal. Further, given that the aggregate level of primary care visits is unchanged, this does not appear to be a main explanation for effects.

[50]Table A.33 breaks clinic size into three categories: 1 PCP, 2-3 PCPs, 4+ PCPs. It shows that effects are largest among patients who belong to clinics with one PCP. However, this case is by definition an individual clinic and patients by definition do not have a replacement PCP at the clinic.

Table 1.7: *Treatment Effect of a PCP Leaving a Clinic*
Utilization of Clinic Based Services by Focal Clinic Practice Size

Within Open Clinics	1-3 PCPs		4-100 PCPs		
Type	Mean	Impact	Mean	Impact	P-Value
No. of Primary Care Visits	5.4	-0.75***	5.3	-0.76***	0.88
		(0.063)		(0.043)	
		-13.9%		-14.3%	
No. of PCP Visits at Clinic	4.1	**-2.0*** **	4.3	**-1.8*** **	0.04
		(0.068)		(0.048)	
		-48.8%		-42.8%	
No. of PCP Visits at Other Clinics	1.3	**1.2*** **	1.019	**1.072*** **	0.004
		(0.051)		(0.037)	
		98.1%		105.2%	
No. of Specialist Visits	9.7	**0.45*** **	9.3	**0.23*** **	0.007
		(0.064)		(0.059)	
		4.7%		2.5%	
No. of Emergency Department Visits	0.80	0.036***	0.73	0.028***	0.49
		(0.0099)		(0.0088)	
		4.5%		3.8%	
Treated PCP Sample Size	1931		5655		
Control PCP Sample Size	8398		2039		

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "No." indicates that the outcome is the yearly number. "No." indicates that the outcome is the yearly number and "Avg." indicates that the outcome is the average over the year. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Clinics with 1-3 PCPs were compared to those with 4-100 PCPs because 3 PCPs was the median practice size (7 is the mean). Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. See Section 1.8.3 for more details on how small and large groups were created.

1.8.4 Ruling Out Differences in Leaving and Replacement PCP Practice Patterns

There is a growing literature finding that PCP practice styles explain 2-3% of the variation in long-run total utilization (Kwok, 2019; Fadlon and Parys, 2019). As a result, it is possible that the sustained decrease in primary care visits and increase in specialty care visits are driven by differences between exiting and replacement PCPs propensity to refer to specialists.

I test for the importance of this hypothesis by controlling for specialist and primary care utilization of replacement PCPs' indirectly affected patients from $1 \leq t \leq 12$ (i.e. the

leave-out-mean).[51] In order to attribute utilization to a specific PCP, I assign all patients to the modal PCP seen from $1 \leq t \leq 12$.[52] Treated patients who do not see a new PCP from $1 \leq t \leq 12$ (31% of patients) and treated patients with replacement PCPs who only see treated patients over the relevant period (2% of patients) are not included in analyses. This is a limitation of this method and of the literature more generally.

The magnitude of the coefficients does not significantly change when controlling for the leave-out-mean of specialist and primary care utilization (Table 1.8). The number of primary care visits decline by 0.90 visits (SE 0.032), which is not significantly different from the decline of 0.91 visits (SE 0.032) in the uncontrolled results. The number of specialist visits is also not significantly different when controlling for indirectly affected patients use of specialists (0.49 vs 0.49 visits). These results show that practice styles explain virtually none of the observed long run decrease in average primary care and specialty use. This rules out changing practice styles as an explanation for the long-term decline in primary care visits and increase in specialty visits.

[51] I control for 100 quantiles of SP and PCP utilization.

[52] Utilization from $1 \leq t \leq 12$ was used to follow other work, namely Kwok (2019). Using this assignment mechanism, I find the median PCP sees 76 non-focal patients over this period.

Table 1.8: *Treatment Effect of a PCP Leaving a Clinic Controlling for Replacement PCP Practice Patterns*

		Without Controls	With Controls
Type	Mean	Impact	Impact
No. of Primary Care Visits	5.4	-0.91***	-0.90***
		(0.032)	(0.032)
		-16.8%	-16.5%
No. of Specialist Visits	9.5	0.49***	0.49***
		(0.035)	(0.035)
		5.1%	5.2%
Treated PCP Sample Size	10334		10334
Control PCP Sample Size	10421		10421

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. The leave-one-out average of utilization pf replacement PCP's patients was controlled for non-parametrically using 100 quantiles of the replacement PCP's non-focal patient's PCP and specialist utilization. See Section 1.8.4 for additional definitions.

1.8.5 Alternative Mechanisms

There are at least two other hypotheses that could explain why patients substitute from the primary care to specialty setting in response to a PCP's exit. First, replacement PCPs may not compensate for the loss of the main PCP because they do not realize they are now responsible for the entirety of a patient's primary care. If this was the case, one would expect patients who stay at the main clinic to be unaffected because PCPs who care for patients that remain should be aware that their co-working PCP left. Table 1.4 shows that patients who belong to clinics that stay open after a PCP's exit still substitute away from the primary care towards the specialty care setting, which goes against what we would expect if this was a primary mechanism.

Second, replacement PCPs may treat new patients differently from their other patients. PCP visits would therefore decline if replacement PCPs are not prioritizing new patients.

If this was a main explanation, the rate of PCP visits should slowly recover over the four years post-departure as the replacement PCP-patient relationship grows. Further, PCP visits should mostly decrease on the intensive margin. Instead, I show that patients decrease their use of primary care long-term and primarily on the extensive margin.

More generally, if PCP behavior was driving patients' shift to specialists, one would expect preventive care—such as preventive screens and flu vaccines—to still be administered by replacement PCPs. This is because preventive screens and flu vaccines are clearly targeted metrics that are straightforward for new PCPs to administer. They are also not in a specialist's scope-of-practice, so for specialists to absorb these procedures, they are presumably doing so at the patient's request. Therefore, mechanisms that surround PCP, opposed to patient, behavior are hard to reconcile with PCPs administering less and specialists administering more preventive care.

1.9 Conclusion

Relationships patients have with their PCPs matter. The provision of primary care within the context of a long-term relationship leads to higher quality health care. This is because relationship-specific capital exists, grows over time, contains health-specific information, and is nontransferrable.

Evaluating the importance of relationship-specific capital among Medicare patients may be a context where relationships are especially important because Medicare patients are older and in worse health. Medicare patients also face low and relatively constant out-of-pocket costs across providers and have large teams of PCPs and specialists. While constant costs more precisely identifies patient preferences, it is unclear how results would generalize to younger patients who may face network constraints, need a PCP for referrals, and have less attachment to individual PCPs. For instance, requiring that a patient see a PCP to be referred to a specialist may effectively push patients back to primary care, encouraging patients to start new relationships. Further, younger patients may be more likely to choose convenience over building a relationship, using retail and urgent care clinics as a regular

source of care.

Patients who are more able to substitute to pre-existing PCPs are able to maintain better health after the loss of a PCP. In addition to teams, which this work speaks to, organizations such as independent practice associations may be a helpful bridge for patients as they attempt to re-establish care. This is especially relevant for patients who belong to clinics that close when a PCP exits, who cannot switch to PCPs within the main clinic.

Findings clearly affirm the theoretical role of the PCP as a point of contact to the rest of the health system as well as an administrator of preventive care (Starfield, 1994). This role is receiving more focus as recent and proposed policy reforms directly target the PCP-patient relationship. For instance, Medicare's Shared Savings Program shares savings with PCPs who keep total costs below a financial benchmark, encouraging PCPs to take greater control over the continuum of patient care (McWilliams, Hatfield, Chernew, Landon and Schwartz, 2016).

Continuing to develop our understanding of PCP-patient relationships helps firms and policymakers mitigate the harms of disrupting them. PCP departures are especially ubiquitous in the health care context where volatile insurance networks and non-compete agreements may artificially sever a PCP-patient relationship, even if the PCP remains in practice. For instance, 26% of commercially insured individuals switch insurance plans in a given year and 39 states enforce non-compete agreements (Barnett, Song, Rose, Bitton, Chernew and Landon 2017 and Hausman and Lavetti 2016). PCP retirements are also projected to increase over the next decade, with 32% of PCPs currently over age 60 (Sabety, 2020). As a result, interventions such as team care may be better able to maintain patients' health as a growing number of PCPs leave clinical practice.

Chapter 2

The ActionHealthNYC Experiment: Coordinating Undocumented Immigrants' Health Care[1]

[1]This chapter is joint work with Jonathan Gruber, Rishi Sood, and Jin Yung Bae. This work was supported in part by the Robin Hood, Rockefeller, and Altman Foundations as well as National Science Foundation Graduate Research Fellowship Program (Grant No. DGE1144152). The content is solely the responsibility of the authors and does not necessarily represent the official views of the New York City Department of Mental Health and Hygiene.

Dissertation Advisor:
David Cutler

Author:
Adrienne Hope Sabety

Essays on the Institutional Design of Health Care Markets

Abstract

Using a randomized evaluation, we measure the impact of coordinating undocumented immigrants' health care on utilization and health status. We show that coordinating undocumented immigrants' health care increases utilization of primary care, self-reported access to care, and preventive care, leading to a 7% decrease in individuals' long-run probability of death in back-of-the-envelope calculations. The intervention also causes high risk individuals to decrease their use of the emergency department by 26%, leading to $19,000 in government savings on net.

2.1 Introduction

Around 7 million undocumented immigrants remain uninsured a decade after the Patient Protection and Affordable Care Act expanded health insurance coverage to 17 million Americans (Tolbert, Orgera, Singer and Damico, 2019). Most of these immigrants face large informational and psychological barriers to accessing health care (Barcellos, Goldman and Smith, 2012). This may not only adversely affect undocumented immigrants' health and human capital but also spillover onto their citizen children, where marginal returns to health capital investments are highest (Watson, 2014; Goodman-Bacon, 2018).

However, it is challenging to design policy targeting undocumented groups. We know very little about undocumented immigrants' health status and needs and, what we do know, relies on imputing documentation status from surveys, which may introduce bias. For instance, imputation methods often categorize H1-B workers as undocumented, who tend to be well educated and work in relatively high-paying jobs (Capps, Bachmeier, Fix and Hook, 2013).

Survey data, where documentation status is imputed, show that undocumented immigrants appear to be in relatively better health and use less health care than natives (Rodríguez, Bustamante and Ang, 2009; Artiga, Damico, Young, Cornachione and Garfield, 2016; Goldman, Smith and Sood, 2006). In contrast, policymakers have claimed that poor primary care access leads undocumented immigrants to use the emergency department (ED) for primary care treatable conditions (Goldstein, 2019). It is therefore difficult to design policy around utilization patterns and health profiles of undocumented immigrants without understanding baseline health needs.

In this paper, we provide the only information on undocumented immigrants' demographics and health status from survey and administrative data precisely observing individuals' documentation status. We then design a randomized intervention to understand how expanding access to high quality, coordinated care affects undocumented immigrants' health care utilization and health status. Our intervention, ActionHealthNYC, ran for 14 months and included 2,404 low-income, uninsured, undocumented immigrants in New

York City (NYC).[2]

Based on baseline survey responses, we show that undocumented immigrants have less access to primary care than their low-income, citizen counterparts. Only 25% of individuals report having a PCP, a widely used measure of access to health care services, compared to 60% among a low-income, uninsured, native population (Sommers, Blendon and Orav, 2016).[3] Despite low levels of self-reported access, 60% of ActionHealthNYC individuals reported visiting a physician, 20% the ED, and 6% experienced a hospitalization in the past year. These rates are similar to low-income, uninsured, native populations, where 56% visited a physician, 21% visited the ED, and 17% experienced a hospitalization in the past year (Sommers et al., 2016).

We then leverage our experimental results to quantify how ActionHealthNYC causally affected patients' use of primary care, preventive care, and emergency care. ActionHealthNYC expanded access to coordinated care with the goal of improving health and decreasing ED use. To "coordinate" individuals' care we made initial appointments for treated individuals at one of nine primary care homes (PCHs), which provided USPSTF A+B recommended preventive screens.[4] For patients at the highest risk of using the ED, those who were housing insecure or with one or more chronic conditions, received an enhanced care coordination (ECC) intervention.[5] ECC included pre-visit planning, referrals to other social services, and a minimum of six encounters with an assigned care coordinator. Analyses present comparisons of outcomes between those selected and not selected by randomization to participate in ActionHealthNYC.

[2]NYC is an ideal setting for our experiment because it contains more than a million undocumented immigrants, making it the metropolitan area with the largest number of undocumented immigrants nationwide (Passel and Cohn, 2017).

[3]This discrepancy is particularly notable given the robust safety-net health system in NYC.

[4]PCHs refer to clinics that provide primary care services, with an emphasis on care coordination. Our PCH system included two Health+Hospital (H+H) and seven federally qualified health centers (FQHCs) chosen to be in neighborhoods with large undocumented populations.

[5]We define chronic conditions to include a diagnosis of hypertension, diabetes, asthma, mental illness, cardiovascular disease, chronic obstructive pulmonary disease, congestive heart failure, Atrial Fibrillation, cancer, HIV+, Hepatitis C, or substance abuse disorder.

We find that the intervention causally increased self-reported visits to a physician by 22% among treated individuals. ActionHealthNYC also improved self-reported health care access: treated individuals were 55% more likely to report having a PCP than control individuals. Increased primary care access translated into individuals receiving a greater amount of preventive care. Treated individuals had a 62% increase in diabetes screens and a 72% increase in blood pressure screens relative to the rate reported on the baseline survey.[6] Relying on estimates from the clinical literature, this translates into a 7% decrease in the long-run probability of death. Effects were strongest among high risk individuals.

We find that the intervention caused a 26% decrease in ED visits among high risk patients, a decrease driven by a 59% decline in ED visits for primary care treatable conditions. In turn, ED charges decreased $55,500, translating into $19,000 net savings among *all* patients. The literature to-date has found a mixed relationship between access to care and ED use.[7] We significantly add to the literature by focusing on an undocumented, under-served population where informational and psychological barriers are likely quite large. Our results suggest that expanding access to *coordinated* primary care has the potential to decrease ED use among individuals that are the most likely to use the ED.

The paper proceeds as follows. Section 2 outlines the experimental design. Section 3 describes the survey and administrative data sources. In Section 4, we describe the empirical strategy and identification assumptions. Section 5 presents the results and Section 6 concludes.

[6]We focus on diabetes and blood pressure screens because these were the only two screens surveyed on the baseline survey.

[7]On one hand, Finkelstein, Taubman, Wright, Bernstein, Gruber, Newhouse, Allen, Baicker and Group (2012) find that increasing access to Medicaid increases ED use. On the other hand, quasi-experimental studies, such as Miller (2012) focused on Massachusetts' Health Reform, find that ED utilization is negatively correlated with access to insurance (Sommers and Simon, 2017).

2.2 Experimental Design

2.2.1 Enrollment

Our target population included 345,000 undocumented and uninsured New Yorkers over age 19 and under 200% of the federal poverty line (FPL). ActionHealthNYC enrollment began in May 2016, the follow-up survey occurred in May 2017, and the program ended in June 2017 (Figure B.1).

Designing an effective outreach strategy for an undocumented population with a high level of mistrust was challenging. Our most effective strategy was partnering with six community based organizations (CBOs) who referred their clients to ActionHealthNYC.[8] We also used paid and earned media coverage, social media outreach as well as IDNYC (NYC's municipal card) and Emergency Medicaid mailings to reach potentially eligible individuals. Our outreach strategy may mean that our population is less healthy and more engaged than the average NYC undocumented immigrant. This may have positively affected our ability to obtain follow-up survey responses while also affecting the generalizability of our findings.

As shown in Figure 2.1, 6,094 individuals applied for ActionHealthNYC. Individuals enrolled by scheduling an appointment at one of seven enrollment sites, which included two human resources administration (HRA) Medicaid offices,[9] one Department of Health and Mental Hygiene (DOHMH) health clinic, and four of NYC's Health+Hospital (H+H) facilities (Bellevue Hospital, Gourverneur Health, Elmhurst Hospital, and Queens Hospital). The enrollment process required two separate appointments. In the first appointment, certified application counselors (CAC) enrolled applicants in insurance through New York's insurance marketplace.[10] If the individual was ineligible for insurance through the mar-

[8]However, many of those who showed up did not enroll. For instance, Make the Road, our most active CBO, was responsible for 1,018 initial appointments, of which only 285 enrolled.

[9]Part of the Department of Social Service.

[10]We also enrolled all individuals in IDNYC and pre-enrolled individuals in Emergency Medicaid (EM), which covers undocumented immigrants' for qualifying conditions. IDNYC doubles as a prescription drug discount card, offering discounts of 18% off brand name drugs and 55% off generic drugs. By pre-enrolling individuals in EM, Medicaid would cover hospitalizations.

ketplace they were, by process of elimination, undocumented and therefore eligible for ActionHealthNYC.[11] ActionHealthNYC was targeted towards individuals with income under 200% of the federal poverty line, which we verified through tax documents or letters from individuals' employers with their income.[12] We also enrolled 266 in other programs, such as Medicaid and Medicare, 234 were ineligible, and 358 declined to continue to the second ActionHealthNYC enrollment appointment. Those who declined enrollment likely had a level of discomfort that exceeded the expected gains from enrolling.[13]

Among the 2,404 individuals that returned for a second appointment, all individuals completed the baseline survey and were subsequently randomized to the treatment (1,265 individuals) or the control arm (1,139 individuals).[14] We randomized couples into the same arm, where 598 individuals were randomized with a partner and 1,806 individuals were randomized alone. Enrollers offered to make initial appointments for treated individuals at one of our nine primary care homes (PCHs) at the end of the appointment.

2.2.2 Program Design

We created "Primary Care Homes" (PCHs) for individuals by contracting with NYC's safety net health care providers: federally qualified health centers (FQHCs) and H+H facilities. We offered two H+H and seven FQHC PCH sites specifically chosen because of their location in neighborhoods with large undocumented populations (Figure B.2 shows their geographic distribution in NYC). If an individual obtained a referral from their PCH, they could access care at 46 different sites in NYC. The two H+H facilities provided primary

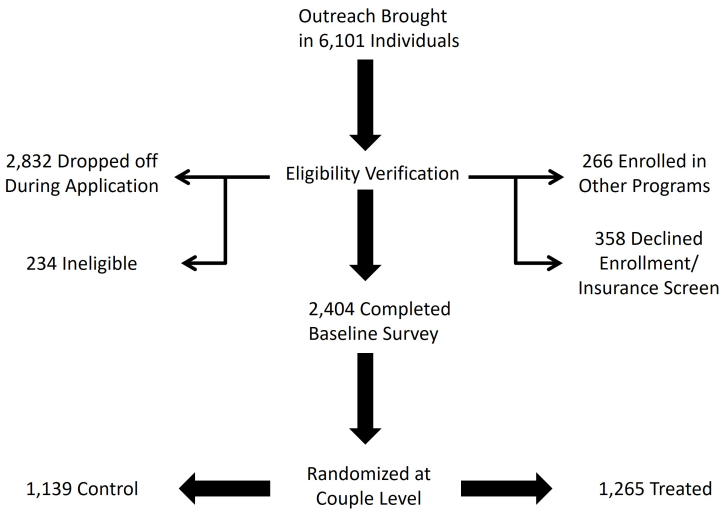
[11]For instance, in New York's Medicaid program covers green card holders and those permanently residing under color of law.

[12]Information on other social programs individuals qualified for—such as public education, food stamps, free legal immigration help, as well as domestic violence and homeless prevention programs—were also provided at the end of the initial appointment.

[13]Unfortunately, we are not able to observe individuals enrolled in other programs because we did not obtain their consent.

[14]The drop off between the first and second appointment may be partially due to the two appointment nature of enrollment.

Figure 2.1: *Enrollment Flow Chart*



and specialty care, while FQHCs provided exclusively primary care. Of the two H+H PCH sites, one was H+H Gouverneur Health located in lower Manhattan and the other was H+H Elmhurst Hospital Center located in Elmhurst, Queens–a neighborhood with the most immigrants as a percentage of the population in NYC (Bloomberg, Burden and Shama, 2013). Our seven FQHC sites included three Charles B. Wang Community (CBW) Health Centers–two in Flushing Chinatown (Queens) and one in Chinatown (Manhattan).[15] Community Healthcare Network (CHN) also sponsored a PCH in Corona (Queens) another predominately immigrant neighborhood (Bloomberg et al., 2013).[16] Ryan-NENA Community Health Center, located in Alphabet City (Manhattan), was also a PCH.[17]

ActionHealthNYC targeted improving undocumented immigrants' health care from both the supply and demand side. On the supply side, we gave PCHs a capitated $35.25

[15]CBW has five sites throughout Manhattan that members could access if referred.

[16]CHN operates 12 locations in total that were open to ActionHealthNYC members if referred.

[17]Ryan-NENA Community Health Center is part of the William F. Ryan Community Health Network (WR), which has five locations in total that were open to members with referrals.

per member, per month payment based on the number of members choosing the PCH in the initial appointment. H+H received an additional $24 per member per month for providing specialty services to our entire population. For PCHs to receive payments, they had to report the number and type of USPSTF A+B recommended preventive screenings received by members. Individuals in FQHC PCHs could access specialty care from H+H if referred. Anecdotal evidence suggests that members valued access to specialty services. DOHMH not only made H+H specialty appointments for program members but also leaned on specialists to prioritize members above non-members. We also designed a targeted enhanced care coordination (ECC) intervention for high risk individuals or individuals who were homeless or had one or more uncontrolled chronic conditions. Our definition of having a chronic condition included a diagnosis of hypertension, diabetes, asthma, mental illness, cardiovascular disease, chronic obstructive pulmonary disease, congestive heart failure, Atrial Fibrillation, cancer, HIV+, Hepatitis C, or substance abuse disorder. Based on these guidelines, enrollment in the ECC program occurred during the initial visit depending on the physician's discretion. The ECC program included assigning members a care coordinator who identified barriers/facilitators to care, patient-led goals, nursing or social work interventions, and social determinants of health. The ECC intervention also included six or more contacts from the clinic, pre-visit planning, as well as prescription medication refills and pick-up.

On the demand side, individuals chose PCHs with the understanding that the PCH would provide the majority of their care.[18] Each PCH adopted the ActionHealthNYC sliding fee-scale for members providing a lower and, more importantly, *certain* price for health care. Table B.1 shows the costs for members compared to what the control group faced.[19] Individuals with incomes between 0% and 150% of the FPL faced different costs than those

[18]The program does not cover out-of-network care, but individuals can seek care at non-PCH sites affiliated with their primary PCH when referred.

[19]The costs that the control group sees are equivalent to the cost-sharing schedule that all individuals paid before ActionHealthNYC existed.

between 150% and 200% of the FPL.[20,21]

2.3 Data

Our survey and administrative data identify ActionHealthNYC's impact on individuals' health status, utilization, and financial burden. We linked data sets using an individual's date of birth, sex, and first and last name. The four data sets and the periods of time covered is diagramed in Figure B.1.

The survey data included a baseline and follow-up survey. The baseline survey asked 75 questions completed by *all* individuals who signed up for the program. We translated the survey into 32 languages and asked four main categories of questions covering demographics, utilization, self-reported health, and financial burden. We administered the follow-up survey nine months after the program began enrolling members in May 2016. It included 23 questions covering the same themes as the baseline survey with a response rate of 45% for the controls and 55% for the treated group. Given the national dialogue, the response rate may have been lower than it would have been otherwise. We intend to adjust for non-response bias using baseline results for the entire population.

Our primary administrative data source was the New York State Department of Health's Statewide Planning and Research Cooperative System (SPARCS) hospital discharge database from 2014-2017. Every short-term nonfederal hospital in New York State submits discharge data through SPARCS, which the Department of Health reviews for quality and completeness. The data contain rich, claims level information on each ED visit and inpatient admission, including name, date of birth, race, sex, ICD-9 and ICD-10 diagnosis codes, procedure codes, charges, primary payer, and out-of-pocket costs. We classified ED visits for primary care treatable conditions by linking ICD-9 and ICD-10 codes to an algorithm

[20] 94% of the population enrolled in the experiment is below 150% of the FPL.

[21] However, non-payment is likely always an issue–patients can always choose to not pay the bill and FQHCs rarely send the bills to collection. For example, William F. Ryan Community Health Network and Urban Health Plan offer a discount if the patient pays upfront, suggesting that non-payment may be a significant issue.

built by Billings et al. (2000) and updated by Johnston et al. (2017). The main limitation of SPARCS is that care delivered outside of New York State is not observed.

PCH administrative data is our secondary administrative data source because it only captured visits made to our seven PCHs over our program period, May 2016 - June 2017. Data included a count of specialty and primary care visits as well as chronic condition diagnoses and preventive screens. The H+H data included visits made to Elmhurst and Gouverneur hospitals. The FQHC data included visits made to Community Health Network (CHN), Charles B. Wang (CBW), William F. Ryan Community Health Center (WFR), and Urban Health Plan (UHP). This information was only available for control and treatment individuals who used our PCH facilities, so it may not be representative of the full sample.

2.4 Empirical Strategy

2.4.1 Estimating Equations and Identification

Analyses measure the causal effect of enrollment in ActionHealthNYC on health care utilization and health status. We randomize individuals into eligibility, so the intent to treat (ITT) and treatment on the treated (TOT) are equivalent. We assume that the local average treatment effect is also equal to ITT and TOT by assuming that outcomes are only affected through enrolling in ActionHealthNYC and its subsequent effect on utilization. Further, monotonicity holds in this setting because ActionHealthNYC eligibility is conditional being in the treatment group.

Regressions are at the individual i, time t level:

$$y_{it} = \mu_i + \lambda_t + \beta_1 \text{Enrolled in Program}_i \times \mathbb{1}(Post)_t + \epsilon_{it}$$

μ_i are individual and λ_t are time fixed effects. Time fixed effects include two and one year prior to enrollment as well as one year after enrollment for outcomes using SPARCS data and one year prior to as well as one year after enrollment for outcomes using PCH data.

Enrolled in Program is one for individuals enrolled in ActionHealthNYC. In all analyses we used linear probability models (LPMs) and clustered at the couple level to account for the randomization of couples together.

2.4.2 Balance Tests

In order to infer causal effects, the treatment must be randomly assigned and treatment individuals should not be differentially selected within sub-samples. We are able to consistently estimate the causal effect of interest under the assumption that potential outcomes are independent of treatment assignment conditional on stratification variables. To test that treatment is as good as randomly assigned, Table A.6 illustrates how the treatment and control groups differ at baseline, conditional on completing the follow-up survey, and conditional on using one of our nine PCHs during the program period. P-values and joint f-statistics are also reported.

Columns 2-4 of Table A.6 show that treatment groups are well-balanced on baseline observables with three exceptions. Treated individuals are less likely to be randomized alone (3.7pp), have a higher rate of employment (5.1pp), and have slightly fewer inpatient visits, conditional on having at least one (0.7pp). Columns 5-7 of Table A.6 compares baseline observables between treatment groups conditional on having completed the follow-up survey: 581 treated and 486 control individuals. Again, these groups look well balanced, with the exception of controls being more likely to speak English well than those treated and that treated individuals self-report better health status. All analyses include individual fixed effects to account for these baseline differences.[22]

Columns 8-10 illustrate how treatment and control groups differed at baseline among individuals who used our PCHs during the treatment window. The comparison enabled us to asses the viability of comparing treated and control individuals using the PCH data (i.e. Table 2.4 and 2.4). Groups look fairly similar, except for PCP and specialist utilization from the PCH data, which is mechanical.

[22] Outstanding: correcting for non-response

Table 2.1: *Balance Table for Individuals Survey Data*

	Baseline Survey Respondents			Follow-up Survey Respondents			Conditional on Using PCH in Treatment Window		
	Treated	Control	P-Value	Treated	Control	P-Value	Treated	Control	P-Value
No. of Individuals	1,265	1,139		581	486		809	242	
Rand. Alone (%)	**73.4**	**77.1**	0.03	73.5	77.6	0.12	73.0	74.4	0.68
Covariates at Baseline									
Age (yr)	44.6	44.3	0.43	43.7	43.3	0.60	45.2	45.4	0.80
High Risk (%)	43.6	42.6	0.60	42.3	42.8	0.88	42.5	45.9	0.36
Single (%)	52.8	53.6	0.71	54.9	51.0	0.21	52.2	50.4	0.63
Yrs in U.S. (yr)	13.8	13.2	0.07	14.1	13.5	0.16	**13.4**	**12.0**	0.02
English (%)	29.7	30.7	0.60	**26.8**	**32.6**	0.04	28.8	29.8	0.77
High School (%)	47.9	47.2	0.74	46.0	44.6	0.67	48.0	47.1	0.82
Employed (%)	**61.3**	**56.2**	0.01	63.2	58.0	0.09	**61.7**	**49.2**	$p< 0.001$
Female (%)	55.2	56.3	0.66	60.1	56.8	0.37	55.3	62.8	0.08
Hispanic (%)	49.7	49.0	0.72	51.8	47.1	0.13	**44.1**	**31.0**	$p< 0.001$
$<$ 100 FPL (%)	81.2	82.0	0.64	83.3	83.3	1.00	83.7	86.5	0.33
Housing Insecure (%)	12.9	10.9	0.14	11.9	12.4	0.81	12.4	12.2	0.93
Joint Test: P-Values			0.33			0.28			$p< 0.001$
Utilization at Baseline									
Has a PCP (%)	25.7	24.9	0.67	25.7	26.6	0.76	24.1	30.3	0.05
Prob. Doc Visit (%)	59.8	60.6	0.70	61.1	60.8	0.92	58.3	56.4	0.61
Prob. ED Visit (%)	19.7	20.0	0.85	20.5	17.9	0.28	18.5	15.3	0.26
Prob. Inpatient Admission (%)	6.4	5.5	0.36	5.0	4.4	0.61	5.7	5.1	0.70
Joint Test: P-Values			0.75			0.87			0.08
Health Status at Baseline									
Diabetic (%)	8.8	8.8	0.95	9.5	8.6	0.64	8.2	7.9	0.88
Hypertensive (%)	16.8	18.3	0.33	14.5	18.7	0.06	16.6	18.6	0.46
Asthmatic (%)	4.5	4.6	0.94	4.8	3.7	0.37	**5.3**	**2.1**	0.03
Mental Illness (%)	4.4	4.0	0.63	4.1	3.9	0.85	4.5	4.1	0.83
Smoker (%)	11.1	12.1	0.42	12.2	12.8	0.79	13.5	12.8	0.78
$\geq$ Good Quality of Care	82.2	82.1	0.97	84.2	86.0	0.47	81.6	82.0	0.91
$\geq$ Fair Health	37.8	34.3	0.07	**41.9**	**35.9**	0.04	39.3	39.4	0.97
PROMIS Physicial Health Score	14.4	14.6	0.12	14.3	14.7	0.05	**14.4**	**14.8**	0.04
PROMIS Mental Health Score	13.5	13.6	0.44	13.4	13.5	0.80	13.4	13.1	0.25
Joint Test: P-Values			0.75			0.20			0.09
Self Reported Costs at Baseline									
Amount Spent on Medical Care ($)	115.8	119.2	0.81	112.9	136.2	0.30	106.6	81.3	0.22
Amount Spent on RX ($)	61.0	73.2	0.33	55.4	62.5	0.66	66.7	56.1	0.67
Did Not Fill RX Bc of Cost? (%)	14.4	14.0	0.76	14.2	13.1	0.60	14.7	9.8	0.06
Postpone Paying Bills? (%)	12.6	11.6	0.43	13.5	10.7	0.17	**13.0**	**5.4**	0.001
Joint Test: P-Values			0.96			0.62			0.05
Utilization From Admin Data									
Prob. ED Visit (%)	17.1	17.7	0.67	18.9	17.3	0.49	16.4	12.4	0.13
No. ED Visits	1.7	1.6	0.76	1.8	1.8	0.97	**1.6**	**2.3**	0.03
Prob. IP Visit (%)	5.1	4.7	0.58	5.5	4.9	0.68	4.8	4.1	0.66
No. IP Visits	**1.2**	**1.9**	0.02	1.2	1.7	0.20	**1.1**	**3.1**	0.006
Utilization in Population that Used PCHs									
Prob. PCP Visit (%)	8.8	7.5	0.25	12.4	9.9	0.20	**11.4**	**27.7**	$p< 0.001$
No. PCP Visits	4.6	4.3	0.63	4.2	4.0	0.76	4.4	4.3	0.77
Prob. Specialist Visits (%)	5.1	4.0	0.23	5.9	5.1	0.61	**6.2**	**14.9**	$p< 0.001$
No. Specialist Visits	5.3	5.7	0.82	4.4	4.7	0.78	5.8	6.4	0.73

Notes: Outcomes are from the baseline survey and 44% of individuals completed the follow-up survey. The randomization was done at the couple level, so couples would be randomized into the same arm. Bolded estimates indicate that values are significantly different at the 5% level. "No." abbreviates number and "Prob." abbreviates probability. A higher physical or mental health PROMIS score means individuals experienced worse physical or mental health.

2.5 Results

2.5.1 New Information on Hard to Reach Population

To our knowledge, all information on undocumented immigrants residing in the United States relies on survey responses imputing individuals' documentation status. Methods assigning immigration status to non-citizens may make immigrant groups look healthier by mistakenly including documented individuals in the undocumented group, such as H-1B workers who tend to be well educated and work in relatively high-paying jobs (Capps et al., 2013). Given that undocumented immigrants face real fears of deportation, non-response rates may also be higher among undocumented groups, biasing estimates even further towards documented immigrants.

We build on prior work relying on imputation methods by targeting undocumented immigrants through outreach and, ultimately, enrollment. We designed our enrollment process to determine documentation status without explicitly asking about documentation status, a question that is widely considered harmful for turnout and engagement. Our method relied on attempting to enroll individuals in other types of insurance before considering them eligible for our program. The idea was that documented individuals should qualify for insurance options, whereas undocumented groups would not. This enabled us to obtain a group of undocumented immigrants through process of elimination.

While precisely observing documentation status improves on previous limitations, estimates may not generalize nationally. ActionHealthNYC targeted undocumented immigrants with incomes below 200% of FPL who did not qualify for other insurance options. Further, we provided access to health care, so our group of immigrants may be selected to be in worse health than the average undocumented immigrant in NYC.

In order to assess the generalizability of our population, we start by benchmarking our population's demographics against national and statewide surveys of undocumented groups, which impute immigration status. National and state estimates are from the Migration Policy Insitute (MPI).

Income levels are the most significant difference. ActionHealthNYC exclusively enrolled individuals under 200% of FPL, whereas nationally 62% and statewide 55% of undocumented immigrants have incomes less than 200% FPL. Apart from income, demographic characteristics of race, education, age, sex, marital status, employment, and English proficiency look fairly similar between ActionHealthNYC individuals and national and state averages. Nationwide, 76% of undocumented immigrants are Hispanic and 11% are Asian; statewide, only 65% of undocumented immigrants are Hispanic and 22% are Asian. We find that 48% of our population immigrated from Central or South America and 27% of our population immigrated from China or Korea. The slightly higher share of Chinese and Koreans relative to Hispanics is likely specific to NYC, but may also be because we included several PCHs that specifically cater Chinese and Koreans (e.g. Charles B. Wong Community Health Center in Chinatown). As for the percent with at least a high school education, 53% nationwide and 62% statewide undocumented immigrants report having at least a high school education. Comparatively, a slightly lower share of 47% ActionHealthNYC individuals had at least a high school education. Nationally, 72% of undocumented immigrants are between 19-44 and, statewide, 42% of undocumented immigrants are between age 35-54. Although comparing to age ranges is not perfect, ActionHealthNYC individuals seem similar, with an average age of 44. Slightly less than half of nationwide and statewide undocumented immigrants are female (45% and 47%), whereas slightly more than half of ActionHealthNYC individuals are female (55%). Employment rates of 65% nationally and statewide are similar to the 60% of ActionHealthNYC individuals who report being employed. English proficiency is 30% nationally and 35% statewide. Similarly, 30% of ActionHealthNYC individuals report speaking English at least well. Taken together, these characteristics suggest that our population is fairly similar to national and New York state undocumented groups, with the one exception of income.

We then report new information on undocumented immigrants, focusing focus on housing insecurity, self-reported utilization of health care and health status, and self-reported and administrative-based measures of health expenditures. The full set of outcomes are in

Table XinprogressX.

We find that 12% of individuals are housing insecure, whereas nationally 0.2% of individuals are homeless (Cox, Rodnyansky, Henwood and Wenzel, 2017). This suggests that housing insecurity is likely an important determinant of undocumented immigrants' health (Cutts, Meyers, Black, Casey, Chilton, Cook, Geppert, De Cuba, Heeren, Coleman, Rose-Jacobs and Frank, 2011). We use a widely used metric of access to healthcare—having a PCP—and show that only 25% of individuals report having a PCP. This is lower than the 58% of undocumented Latinos (imputed) and 60% of low-income, Medicaid-eligible individuals found by Rodríguez et al. (2009) and Sommers et al. (2016), suggesting that access to health care is a large issue for our population.

Inspite of low levels of self-reported access, 60% of ActionHealthNYC individuals reported visiting a doctor, 20% reported visiting the ED, and 6% reported a hospital admission in the past year.[23] These estimates are similar in our SPARCS administrative data, where 17% of individuals visited an ED and 5% were hospitalized over the previous year. Our undocumented population seems to use the health care system at a similar rate to low-income, uninsured, native populations, where 56% visited a physician, 21% visited the ED, and 17% were hospitalized in the past year (Sommers et al., 2016).

Self-reported health status among ActionHealthNYC individuals seems slightly better than natives. Table 2.2 shows that 9% of ActionHealthNYC individuals self-report being diabetic and 18% self-report being hypertensive. We then adjust these estimates by the number of individuals found to have diabetes or hypertension when visiting one of our PCHs to determine something closer to the true prevalence in the population. With this adjustment, we find that 13% of individuals are diabetic and 25% are hypertensive. We compare these estimates to the native born from the National Health and Nutrition Examination Survey

[23]Other work focusing on utilization in an "undocumented" group imputes documentation status from the Medical Expenditure Panel Survey, finding a 62% insurance rate among the imputed documented group (Tarraf, Vega and Gonza, 2014). This is highly questionable because most other estimates find a 29% insurance rate among undocumented immigrants (e.g. Sommers and Parmet (2015)). As a result, we view a low-income, uninsured group to be a better bedchamber for our low-income, uninsured population because insurance status is likely quite important in driving utilization decisions.

(NHANES). We focus on NHANES respondents over age 18, with incomes under 200% FPL, and weight responses among Hispanics and Asians to be similar to the composition in the ActionHealthNYC sample. Table 2.2 shows that self-reported and clinically adjusted health of ActionHealthNYC enrollees is very similar, if not healthier, than the native born.

Lastly, we show that self-reported health expenditures are quite low, with only 14% not filling a prescription because of cost and 12% postponing the payment of other bills in order to pay for health care. In contrast, Sommers et al. (2016) surveys a low income population and finds that 39% of individuals did not fill a prescription because of cost and 42% had trouble paying medical bills. The lower rates of financial strain among ActionHealthNYC individuals may be a function of undocumented immigrants low levels of access as well as NYC's robust safety net.

Table 2.2: *Health Status Differences Across ActionHealthNYC and NHANES Immigrant and Native Born*

Outcome	ActionHealthNYC Undocumented	NHANES Immigrants	NHANES Native Born
Diabetic			
Self-reported	9%	9%	8%
Clinically adjusted	13%	13%	16%
Hypertensive			
Self-reported	18%	13%	21%
Cinically adjusted	25%	18%	30%

Notes: NHANES immigrants and NHANES native born weights hispanic and asian estimates by the 48% and 27% shares observed in ActionHealthNYC data in order to make estimates comparable across groups.

2.5.2 Experimental Results

We now turn to our experimental results showing that coordinating undocumented immigrants' health care affects individuals' health care utilization and health status. We present causal, ITT estimates comparing control and treated individuals at baseline and follow-up. We report findings relying on survey data, Table **??**, as well as administrative data, Table

2.3. For results relying on survey data, analyses only contain the 44% of individuals who completed the baseline survey, whereas administrative data results contain the entire sample.

Table **??** shows that the intervention lead to a 22% aggregate increase in doctors office visits among both high and low risk individuals. This appears to be mainly driven by extensive margin increases, although there is a positive (but insignificant) increase in the number of doctors' office visits among individuals who had at least one. ED visits insignificantly decline on both the extensive and intensive margins, but the point estimates are large suggesting that we lack precision. Treated individuals seem to spend slightly more on medical care, but it does not translate into treated individuals being more likely to borrow money, skip paying other bills, or pay other bills late in order to pay health care bills (Table B.2). Treated individuals are 55% more likely to report having a PCP, with the treatment effect being largest among high risk individuals. Other health metrics do not significantly change, but the sign of the estimates suggest that our intervention positively affected patients' health.

Table 2.3 uses SPARCS administrative data to show that the intervention caused high risk individuals to decrease their use of the ED by 26% in response to ActionHealthNYC enrollment. The decline is especially large among primary care treatable conditions, which decline 59%, leading to a $55,574 decrease in ED charges among high risk individuals.[24]

We then show that the program was cost-effective in a back-of-the-envelope calculation. The costs of the program encompassed the per member, per month payment made to PCHs to coordinate individuals care. We paid PCHs $35.25 to provide primary care services and $24 to provide specialty care services, per member per month, which totaled $711 per member over the 12 month program. We also bought down fee-scales when PCHs' official fee-scale exceeded ActionHealthNYC's, at an additional $83 per member. This totaled $794 per member over the program period. Benchmarking this against the ED savings among an average individual ($19,836), suggests that the program saved $19,042 per member.

[24]When I'm able to access the data again, I will spend more time with this result, which seems implausibly large. However, it is worth noting that it does not seem sensitive to large outliers. Table B.2 shows that winsorized charges also significantly increase, where winsorized estimates top code the top 0.1% of charges.

Table 2.3: *SPARCS Administrative Data*

	All Individuals		Low Risk Individuals		High Risk Individuals	
	Mean	ITT	Mean	ITT	Mean	ITT
Utilization						
Prob. ED Visit	0.17	-0.018	0.12	**0.014**	0.23	**-0.059****
		(0.018)		(0.023)		(0.027)
		-10.3%		11.0%		-25.6%
No. ED Visits	1.7	-0.49	1.6	-0.033	1.8	-0.78
		(0.37)		(0.50)		(0.51)
		-29.2%		-2.1%		-43.7%
No. ED Visits, PCP Treatable	0.31	-0.088	0.17	**0.060**	0.48	**-0.28*****
		(0.053)		(0.050)		(0.10)
		-29.0%		35.4%		-59.0%
Charges						
ED Tot. Charges	52494.7	-19836.3**	30730.8	**7211.6**	80606.2	**-55574.3*****
		(9046.1)		(8568.8)		(17519.7)
		-37.8%		23.5%		-68.9%
Out of Pocket, ED	0.87	-0.23*	0.58	**0.054**	1.2	**-0.60****
		(0.12)		(0.12)		(0.23)
		-26.4%		9.3%		-49.0%
Treated Sample Size	1265		713		552	
Control Sample Size	1139		654		485	

Notes: Estimates are relative to the pre-period, which occurs 14 months before May 2016, the date when the intervention began. The intervention window spans May 2016 through June 2017. Emergency department (ED) and inpatient (IP) results use two years of additional pre-period data to estimate regressions with a linear time trend. "Prob." captures the extensive margin change in utilization across the period and "No." represents the number of visits conditional on non-zero utilization, or intensive margin change. The high risk patient category includes patients with one or more chronic condition, so low risk patients did not have a chronic condition diagnosis at baseline by definition. Bolded estimates indicate that High Risk and ECC Group estimates are significantly different at the 5% level. Regressions include individual fixed effects and are clustered at the couple level because couples were randomized together. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. SPARCS administrative data captures any ED or IP visit in New York State, whereas EHR data from PCHs only contains information on individuals who used our PCH facilities. See Table B.3 for additional outcomes.

2.5.3 New Diagnoses and Screens

This section focuses on results showing that individuals who visited a PCH were more likely to receive a chronic condition diagnosis (Table 2.4). To do this, we use PCH electronic health records (EHR) data. Importantly, data is selected to contain information on individuals who visited one of our PCHs, 809 (64%) treated and 242 (21%) control individuals (Table 2.5), so it is not an apples-to-apples comparison.

Table 2.4 shows the increase in the probability an individual received a new diagnosis or screen in a PCH compared to their response on the baseline survey. For instance, we asked individuals if they were hypertensive at baseline. We then compared that response to whether or not PCHs diagnosed the individual as hypertensive. We defined "high risk

individuals" to be individuals who were housing insecure or had one or more chronic condition. As a result, by definition, low risk individuals did not report having a chronic condition at baseline (excluding whether the individual was a smoker).

The probability an individual gets a diabetes screen increases by 4.3 pp and a blood pressure screen increase by 4.9 pp. We then pair the increases in diabetes and blood pressure screens with estimates drawn from Dehmer, Maciosek, LaFrance and Flottemesch (2017) and Kahn, Alperin, Eddy, Borch-Johnsen, Buse, Feigelman, Gregg, Holman, Kirkman, Stern, Tuomilehto and Wareham (2010) on the long run impact of screening on mortality. Calculations illustrated in Appendix B.1 show that the increased rate of diabetes and blood pressure screening translates to a 7% decrease in individuals' long-run probability of death.

Table 2.5 shows more care coordination details based on care delivered over the treatment period. Columns 2-4 compare rates among all individuals and columns 5-7 compare rates among individuals who used a PCH in the treatment window. The Table shows that care coordinators were almost 14 times as likely to contact ActionHealthNYC members.

Table 2.4: *PCH Electronic Health Records (EHR) Administrative Data*

	All Individuals		Low Risk Individuals		High Risk Individuals	
	Mean	Impact	Mean	Impact	Mean	Impact
Utilization, Relative to Baseline PCH Data						
Prob. PCP Visits	0.089	0.42***	0.067	**0.45*** **	0.12	**0.38*** **
		(0.021)		(0.028)		(0.031)
		476.6%		673.4%		328.2%
No. PCP Visits	4.6	0.75	3.9	-0.88	5.1	2.0
		(1.2)		(1.3)		(1.9)
		16.5%		-22.6%		40.5%
Prob. SP Visit	0.051	0.23***	0.029	0.22***	0.078	0.24***
		(0.019)		(0.025)		(0.028)
		450.8%		756.0%		301.9%
No. SP Visits	5.3	0.11	3.3	-0.50	6.3	0.39
		(5.7)		(5.7)		(7.2)
		2.0%		-15.3%		6.2%
New DX, Relative to Baseline Survey Responses						
Prob. Any DX	0.53	0.28***	0.100	**0.17*** **	1.087	**0.41*** **
		(0.050)		(0.044)		(0.099)
		52.3%		175.3%		38.1%
Prob. Hypertension DX	0.17	0.076***	–	**0.015**	0.38	**0.16*** **
		(0.020)		(0.013)		(0.044)
		45.4%		–%		40.8%
Prob. Diabetes DX	0.089	0.040**	–	0.016	0.20	0.072**
		(0.015)		(0.0098)		(0.033)
		45.5%		–%		35.6%
Prob. Asthma DX	0.045	0.0051	–	0.0027	0.10	0.0084
		(0.012)		(0.0041)		(0.027)
		11.4%		–%		8.1%
Prob. Substance Abuse DX	0.11	0.087***	0.093	0.070**	0.14	0.11***
		(0.023)		(0.030)		(0.037)
		76.8%		75.2%		78.5%
Screens, Relative to Baseline Survey Responses						
Prob. Diabetes Screen	6.9	4.3***	5.9	4.1***	8.3	4.5***
		(0.32)		(0.42)		(0.49)
		61.7%		70.0%		54.3%
Prob. Blood Pressure Screen	6.8	4.9***	5.5	5.0***	8.4	4.7***
		(0.33)		(0.43)		(0.50)
		71.8%		90.6%		55.9%

Notes: Estimates are relative to baseline, which was assessed on the baseline survey. Regressions include member fixed effects and are clustered at the couple level because couples were randomized together. Post-period data was obtained from electronic health records from participating primary care homes (PCHs). The high risk patient category includes patients with one or more of the following conditions met, based on baseline survey responses: mental illness, hypertensive, cardiovascular disease, COPD, congestive heart failure, previously had a heart attack or stroke, Atrial Fibrillation, diabetes, asthma, cancer, HIV+, Hep C, or homeless. Visits and diagnoses that occurred outside of our official sites are not captured. See Table **??** to get a sense of how many visits we miss. Diagnoses not included in this table are cancer, HIV, congestive heart failure, mental health conditions, and Atrial Fibrillation, which are rarer conditions than those listed. They do not significantly change in response to our program. Bolded estimates indicate that means are significantly different from each other at the 5% level.

Table 2.5: *PCH Administrative Data*

	All Individuals			Conditional on Using PCH in Treatment Window		
	Treated	Control	P-Value	Treated	Control	P-Value
No. of Individuals	1,265	1,139		809	242	
Screens						
Prob. ECC Contact	**0.48**	**0.0079**	p< 0.001	**0.45**	**0.033**	p< 0.001
Prob. Prescr. Aspirin	0.040	0.028	0.10	0.049	0.070	0.21
Prob. Mental Health Screen	**0.62**	**0.27**	p< 0.001	0.82	0.84	0.33
Prob. Cancer Screen	**0.30**	**0.15**	p< 0.001	**0.42**	**0.52**	0.009
Prob. Substance Abuse Screen	**0.66**	**0.32**	p< 0.001	**0.87**	**0.94**	0.004
Prob. Hep C Screen	0.0032		0.22	0.0049	0.0041	0.87
Prob. HIV Screen	**0.24**	**0.11**	p< 0.001	**0.33**	**0.22**	0.001
Prob. CVD Screen	**0.59**	**0.25**	p< 0.001	0.80	0.83	0.34
Prob. STD Screen	**0.22**	**0.10**	p< 0.001	0.33	0.37	0.22

Notes: Estimates are relative to baseline, which was assessed on the baseline survey. Regressions include member fixed effects and are clustered at the couple level because couples were randomized together. Post-period data was obtained from electronic health records from participating primary care homes (PCHs). The high risk patient category includes patients with one or more chronic condition based on baseline survey responses. Visits and diagnoses that occurred outside of our official sites are not captured. See Table **??** to get a sense of how many visits we miss. Diagnoses not included in this table are cancer, HIV, congestive heart failure, mental health conditions, substance use disorder, and Atrial Fibrillation, which are rarer conditions than those listed. They do not significantly change in response to our program. All screens are dummies for an individual having one or more screen. Bolded estimates indicate that means are significantly different from each other at the 5% level.

2.6 Discussion and Conclusion

In this study, we examine the impact of coordinating undocumented immigrants' health care on utilization and health status using a randomized control trial. We find that coordinating undocumented immigrants' health care increases self-reported access to care, the use of primary care, as well as preventive care. Effects were strongest among high risk individuals. The intervention also caused high risk individuals to decrease their use of the emergency department by 26%.

These findings, paired with novel information on the health status of undocumented immigrants, suggest that undocumented immigrants have a high degree of unmet need for health care services. As a result of poor access to primary care, our results suggest that undocumented immigrants subsequently rely on the ED for primary care services.

By providing undocumented immigrants' access to primary care services upfront, our back-of-the-envelope calculations suggest that individuals' long-run probability of death may decrease by 7%. Simultaneously, this may save policymakers money by decreasing individuals' reliance on the emergency department. This is especially relevant for cities like NYC where the public hospital system, and therefore NYC's taxpayers, internalize a large share of the cost of uncompensated care.

Chapter 3

Should There be Vertical Choice in Health Insurance Markets?[1]

[1]This chapter is joint work with Victoria Marone. We would like to thank Leemore Dafny, Igal Hendel, Gaston Illanes, and Amanda Starc for their invaluable mentorship and advice. We are also grateful to Vivek Bhattacharya, David Cutler, David Dranove, Liran Einav, Tal Gross, Matthew Leisten, Matt Notowidigdo, Chris Ody, Rob Porter, Elena Prager, Mar Reguant, Bill Rogerson, and Gabriel Ziegler for excellent advice and suggestions. We thank our discussant Sebastian Fleitas, seminar participants at the 8th Conference of the American Society of Health Economists, and the Northwestern Industrial Organization working group for helpful comments. Finally, we thank Jason Abaluck and Jon Gruber for access to the data and for their support of this research project.

Dissertation Advisor:
David Cutler

Author:
Adrienne Hope Sabety

Essays on the Institutional Design of Health Care Markets

Abstract

Choice over coverage level—"vertical choice"—is widely available in U.S. health insurance markets, but there is limited evidence of its effect on welfare. The socially efficient level of coverage for a given consumer optimally trades off the value of risk protection and the social cost from moral hazard. Providing choice does not necessarily lead consumers to select their efficient coverage level. We show that in regulated competitive health insurance markets, vertical choice should be offered only if consumers with higher willingness to pay for insurance have a higher efficient coverage level. We test for this condition empirically using administrative data from a large employer representing 45,000 households. We estimate a model of consumer demand for health insurance and healthcare utilization that incorporates heterogeneity in health, risk aversion, and moral hazard. Our estimates imply substantial heterogeneity in efficient coverage level, but we do not find that households with higher efficient coverage levels have higher willingness to pay. It is therefore optimal to offer only a single coverage level. Relative to a status quo with vertical choice, offering only the optimal single level of coverage increases welfare by $302 per household per year. This policy shift also leads to a more even distribution of health-related spending (premiums plus out-of-pocket costs) in the population, suggesting an increase in equity as well as in efficiency.

3.1 Introduction

The availability of choice over financial coverage level—which we term "vertical choice"—is widespread in U.S. health insurance markets.[2] A leading example is the metal tier plans (e.g., Bronze, Silver, Gold) offered on Affordable Care Act exchanges. In contrast, national health insurance schemes typically offer a single level of coverage. Regulation plays a central role in determining the extent of vertical choice in health insurance markets, but the literature in economics provides limited guidance to regulators on this topic. In this paper we develop a theoretical and empirical framework for evaluating the welfare effects of vertical choice.

The basic argument in favor of vertical choice is the standard argument in favor of product variety: with more product choices, consumers can more closely match with their socially efficient product by revealed preference (Dixit and Stiglitz, 1977). However, this argument relies critically on the condition that privately optimal choices align with socially optimal choices. In competitive markets in which costs are independent of private values, this alignment is standard. In markets with selection, this alignment may not be possible. Health insurance markets are classic examples of selection markets. Costs are inextricably related to private values, and asymmetric information prevents prices from reflecting marginal costs (Akerlof, 1970; Rothschild and Stiglitz, 1976). We show that even if such markets are competitive, regulated, and populated by informed consumers, whether more choices can lead to a more efficient allocation is theoretically ambiguous.

Our welfare metric derives from the seminal literature on optimal health insurance, which holds that the efficient level of coverage equates the marginal benefit of risk protection and the marginal social cost of healthcare utilization induced by insurance (Arrow, 1965; Pauly, 1968, 1974; Zeckhauser, 1970). This central tradeoff between the "value of risk

[2]Financial coverage level is determined by plan features such as deductibles and caps on out-of-pocket payments. Vertical choice is also a key point of differentiation among current federal policy proposals. The "Medicare for all" proposal (endorsed by Bernie Sanders and Elizabeth Warren) would not feature vertical choice, while the plan to introduce a public option to existing exchanges (endorsed by Joe Biden) and the "American Health Care Act" (endorsed by Donald Trump) would continue to do so.

protection" and the "social cost of moral hazard" plays out on a consumer-by-consumer basis, and the efficient level of coverage likely varies across consumers. Socially optimal regulation aims to design plan menus such that consumers self-select into their efficient level of coverage. Private incentives are such that under any plan menu, consumers with higher willingness to pay for insurance choose (weakly) higher levels of coverage. However, consumers with higher willingness to pay do not necessarily have a higher efficient level of coverage. It is precisely this statement that captures the theoretical ambiguity of whether vertical choice should be offered.

We ask whether vertical choice should be offered from the perspective of a market regulator that can offer vertically differentiated plans and set premiums.[3] The regulator's objective is to maximize allocational efficiency of consumers to plans. As is standard in national health insurance schemes and employer-sponsored health insurance, consumer premiums need not equal plan average cost. If the regulator sets premiums such that more than one plan is demanded, we say it has offered vertical choice. Using a graphical framework in the spirit of Einav, Finkelstein and Cullen (2010b), we show that the key condition determining whether vertical choice should be offered is whether consumers with higher willingness to pay have a higher efficient coverage level. The principal empirical focus of this paper is to determine whether this is likely to be true.

We begin by presenting a model of consumer demand for health insurance, building closely on the models of Cardon and Hendel (2001) and Einav, Finkelstein, Ryan, Schrimpf and Cullen (2013). [4] The model features two stages. In the first stage, consumers make a discrete choice over plans under uncertainty about their health. In the second stage, upon realizing their health, consumers make a continuous choice of healthcare utilization. We use the model to show that willingness to pay for insurance can be partitioned into two parts:

[3]By market regulator, we mean the entity that administers and operates a particular health insurance market. In employer-sponsored insurance, this is the employer; in Medicare, this is the Center for Medicare and Medicaid Services; in Norway, this is the Norwegian government. As we will discuss, the regulator can set premiums in a competitive market by strategically taxing and subsidizing plans, or can supply plans itself.

[4]The goal of the model is to capture heterogeneity across consumers in the determinants of private and social surplus generated by insurance.

one that is both privately *and* socially relevant (the value of risk protection), and one that is only privately relevant (expected reduction in out-of-pocket spending). Because a portion of private benefit is a transfer, it is not necessarily the case that higher willingness to pay implies higher social surplus. For example, a very sick but risk neutral person obtains a large private benefit from higher coverage, but generates no social benefit. The burden of her expected spending is simply shifted. If she consumes more healthcare than she values in response to higher coverage, the regulator would prefer she had lower coverage.

We estimate the model using data from the population of public-school employees in Oregon. The data contain the health insurance plan menu, plan choice, and subsequent healthcare utilization of 45,000 households between 2008 and 2013. Crucially for identification, we observe plausibly exogenous variation in the plan premiums and plan options offered to employees. This variation is driven by the fact that plan menus are set independently by each of the 187 school districts in the state, where districts select plans from a common superset determined at the state level. In addition, employees are offered several different coverage levels by the same insurer with the same provider network, providing isolated variation along our focal dimension.

Our empirical model incorporates both observed and unobserved heterogeneity along three key dimensions of household type: health status, moral hazard, and risk aversion. We use the model to recover the joint distribution of household types in the population. For each household, we then construct willingness to pay for and the social surplus generated by different levels of coverage. We construct these objects for a set of coverage levels that span the range offered on the Affordable Care Act exchanges and in our empirical setting. Each coverage level, or plan, is characterized by a deductible, a coinsurance rate, and an out-of-pocket maximum. The least generous plan we consider is a "Catastrophic" plan, with a deductible and out-of-pocket maximum of $10,000. The most generous plan we consider is full insurance.

We do not find that households with higher efficient coverage level have higher willingness to pay. Households with high willingness to pay are primarily so because of high

expected insured spending, as opposed to a high value of risk protection. While they do tend to be more risk averse, they are so likely to hit their out-of-pocket maximum that they would face little uncertainty over out-of-pocket costs under any of the plans. Households with low willingness to pay are more prone to moral hazard and less risk averse, but also face more uncertainty over out-of-pocket costs. The correlation between willingness to pay and *risk* is of central importance to our results. Variation in risk aversion and moral hazard are less important. We find that a single plan is on average the efficient coverage level across the entire distribution of willingness to pay. Optimal regulation is therefore to offer only this plan. Introducing any other plan leads to over- or under-insurance (on average) among households that would select the alternative. The optimal single plan has an actuarial value (AV) of 85 percent.[5] Households' efficient coverage levels range between 70 percent AV and full insurance. There are no households for whom the efficient level of coverage is below 70 percent AV.[6]

The first-best allocation of households to plans generates $1,796 in welfare per household per year relative to allocating all households to the Catastrophic plan. Because households with the same willingness to pay can have different efficient levels of coverage, this allocation cannot be supported unless premiums can vary by households' specific risk aversion and moral hazard types. Under optimal regulation (the single plan), 31 percent of households are not allocated to their efficient coverage level. Nevertheless, we find that the optimal single plan generates 96 percent of the social surplus of the first-best allocation. The value of risk protection is increasing in coverage level, but at a decreasing rate. The social cost of moral hazard is also increasing in coverage level, and at an increasing rate. At the optimal allocation, the magnitude of risk protection is roughly six times as large as the social cost of moral hazard. As a result, among plans near the optimal single plan, the welfare stakes of misallocation are small. Allocating all households to a 65 percent AV, 70 percent AV, 85

[5]Actuarial value measures the percent of a population's total healthcare spending that would be insured under a particular insurance plan. An actuarial value of 100 percent is full insurance. The Catastrophic plan we consider has an actuarial value of 65 percent.

[6]70 percent AV is the coverage level provided by the Silver plans on the Affordable Care Act exchanges.

percent AV, and full insurance plan respectively generates 51, 92, 96, and 91 percent of first best social surplus.

We compare outcomes under several alternative policies, including competitive pricing and full vertical choice over all plans. Under competitive pricing, vertical choice is permitted but prevailing premiums must equal plan average costs. In our population, the market unravels to the lowest level of coverage (the Catastrophic plan) due to adverse selection. Under full vertical choice, we implement subsidies that can support an allocation in which all plans are traded. Using subsidies designed to mimic the enrollment shares observed on Affordable Care Act exchanges, vertical choice generates 80 percent of first best surplus. We find that all households prefer vertical choice to the unraveled market, and that 81 percent of households prefer the optimal single plan to vertical choice. Social surplus is $302 higher per household per year under optimal regulation than under vertical choice.

Related Literature. Our two-stage model of household demand for health insurance and healthcare utilization is closely related to those of Cardon and Hendel (2001) and Einav et al. (2013). We present a generalized formulation of these models to highlight the fact that the decompositions of willingness to pay and social surplus do not depend on particular functional forms for moral hazard, plan design, or uncertainty over health outcomes. From a methodological perspective, we extend the empirical approach to modeling distributions of household health outcomes. While the healthcare utilization decision occurs at the household level, health status predictors (such as age) are measured at the individual level. We recast household health as the sum of individuals' health and operationalize our approach using an approximation to the sum of lognormal distributions. This method allows us to exploit detailed information on a household's composition of individuals while still limiting the number of parameters to estimate.[7]

Our graphical analysis is based on the widely-used framework developed by Einav et al. (2010b). We extend the framework by incorporating a "social surplus curve" that captures

[7]Given the large size of our data (45,000 households choosing among 14 plans over 5 years), limiting the number of parameters to estimate was an important consideration for computational tractability.

the social surplus generated by allocating consumers to a given plan (Figure 3.1). We also incorporate the framework into our empirical analysis by using our estimates to construct the empirical analogs. Our main findings can be read directly from our empirical social surplus curves (Figure 3.7).

The notion of equilibrium in our model is related to, but departs from standard competitive equilibria studied in health insurance markets in which the premium of each plan must equal the average cost of those who demand it (e.g., Rothschild and Stiglitz (1976), Handel, Hendel and Whinston (2015), Azevedo and Gottlieb (2017)). In our model, a regulator can set premiums arbitrarily. Removing price as an equilibrium object makes a larger set of allocations feasible. We find this desirable both because it reflects realistic regulatory powers and because it focuses attention on the important economic constraint of unobserved types.

Our framework is most closely related to that of Azevedo and Gottlieb (2017). They also use a two-stage model to describe demand for health insurance in a setting with vertically differentiated contracts and multiple dimensions of consumer heterogeneity. While their focus is on competitive equilibria with break-even pricing, their numerical simulations also consider optimal pricing. They document that under certain parameterizations of the distribution of consumer types, offering choice is optimal, while under others it is not.[8] Our paper focuses directly on why this is the case. We are (to our knowledge) the first to characterize the conditions under which it is optimal to offer vertical choice. We also bring to bear a rich empirical approach that permits flexible heterogeneity in the distribution of consumer types.[9]

Finally, our paper also closely relates to the literature on health insurance menu design. Bundorf, Levin and Mahoney (2012) investigate the socially optimal allocation of consumers to insurers in one market and find that optimal allocations cannot be achieved under uniform pricing. Our paper is similar in spirit (and in findings), but analyzes optimal

[8]Their simulated population of consumers is characterized by a lognormal distribution of types, where moments of the distribution are set to match those estimated empirically in Einav et al. (2013).

[9]Ericson and Sydnor (2017) also consider the question of whether vertical choice is welfare-improving. They focus on consumer confusion as a source of inefficiency, while we focus on a setting with informed consumers.

allocations of consumers to coverage levels. Einav, Finkelstein and Levin (2010a) discuss, and Geruso (2017) studies empirically, the idea that difficulties in optimal screening can arise when observably different consumers have the same willingness to pay for insurance; this is a central issue in our setting. In concurrent work, Ho and Lee (2019) use a closely related framework to study the choice of optimal coverage level from the perspective of an employer offering a single coverage option, with a similar focus on the tradeoff between risk protection and moral hazard. We concentrate on whether or not optimal regulation involves choice.

The paper proceeds as follows. Section 2 presents our theoretical model and derives the objects needed to determine whether vertical choice should be offered. Section 3 describes our data and provides descriptive evidence of the extent of variation it provides. Section 4 presents the empirical implementation of our model. Section 5 presents the model estimates and constructs willingness to pay and social surplus. Section 6 evaluates welfare and distributional outcomes under alternative pricing policies. Section 7 concludes.

3.2 Theoretical Framework

3.2.1 Model

We consider a model of a health insurance market where consumers are heterogeneous along multiple dimensions and the set of traded contracts is endogenous. Contracts differ along a single dimension and we take the set of potential contracts as given. We assume that the regulator cannot (or will not) vary premiums by consumer characteristics and assert that each consumer will select a single contract.[10]

We denote the set of potential contracts by $X = \{x_0, x_1, ..., x_n\}$, where x_0 is a null contract that provides no insurance. Within X, contracts are vertically differentiated only by the level

[10]The regulator may not be able to condition premiums on consumer attributes if consumers have private information (see Cardon and Hendel (2001)). It may not want to do so to prevent exposing consumers to costly reclassification risk (see Handel et al. (2015)). Otherwise, the market could be partitioned according to observable characteristics, and each submarket could be considered separately.

of insurance coverage provided. Consumers are characterized by type $\theta : \{F, \psi, \omega\}$, where F is a distribution over potential health states, $\psi \in \mathbb{R}_{++}$ is a risk aversion parameter, and ω is a parameter describing consumer preferences over healthcare utilization (capturing the degree of moral hazard). We define a population by a distribution $G(\theta)$.

Demand for Health Insurance and Healthcare Utilization. Consumers are subject to a stochastic health state l, drawn from their distribution F. After their health state is realized, consumers decide the amount $m \in \mathbb{R}_{+}$ of healthcare utilization ("spending") to consume, where m is measured in dollars. Consumers value healthcare spending m and residual income. In deciding how much healthcare to utilize, consumers trade off the associated benefit $b(m, l, \omega)$ and the out-of-pocket (OOP) cost $c(m, x)$, where both are both increasing in m. The privately optimal amount of healthcare to consume is $m^*(l, x, \omega) = \text{argmax}_m \left(b(m, l, \omega) - c(m, x)\right)$.[11] Because insurance reduces the cost of healthcare, privately optimal spending is increasing in coverage level.[12] Optimal spending implies indirect benefit $b^*(l, x, \omega)$ and indirect out-of-pocket cost $c^*(l, x, \omega)$.

We take $\omega = 0$ to mean there is no moral hazard, meaning that optimal spending does not vary over contracts: $m^*(l, x, 0) = m^*(l, x_0, \omega) \; \forall \; x$. In order to reach an expression for the social cost of moral hazard in terms of fundamentals, we decompose healthcare spending $m^*(l, x, \omega)$ into two parts: (i) "unavoidable spending" $m^*(l, x, 0)$ that would occur even absent insurance, and (ii) "moral hazard spending" $m^*(l, x, \omega) - m^*(l, x, 0)$ that is induced by insurance.[13] Moral hazard spending is not entirely wasteful. Consumer utility from

[11] We assume $m^*(l, x, \omega)$ is unique. We also note that the *socially* optimal amount of healthcare to consume is $m^{eff}(l, \omega) = \text{argmax}_m \left(b(m, l, \omega) - m\right) = m^*(l, x_0, \omega)$. This paper does not tackle allocational efficiency with respect to healthcare utilization; implicitly we assume the realized health state l is not contractible.

[12] Following convention, we refer to ω as a "moral hazard" parameter, but note that in our model it captures only price sensitivity to the out-of-pocket price of healthcare, and not a hidden action. See Section I.B of Einav et al. (2013) for a fuller discussion of this abuse of terminology in the health insurance literature.

[13] We assume m^*, b^*, and c^* are weakly increasing in ω. This is a normalization.[14]

moral hazard spending (net of associated out-of-pocket cost) is equal to

$$v(l,x,\omega) = \underbrace{b^*(l,x,\omega) - b^*(l,x,0)}_{\substack{Benefit\ from\ moral \\ hazard\ spending}} - \underbrace{\left(c^*(l,x,\omega) - c^*(l,x,0)\right)}_{\substack{OOP\ from\ moral \\ hazard\ spending}}.$$

Because lower out-of-pocket costs make consumers weakly better off, $v(l,x,\omega)$ is weakly positive. Before the health state is realized, expected utility from contract x at premium p equals

$$U(x,p,\theta) = \mathbb{E}\left[\, u_\psi(y - p + \ b^*(l,x,0) - c^*(l,x,0) + v(l,x,\omega))|\ l \sim F\right], \tag{3.1}$$

where y is initial income and u_ψ is a strictly increasing and concave Bernoulli utility function with curvature governed by ψ.

Private vs. Social Incentives. Calculations in Appendix C.1.1 show that if consumer preferences u_ψ feature constant absolute risk aversion, willingness to pay for contract x relative to the null contract x_0 can be expressed as

$$WTP(x,\theta) = \underbrace{\bar{c}(F,x_0,0) - \bar{c}(F,x,0)}_{\substack{Mean\ reduced\ OOP \\ from\ unavoidable\ spending}} + \underbrace{\bar{v}(F,x,\omega)}_{\substack{Mean\ utility\ from \\ moral\ hazard\ spending}} + \underbrace{\Psi(x,\theta)}_{\substack{Value\ of\ risk \\ protection}}, \tag{3.2}$$

where $\bar{c}(F,x,\omega)$ is the expected value of $c^*(l,x,\omega)$ with respect to l, and $\bar{v}(F,x,\omega)$ is similarly defined. Each contract represents a gamble over financial payoffs and utility from healthcare utilization.[15]

Willingness to pay is composed of three terms: mean reduced out-of-pocket cost from unavoidable spending, mean utility from moral hazard spending, and the value of risk protection.[16] The first term, mean reduced out-of-pocket from unavoidable spending, is a financial expected value that will appear as an equal and opposite cost to the insurer. It is

[15] WTP represents a certainty equivalent, equal to an expected value plus a risk premium. The role of constant absolute risk aversion is to ensure that the risk premium does not depend on the plan premium.

[16] Azevedo and Gottlieb (2017) also discuss how willingness to pay in this setting is composed of these three terms. Our formulation generalizes the decomposition in that it does not depend on particular functional forms for b, c, or F.

a transfer that is not relevant to social welfare.[17] In contrast, the second and third terms depend on consumer preferences and are relevant to social welfare. Consumers may value the ability to consume more healthcare when they have higher coverage as well as the ability to smooth consumption across health states. Our accounting of social welfare takes this into consideration.

Insurer costs are given by $k(m, x)$, where $m = k(m, x) + c(m, x)$. A reduction in out-of-pocket cost is an increase in insurer cost, so $\bar{c}(F, x_0, 0) - \bar{c}(F, x, 0) = \bar{k}(F, x, 0)$.[18] The social surplus generated by allocating a consumer to contract x (relative to allocating the same consumer to the null contract) is the difference between $WTP(x, \theta)$ and expected insured cost $\bar{k}(F, x, \omega)$:

$$SS(x, \theta) = \underbrace{\Psi(x, \theta)}_{\substack{Value\ of\ risk \\ protection}} - \underbrace{(\bar{k}(F, x, \omega) - \bar{k}(F, x, 0) - \bar{v}(F, x, \omega))}_{\substack{Social\ cost \\ of\ moral\ hazard}}. \tag{3.3}$$

Because the insurer is risk neutral, it bears no extra cost from uncertain payoffs. If there is moral hazard, the consumer's value of mean insured spending falls below the cost of providing it, generating a welfare loss from insurance.

The socially optimal contract x^{eff} for a particular type of consumer is that which optimally trades off risk protection and the social cost of moral hazard: $x^{eff}(\theta) = \text{argmax}_{x \in X} SS(x, \theta)$. Given premium vector $\boldsymbol{p} = \{p_x\}_{x \in X}$, consumers choose the privately optimal contract x^* that optimally trades off private utility and premium: $x^*(\theta, \boldsymbol{p}) = \text{argmax}_{x \in X}(WTP(x, \theta) - p_x)$.

Supply and Regulation. Contracts are supplied by a regulator, which can observe the distribution of consumer types and can set premiums. The regulator need not break even on any given contract, nor break even in aggregate.[19] It could remove a contract from the

[17] The insurer's technology is risk neutrality. It cannot pay doctors a marginal dollar more efficiently than the consumer could do.

[18] To see this, note that $\bar{c}(F, x_0, 0) = \bar{m}(F, x, 0)$. $\bar{k}(F, x, \omega)$ is the expectation of $k^*(l, x, \omega)$ with respect to the distribution of l, where $k^*(l, x, \omega) = k(m^*(l, x, \omega), x)$.

[19] We assume any aggregate deficit can be funded by taxing consumer incomes. Since we assume constant absolute risk aversion, this is not different than increasing premiums on all plans and calling it a tax.

set of contracts on offer by setting a premium of infinity. This model of supply is equivalent to a perfectly competitive insurance market with a regulator that has the power to tax and subsidize plans. Precisely such a model is formalized in Section 6 of Azevedo and Gottlieb (2017).

The regulator sets premiums in order to align privately optimal $x^*(\theta, \boldsymbol{p})$ and socially optimal $x^{eff}(\theta)$ allocations as closely as possible. Equilibrium social welfare is given by

$$W(\boldsymbol{p}) = \int SS(x^*(\theta, \boldsymbol{p}), \theta)\, dG(\theta).$$

Our question is whether, or when, the regulator's solution will involve vertical choice. That is, will the regulator wish to offer (have enrollment in) more than one contract at the optimal allocation.[20]

3.2.2 Graphical Analysis

We characterize the answer graphically for the case of a market with two potential contracts. This case conveys the basic intuition and can be depicted easily using the graphical framework introduced by Einav et al. (2010b).

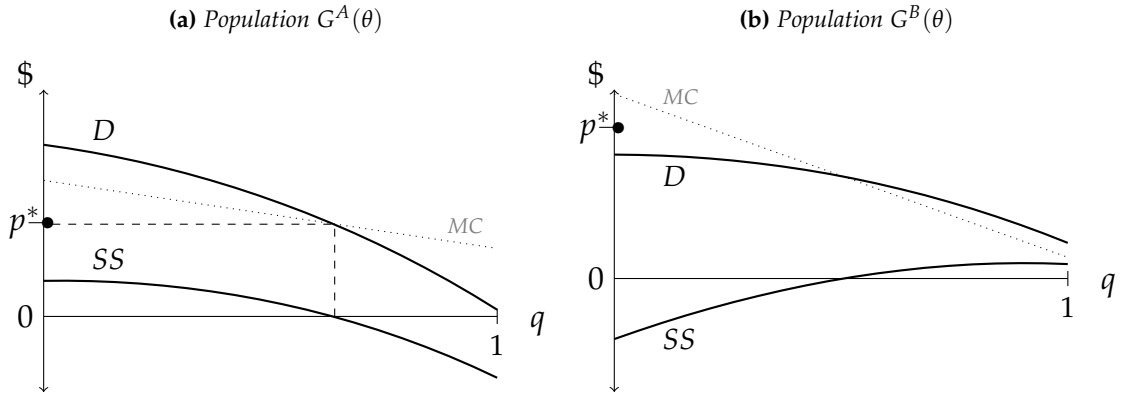
First, it is useful to recognize that moral hazard and consumer heterogeneity are necessary conditions for the regulator to wish to offer vertical choice. If there were no moral hazard, higher coverage would weakly increase social welfare for every consumer. The optimal contract for all consumers would therefore be the maximum possible coverage level. The regulator would set the premium of that contract to zero and the premiums of all other contracts sufficiently high that they are not chosen. If there were no consumer heterogeneity, all consumers would again have the same socially optimal contract, say $\tilde{x}$. The regulator would optimally set the premium of $\tilde{x}$ to be zero and the premiums of all other contracts to be sufficiently high that they are not chosen. In both examples, the optimal allocation does not involve vertical choice. In the following, we explore the more interesting (and more

[20] If the regulator sets premiums such that all consumers choose the same plan, then we say that it has not offered vertical choice. This is to avoid discussion of, for example, whether an option of a plan with a premium of infinity is in fact an option at all.

realistic) cases in which consumers do not all have the same optimal contract.[21]

We consider an example with two possible contracts, x_H and x_L, where $x_H > x_L$. Figure 3.1 depicts two possible markets, corresponding to two populations $G^A(\theta)$ and $G^B(\theta)$. If a consumer does not choose x_H, they receive x_L. Since contracts are vertically differentiated, $WTP(x_H, \theta) \geq WTP(x_L, \theta)$ for all consumers. Each panel shows the demand curve D for contract x_H, representing marginal willingness to pay for x_H relative to x_L. The vertical axis plots the marginal price $p = p_H - p_L$ at which the contracts are offered. The horizontal axis plots the fraction q of consumers that choose x_H.

Figure 3.1: *Markets Where There (a)* Should *and (b)* Should Not *be Vertical Choice*



Notes: This figure shows two health insurance markets where there are two contracts available: x_H and x_L, where $x_H > x_L$. Each panel shows the demand curve D, the marginal cost curve MC, and the social surplus curve SS for contract x_H relative to contract x_L. The left panel depicts an example where the regulator optimally offers vertical choice, and there is enrollment in both contracts. The right panel depicts an example where the regulator optimally does not offer vertical choice, and all consumers choose x_L.

Each panel also shows the marginal cost curve MC and the marginal social surplus curve SS. The marginal cost curve measures the expected marginal cost of insuring consumers under x_H relative to x_L. Because consumers with the same willingness to pay can have different costs, MC represents the average marginal cost among all consumers at a particular

[21] Requiring that all consumers do not have the same optimal contract is a stronger condition than requiring the presence of both moral hazard and consumer heterogeneity. Heterogeneity in optimal contracts is necessary for the regulator to wish to offer vertical choice. As in the examples above, if consumers are heterogeneous but still have the same optimal contract, the regulator will offer only that one.

point on the horizontal axis (a particular willingness to pay). The social social surplus curve SS plots the vertical difference between D and MC. A particular point on the social surplus curve represents the average marginal social surplus $SS(x_H, \theta) - SS(x_L, \theta)$ among all consumers at that point on the horizontal axis.

While D and MC must be weakly positive, the presence of moral hazard means that SS need not be; it is possible for a consumer to be over-insured. Moreover, our precondition that all consumers do not have the same optimal contract guarantees that in both populations, marginal social surplus will be positive for some consumers and negative for some consumers. Given that SS represents the average over consumers at each value of D, this condition does not guarantee that SS will itself cross zero. If SS does not cross zero, a single plan is on average optimal at every level of willingness to pay, and the regulator will offer only that plan.[22] Since it is necessary for SS to cross zero for vertical choice to be optimal, we focus both graphical examples on cases where that occurs.

The key difference between the two populations is whether consumers with high or low willingness to pay have a higher efficient level of coverage. In Figure 3.1a, marginal social surplus is increasing in marginal willingness to pay. The optimal marginal premium p^* can sort consumers with on-average positive SS into x_H, and on-average negative SS into x_L. Because private and social incentives are aligned, it is possible to get consumers to self-select efficiently. In Figure 3.1b, marginal social surplus is decreasing in consumer willingness to pay, and efficient screening is no longer possible.

In population $G^B(\theta)$, any marginal premium between the minimum and the maximum value of D will result in some avoidable amount of "backwards sorting." Consequently, any allocation with enrollment in both plans will be dominated by an allocation with enrollment in only one plan. No sorting dominates backwards sorting because it is always possible to prevent "one side" of the backwards sort by declaring no sorting. To see this, consider the (worst possible) allocation $\tilde{q}$ at the point where SS intersects zero; a slightly higher

[22]For example if SS lies everywhere above zero, the regulator will optimally offer only x_H. Note that this result corresponds to what we find empirically (cf. Figure 3.7).

allocation q' strictly dominates, as more consumers with positive marginal social surplus now enroll in contract x_H. The same logic applies to the left of $\tilde{q}$. The only allocations that cannot easily be ruled out as suboptimal are the endpoints, at which all consumers enroll in the same contract. In the example shown, the integral of SS is negative, meaning that the population would on average be over-insured in contract x_H. The optimal marginal premium p^* is therefore anything high enough to induce all consumers to choose contract x_L.

Considering all cases, if the social surplus curve SS crosses zero at most once, vertical choice should be offered if and only if SS crosses from above. More generally, the key characteristic of a population that determines whether vertical choice should be offered is whether consumers with higher willingness to pay have a higher efficient coverage level. This condition itself is complex, and both theoretically and by our own metrics of common sense, ambiguous. If healthy consumers change their behavior more in response to insurance, as is suggested by findings in Brot-Goldberg, Chandra, Handel and Kolstad (2017), this would tend to positively align willingness to pay and efficient coverage level. If healthy consumers are more risk averse, as is suggested by findings in Finkelstein and McGarry (2006), this would tend to negatively align willingness to pay and efficient coverage level.

There is a question of what characteristics drive variation in willingness to pay, and in turn how those characteristics are correlated with the efficient level of coverage. The net result depends on the joint distribution of expected health spending, uncertainty in health spending, risk aversion and moral hazard in the population. Moreover, it depends on how these primitives map into marginal willingness to pay and marginal insurer cost across nonlinear insurance contracts, as are common in U.S. health insurance markets and present in the empirical setting we study. Ultimately, whether high willingness to pay consumers should have higher coverage than low willingness to pay consumers is an open empirical question.

3.3 Empirical Setting

In this section, we describe our empirical setting. Section 3.3.1 describes the data. Section 3.3.2 presents descriptive evidence of the variation in our data, discusses our primary identifying assumption, and provides reduced form evidence of moral hazard.

3.3.1 Data

Our data are derived from the employer-sponsored health insurance market for public school employees in Oregon between 2008 and 2013. The market is operated by the Oregon Educators Benefit Board (OEBB), which manages benefits for the employees of Oregon's 187 school districts. Each year, OEBB negotiates with insurers and creates a state-level "master list" of plans that school districts can offer to their employees. Each plan has an associated full premium. During our time period, OEBB contracted with three insurers, each of which offered a selection of plans. School districts then independently select a subset of plans from the state-level menu and set their "employer contribution" to plan premiums, creating variation across school districts in the subsidized premiums and set of plans available to employees. Between 2008 and 2010, school districts could offer at most four plans; after 2010, there was no restriction on the number of plans a district could offer, but many still offered only a subset.

The data contain the menu of plan options available to each employee, realized plan choices, plan characteristics, and medical and pharmaceutical claims data for all insured individuals. We observe detailed demographic information about employees and their families, including age, gender, zip code, health risk score, family type, and employee's occupation type.[23,25] An employee's plan menu consists of the plan choice set and plan

[23]Individual risk scores are calculated based on prior-year medical diagnoses and demographics using Johns Hopkins ACG Case-Mix software. This software uses the diagnostic information contained in past claims data as well as demographic information to predict future healthcare spending. [24] See, for example, Brot-Goldberg et al. (2017), Carlin and Town (2008), or Handel and Kolstad (2015) for a more in-depth explanation of the software and examples of its use in economic research.

[25]Possible employee occupation types are licensed administrator, non-licensed administrator, classified, community college non-instructional, community college faculty, confidential, licensed, substitute, or superintendent.

prices. Prices consist of the subsidized premium, potential contributions to a Health Savings Account (HSA) or a Health Reimbursement Arrangement (HRA), and potential contributions towards a vision or dental insurance plan.[26]

The decentralized determination of plan menus provides a plausibly exogenous source of variation in both prices and choice sets. While all the plan menus we observe are quite generous in that the plans are highly subsidized, there is substantial variation across districts in the range of coverage levels offered and in the exact nature of the subsidies.[28] Moreover, school districts can vary plan menus at the family type and employee type level, resulting in variation both within and across school districts. These benefits decisions are made by school district employee and administrator committees, and subsidy designs are influenced by bargaining agreements with the local teachers union. Between 2008 and 2013, we observe 13,661 unique combinations of year, school district, family type, and employee type, resulting in 7,835 unique plan menus.

Household Characteristics. We restrict our sample to households where the oldest member is not older than 65, the employee is not retired, and for whom all members are enrolled in the same plan for the entire year. Further, because we require one prior year of claims data in order to estimate an individual's prospective risk score, we begin our sample in 2009, and require households to have one year of data prior to inclusion. Our sample consists of 44,562 unique households, representing 117,949 unique individuals between 2009 and

---

Within each category, an employee can be either full-time or part-time. Possible family types are employee only; employee and spouse; employee and child(ren); and employee, spouse, and child(ren).

[26]Decisions about HSA/HRA and vision/dental contributions are also made independently by school districts. An HRA is a notional account that employers can use to reimburse employees' uninsured medical expenses on a pre-tax basis; balances typically expire at the end of the year or when the employee leaves the employer. An HSA is a financial account maintained by an external broker to which employers or employees can make pre-tax contributions. [27] The data on employer premium contributions and savings account contributions were hand-collected via surveys of each school district. Additional details about the data collection process can be found in Abaluck and Gruber (2016).

[28]The majority of school districts used either a fixed dollar contribution or a percentage contribution, but the levels of the contribution varied widely. Other districts used a fixed employee contribution. In addition, the districts' policies for how "excess" contributions were treated varied; in some cases, contribution amounts in excess of the full plan premium could be "banked" by the employee in a HSA or HRA, or else contributed towards the purchase of a vision or dental insurance plan, either in full, in part, or not at all.

2013.[29]

Table 3.1 provides annual summary statistics on our panel of households. Across all years, the age of the average employee is 47.4, while the age of the average enrollee (employees and their families) is 39.8. Enrollees are 54 percent female, and 72 percent of households are "families" (purchased health insurance to cover more than the employee alone). Households have on average 2.54 enrollees.

Table 3.1: *Household Summary Statistics*

Sample demographics	2009	2010	2011	2012	2013
Number of households	31,074	29,538	29,279	27,897	24,283
Number of enrollees	78,932	75,129	75,601	72,311	63,264
Enrollee age, mean (med.)	39.7	39.8	39.8	40.1	40.0
	(38.0)	(38.0)	(37.7)	(38.0)	(37.8)
Enrollee percent female	0.54	0.54	0.54	0.54	0.54
Premiums					
Employee premium ($), mean (med.)	885	1,023	523	1,079	905
	(0)	(0)	(0)	(0)	(0)
Full premium ($), mean (med.)	11,170	11,785	10,433	12,253	12,000
	(11,665)	(11,801)	(11,021)	(12,278)	(12,362)
Household health spending					
Total spending ($), mean (med.)	10,563	10,405	10,911	10,984	10,967
	(4,753)	(4,589)	(4,595)	(4,569)	(4,559)
OOP spending ($), mean (med.)	1,152	1,634	1,884	1,897	1,998
	(743)	(1,089)	(1,306)	(1,292)	(1,234)
Plan switches (percent)					
Forced to switch plan	0.06	0.34	0.12	0.05	0.46
insurer	0.01	0.02	0.02	0.02	0.00
Unforced, switched plan	0.13	0.23	0.22	0.22	0.04
insurer	0.06	0.05	0.03	0.01	0.02
Household structure (percent)					
Individual	0.27	0.28	0.28	0.28	0.28
Family	0.73	0.72	0.72	0.72	0.72

Notes: Enrollees are employees plus their family members. Statistics about premiums are for households' chosen plans, as opposed to for all possible plans. Sample medians are shown in parentheses.

Employees received large subsidies towards the purchase of health insurance. The average household paid only $880 per year for their plan; the median household paid nothing. Meanwhile, the average full premium paid to insurers was $11,500, meaning the average household received an employer contribution of $10,620. Households had average

[29]Table C.1 provides additional details on the construction of this sample.

out-of-pocket spending of $1,694, and households plus insurers had average total spending of $10,754.

Households were highly likely to remain in the same plan and with the same insurer that they chose last year, when possible. OEBB can adjust the master list of plans available, and school districts can adjust choice sets over time. Such adjustments forced 19.6 percent of household-years to switch plans and 1.4 percent to switch insurers. Among household-years where the incumbent plan/insurer *was* available, 17.2 percent voluntarily switched plans, and 3.4 percent voluntarily switched insurers. This variation is particularly important in our empirical model in identifying "inertia" associated with switching plans or insurers.

We divide the state into a small number of regions because in our empirical model we allow preferences for each insurer to vary by region. We use three regions based on groups of adjacent Hospital Referral Regions (HRRs): the Portland and Salem HRRs in northwest Oregon (containing 64 percent of households), the Eugene and Medford HRRs in southwest Oregon (containing 26 percent of households), and the Bend, Spokane, and Boise HRRs in eastern Oregon (containing 10 percent of households).[30]

Plan Characteristics. During our sample period, OEBB contracted with three insurers: Kaiser, Providence, and Moda. Kaiser offers HMO plans that require enrollees to use only Kaiser healthcare providers and obtain referrals for specialist care. Moda and Providence offer PPO plans with broad provider networks. Kaiser and Providence each offered between two and three plans per year at high coverage levels. Moda offered between seven and nine plans per year, with wide variation in coverage level across plans. Within each insurer, plans were differentiated only by coverage level.

Table 3.2 summarizes the master list of plans made available by OEBB in 2009. The insurer premium reflects the per-employee premium negotiated between OEBB and the insurer. This full premium varies formulaically by family type; the premium shown is for an employee plus spouse. Plan cost sharing features vary by whether the household is an

[30]As HRRs do not respect state boundaries, some HRRs in our regions have names of cities outside Oregon, but nonetheless contain parts of Oregon. For more information as well as HRR maps, see http://www.dartmouthatlas.org/data/region.

individual (the employee alone) or a family (anything else). The deductible and out-of-pocket maximum (OOP Max.) shown are for in-network services for a family household.

Table 3.2: *Plan Characteristics, 2009*

Plan	AV	Insurer Premium ($)	Deductible ($)	OOP Max. ($)	Market Share
Kaiser - 1	0.97	11,869	0	1,200	0.07
Kaiser - 2	0.96	11,342	0	2,000	0.11
Kaiser - 3	0.95	10,995	0	3,000	0.00
Moda - 1	0.92	13,340	300	500	0.27
Moda - 2	0.89	12,808	300	1,000	0.05
Moda - 3	0.88	12,088	600	1,000	0.11
Moda - 4	0.86	11,578	900	1,500	0.10
Moda - 5	0.82	10,723	1,500	2,000	0.13
Moda - 6	0.78	9,691	3,000	3,000	0.04
Moda - 7	0.68	7,401	3,000	10,000	0.01
Providence - 1	0.96	14,359	900	1,200	0.07
Providence - 2	0.95	14,009	900	2,000	0.02
Providence - 3	0.94	13,779	900	3,000	0.01

Notes: Actuarial value (AV) is calculated as the ratio of average insured spending to average total spending among all households, using counterfactual calculations of insured spending for households that did not choose a certain plan. Insurer premium reflects the premium negotiated between OEBB and the insurer. The deductible and out-of-pocket maximum shown are for in-network services for a family household.

One way to summarize and compare plan coverage levels is using actuarial value (AV), which reflects the share of total population spending that would be insured under a given plan. Less generous plans correspond to those with a lower actuarial value. To calculate actuarial value, we simulate the out-of-pocket spending that *all* households would have had in every potential plan, and then compute average insured spending divided by average total spending across all households for each plan.[31] In this way, the measure is not affected by selection or moral hazard effects.

The plan offerings in later years look qualitatively similar to those in 2009.[32] The notable exception is that Providence was no longer available in 2012 and 2013. Moda maintained a

[31]We calculate counterfactual out-of-pocket spending using the "claims calculator" developed for this setting by Abaluck and Gruber (2016).

[32]Corresponding tables for the plans offered between 2010 and 2013 are available in Table C.2.

roughly 75 percent market share throughout 2009 to 2013; Kaiser and Providence initially split the remaining share, but Kaiser steadily gained share thereafter. For the purposes of our empirical model, we estimate cost-sharing features that best fit the relationship between out-of-pocket spending and total spending observed in the claims data; this procedure is described in Appendix C.1.2.

3.3.2 Descriptive Evidence

This section describes the variation in our data and estimates moral hazard in our setting. These estimates provide a moral hazard elasticity that is directly comparable to others in the literature. They also provide suggestive evidence of heterogeneity in treatment intensity, which is an important aspect of our structural model. While this section is essential for evaluating our identifying assumptions, we note that it is not necessary for understanding our structural model or subsequent analysis, which proceed in Section 3.4.

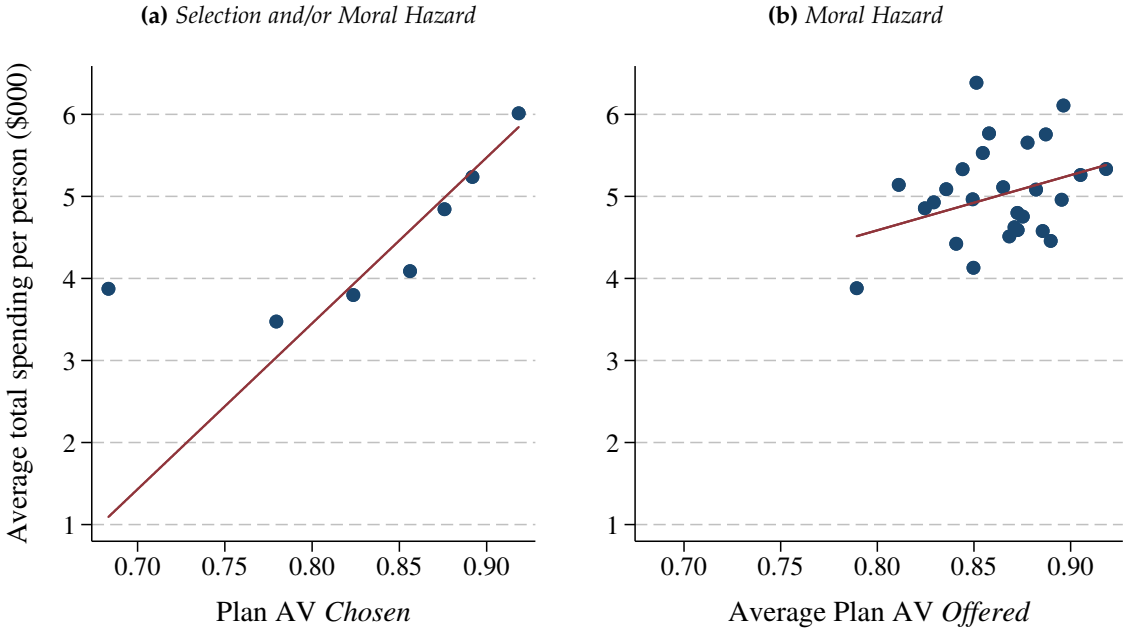
While our primary sample consists of data from 2009–2013, we conduct our descriptive analysis using only data from 2008.[33] The OEBB marketplace began operating in 2008, so in that year, all employees were choosing from among this set of plans for the first time. This "active choice" year permits us to look cleanly at how plan choices and realized healthcare spending depended on plan menus without also having to account for how prior year plan menus affected current year plan choices. While our structural model will capture these dynamics, we feel they are better avoided at this stage.

Variation in Coverage Level and Spending. We first graphically examine the extent of selection and/or moral hazard in the data. Figure 3.2 shows the relationship between healthcare spending and plan actuarial value among the set of households that chose Moda in 2008. We limit our focus to Moda here because we would like to hold the insurer fixed, and there is little variation in coverage level among the plans offered by Kaiser and Providence. The left panel of Figure 3.2 groups households by their *chosen* plan and plots

[33]Cost-sharing features of 2008 plans are presented in Table C.2; they are very similar to the plans offered in 2009. We apply all the same sample construction criteria to our 2008 sample except that the households be present for one prior year. Summary statistics on the 2008 sample can be found in Table C.4.

average spending among households in each plan. There is one observation for each of the seven Moda plans. Households enrolled in more generous plans spend more on average than households enrolled in less generous plans. The lines of best fit in each panel are weighted by the number of households represented.

Figure 3.2: *Average Spending by Coverage Level* Chosen *and* Offered



Notes: This figure shows the relationship between average total spending per person and plan actuarial value among households that selected Moda in 2008. In the left panel, each dot represents a plan. In the right panel, each dot represents a plan menu. Lines of best fit are weighted by the number of households represented.

The right panel groups households by their plan menu, and plots the average actuarial value the households were *offered* against average spending. There is one observation for each unique plan menu. Households that were offered higher coverage had higher spending on average, suggesting that coverage level may have causally affected spending. While suggestive, this graphical analysis raises some important concerns. First and foremost, we must establish that plan menu generosity is not correlated with other factors that determine healthcare spending. In addition, the 'average plan AV offered' may not be a good measure of the coverage level likely to be chosen from a given plan menu. Plan prices vary as well and households also consider plans offered by Kaiser and Providence. We first address the

exogeneity of plan menu generosity and then address these operational issues using an instrumental variables analysis.

Identifying Assumption. Our aim is to recover the causal effect of a household's chosen insurance plan on its total healthcare spending. As in much of this literature, our primary challenge is to disentangle the effects of moral hazard and adverse selection.[34] We address this challenge using choice set variation. We estimate how plan menus—choice sets and prices—affect plan choices, and in turn how plan choices affect total healthcare spending, as described by equations (3.4) and (3.5):

$$plan_k = f(\mathbf{menu}_d, \mathbf{X}_k, \epsilon_k), \tag{3.4}$$

$$y_k = g(plan_k, \mathbf{X}_k, \epsilon_k). \tag{3.5}$$

Here, $plan_k$ represents the plan chosen by household k, $\mathbf{menu}_d$ represents the plan menu available to the school district-family type-employee type combination d (to which household k belongs), $\mathbf{X}_k$ are observable household characteristics, ϵ_k are unobservable household characteristics, and y_k is total healthcare spending. Because household characteristics appear in both equations, the challenge in estimating the effect of $plan_k$ on y_k is that a household's chosen plan is correlated with its unobservable characteristics ϵ_k.

Our identifying assumption is that plan menus are independent of household unobservables ϵ_k conditional on household observables $\mathbf{X}_k$. The most important threat to identification in this paper is that school districts chose plan menu generosity in response to unobservable information about employees that would also drive healthcare spending. Plan choice sets and employer contributions are determined at the school district level by a benefits committee consisting of district administrators and union representatives. Our understanding is that there is little "public input" from employees, who are generally satisfied with their (on average highly generous) offerings. While we cannot observe it, we understand that some variation in benefit generosity is offset by compensating variation in wages. Given the detailed health information provided by claims data, nothing about our understanding of

[34] See Einav and Finkelstein (2018) for a recent review of the empirical literature on moral hazard.

this process leads us to believe that plan menus are endogenous to unobservable employee health.

That said, we investigate by attempting to explain plan menu generosity with observable household characteristics. We argue that if plan menus were not responding to *observable* information about household health, it is unlikely that they were responding to *unobservable* information. We find this argument all the more compelling because we almost certainty have better observable information on household health than did school districts when they made plan menu decisions. We find that conditional on family type, there is no correlation between plan menu generosity and household risk score (see Table C.6).[35] Appendix C.1.3 replicates this analysis for 2009–2013, to the same effect. It also presents additional regressions testing for what *does* explain variation in plan menus. We find that, among other things, plan menu generosity is higher for certain union affiliations, lower for substitute teachers and part-time employees, decreasing in district average house price index, and decreasing in the percent of Republicans in a school district.

Estimates of Moral Hazard. We parameterize $plan_k$ to be an indicator variable for the identity of the insurer and a continuous variable for the actuarial value. We then parameterize equation (3.5) according to

$$\log(y_k) = \delta_f \mathbf{1}_{f(k)=f} + \gamma \log(1 - AV_{j(k)}) \mathbf{1}_{f(k)=Moda} + \mathbf{fi}\mathbf{X}_k + \xi_k, \tag{3.6}$$

where $\mathbf{1}_{f(k)=f}$ is an indicator for the insurer chosen by household k and $AV_{j(k)}$ is the actuarial value of the plan chosen by household k. The parameter δ_f represents insurer-specific treatment effects on total spending.[36] Our parameter of interest is γ, which represents the responsiveness of total spending to plan generosity, holding the insurer fixed (at Moda). We follow the literature in formulating the model such that γ represents the elasticity of total

[35] We calculate household risk score as the average risk score among individuals in that household. As we do not have data before 2008, the 2008 regression uses risk scores calculated using 2008 claims data.

[36] These may arise due to "supply side" effects arising from differences in provider prices, provider networks, care management practices, or due to "demand side" effects from differences in average plan generosity.

spending with respect to the average out-of-pocket price per dollar of total spending.[37]

Our aim is to estimate equation (3.6) using two-stage least squares, instrumenting for the chosen insurer ($\mathbf{1}_{f(k)=f}$) and actuarial value ($AV_{j(k)}$) using $\mathbf{menu}_{d(k)}$. But $\mathbf{menu}_{d(k)}$ is complex. Plan menus contain multiple plans, and plans vary by their coverage level, the identity of their insurer, their employee premium, and their potential HSA/HRA and vision/dental contribution. We transform these multidimensional options into instruments (predicted values of $\mathbf{1}_{f(k)=f}$ and $AV_{j(k)}$) using a conditional logit model. The logit specification allows us to predict the probability that a given household would choose a given plan when presented with plan menu $\mathbf{menu}_d$ *as if* the household had been acting like the average household in the data. Variation in the resulting predicted choice probabilities is driven only by variation in plan menus, and not by household characteristics.

We estimate the following model:

$$plan_k = \underset{j \in \mathcal{J}_d}{\text{argmax}}(\alpha p_{jd} + \alpha^{VD} p_{jd}^{VD} + \alpha^{HA} p_{jd}^{HA} + \nu_j + \epsilon_{jk}), \tag{3.7}$$

where $\mathcal{J}_d$ is the set of plans available in plan menu d. Plan prices are given by the employee premium p_{jd}, the vision/dental subsidy p_{jd}^{VD}, and the HSA/HRA contribution p_{jd}^{HA}. Plan characteristics are captured nonparametrically by plan fixed effects ν_j. All household-specific determinants of plan choice are contained in the error term ϵ_{jk}, which is assumed to have a Type-1 extreme value distribution. The estimated parameters of equation (3.7) are presented in the first column of Table C.5. As expected, households dislike premiums, like HSA/HRA and vision/dental subsidies, and prefer higher coverage plans to lower coverage plans.

We use the choice probabilities predicted by the logit model to construct our instruments, denoting the predicted probability that a household presented with plan menu $\mathbf{menu}_d$ would choose plan j as ρ_{jd}.[38] Our instruments are the probability a household would choose a given insurer and the expected actuarial value of a household's plan choice conditional on

[37] To accommodate the fact that two percent of households have zero spending, we add one to total spending.

[38] Formally: $\rho_{jd} = \frac{\exp(U_{jd})}{\sum_{g \in \mathcal{J}_d} \exp(U_{gd})}$, where $U_{jd} = \alpha p_{jd} + \alpha^{VD} p_{jd}^{VD} + \alpha^{HA} p_{jd}^{HA} + \nu^j$.

insurer, respectively given by:

$$\rho_{fd} = \sum_{j \in \mathcal{J}_d^f} \rho_{jd},$$
$$\widehat{AV}_{fd} = \sum_{j \in \mathcal{J}_d^f} (\frac{\rho_{jd}}{\rho_{fd}}) AV_j, \tag{3.8}$$

where $\mathcal{J}_d^f$ is the set of plans in $\mathbf{menu}_d$ offered by insurer f.

Table 3.3 reports the two-stage least squares estimates of equation (3.6). We instrument for $\mathbf{1}_{f(k)=f}$ using using ρ_{fd} and for $\log(1 - AV_{j(k)})\mathbf{1}_{f(k)=Moda}$ using $\log(1 - \widehat{AV}_{d,Moda})\rho_{d,Moda}$. We report only the coefficient of interest (γ), but all specifications also contain insurer fixed effects, as well as controls for household risk score and family structure. The first column presents the model estimated without instruments, and the second column presents the model estimated using instrumental variables. Comparing the coefficients in columns 1 and 2, moral hazard explains 46 percent of the observed relationship between plan generosity and total spending. Our overall estimate of the elasticity of demand for healthcare spending in the population is -0.27, which is broadly similar to the benchmark of -0.2 estimated by the RAND experiment (Manning, Newhouse, Duan, Keeler and Leibowitz, 1987; Newhouse, 1993).

Columns 3 and 4 introduce heterogeneity in γ by household health. For each household type (individual or family), we classify households into quartiles based on household risk score, where Q_n denotes the quartile of risk (Q_4 is highest risk). We construct separate instruments for each of the eight household types by estimating the logit model only among that subsample of households.[39] We find noisy but large differences in γ across household risk quartiles and between individual and family households.[40]

Variation in γ could reflect either heterogeneity in the intensity of treatment across groups (extent of exposure to varying marginal prices of healthcare across plans), or heterogeneity

[39]The estimates of equation (3.7) for each subsample are presented in Table C.8.

[40]We can reject the hypothesis that the four coefficients are equal at the 10 percent level for families, but not for individuals.

Table 3.3: *Estimates of Moral Hazard*

	OLS *All*	IV *All*	IV *Individuals*	IV *Families*
	(1)	(2)	(3)	(4)
$\log(1 - AV_{j(k)})\mathbf{1}_{f(k)=Moda}$	-0.580	-0.269		
	(0.053)***	(0.084)***		
$\log(1 - AV_{j(k)})\mathbf{1}_{f(k)=Moda} \times Q_1$			-0.220	-0.415
			(0.290)	(0.131)***
$\log(1 - AV_{j(k)})\mathbf{1}_{f(k)=Moda} \times Q_2$			-0.410	-0.235
			(0.189)**	(0.088)***
$\log(1 - AV_{j(k)})\mathbf{1}_{f(k)=Moda} \times Q_3$			-0.253	-0.218
			(0.136)*	(0.090)**
$\log(1 - AV_{j(k)})\mathbf{1}_{f(k)=Moda} \times Q_4$			-0.017	0.074
			(0.346)	(0.145)
R^2	0.19	0.19	0.44	0.37
Observations	35,146	35,146	8,962	26,184

Notes: This table shows the OLS and IV estimates of equation (3.6), describing the relationship between household total spending and plan generosity. The unit of observation is a household, and the dependent variable is log of 1 + total spending. In columns 3 and 4, coefficients can vary by household risk quartile Q_n. Columns 1 and 2 are estimated on all households, while columns 3 and 4 are estimated only on individual or family households, respectively. All specifications also include insurer fixed effects and controls for household risk score and family structure. Standard errors (in parentheses) are clustered by household plan menu, of which there are 533 among individual households and 1,750 among family households. * p<0.10, ** p<0.05, *** p<.01.

in treatment effect across groups (different responsiveness to varying marginal prices of healthcare across plans), or both. While this analysis cannot distinguish between these two effects, we find suggestive evidence that this heterogeneity in some part reflects differential treatment intensity. Appendix C.1.3 presents an analysis comparing realized spending outcomes of households in different risk quartiles with the variation in plan cost-sharing features that gives rise to different (end of year) marginal out-of-pocket prices. We find that the household types for which we estimate higher γ are also more likely to be exposed to varying marginal out-of-pocket costs. Separating variation in treatment intensity from variation in treatment effect is an important advantage of our structural model.

3.4 Empirical Model

3.4.1 Parameterization

We parameterize household utility and the distribution of health states, allowing us to represent our theoretical model fully in terms of data and parameters to be estimated. We extend the theoretical model to account for the fact that in our empirical setting, there are multiple insurers, consumers are households made up of individuals, consumers may value a dollar of premiums and a dollar of out-of-pocket spending differently, and consumers make repeated plan choices over time.

Household Utility. Following Cardon and Hendel (2001) and Einav et al. (2013), we parameterize utility from healthcare spending to be quadratic in its distance above the health state. Household k's valuation of spending level m given health state realization l is given by

$$b(m, l, \omega_k) = (m - l) - \frac{1}{2\omega_k}(m - l)^2, \tag{3.9}$$

where ω_k governs the curvature of the benefit of additional spending and ultimately the degree to which optimal utilization will vary across coverage levels. Given an (increasing and concave) out-of-pocket cost function $c_{jt}(m)$ for plan j in year t, optimal total healthcare spending is given by $m^*_{jt}(l, \omega_k) = \text{argmax}_m \left(b(m, l, \omega_k) - c_{jt}(m)\right)$.[41] Solving yields $m^*_{jt}(l, \omega_k) = \omega_k(1 - c'_{jt}(m^*_{jt})) + l$.

This parameterization of household utilization choice is attractive because it produces reasonable predicted behavior under nonlinear insurance contracts and it is tractable enough to be used inside an optimization routine.[42] Additionally, ω_k can be usefully interpreted as the incremental spending induced when moving a household from no insurance (when

[41] The out-of-pocket cost function $c_{jt}(m)$ is indexed by t because cost-sharing parameters vary within a plan across years. Note that $c_{jt}(m)$ in fact also varies by household type (individual versus family), but we omit an additional index to save on notation.

[42] The model predicts that if a consumer realizes a health state just under the plan deductible, she will take advantage of the proximity to cheaper healthcare and consume a bit more (putting her into the coinsurance region). Likewise if she realizes a health state just under the out-of-pocket maximum. Figure C.2 provides a depiction of optimal spending behavior predicted by this model.

marginal out-of-pocket cost is one and $m^* = l$) to full insurance (when marginal out-of-pocket cost is zero and $m^* = \omega + l$). Substituting for m^*, we denote the benefit of optimal utilization as $b^*_{jt}(l, \omega_k)$ and the associated out-of-pocket cost as $c^*_{jt}(l, \omega_k)$. Households face uncertainty about payoffs through uncertainty in $b^*_{jt}(l, \omega_k) - c^*_{jt}(l, \omega_k)$.[43]

We further assume that households have constant absolute risk aversion (CARA) preferences. Facing uncertainty about their healthcare needs, household k in year t derives the following expected utility from plan choice j:

$$U_{kjt} = \int_0^\infty -\exp\left(-\psi_k x_{kjt}(l)\right) dF_{kft}(l), \tag{3.10}$$

where ψ_k is the coefficient of absolute risk aversion, x_{kjt} is the payoff associated with realization of health state l, and F_{kft} is the distribution of health states. Health state distributions can vary by insurer $f(j)$ in order to capture differences in provider prices across insurers (discussed further below).

The payoff of health state realization l when enrolled in plan j is given by

$$x_{kjt}(l) = -p_{kjt} + \alpha^{OOP}\left(b^*_{jt}(l, \omega_k) - c^*_{jt}(l, \omega_k)\right) + \delta^{f(j)}_{kj} + \mathbf{fl}^{inertia}_{kjt} + \mathbf{fi}\mathbf{X}_{kjt} + \sigma_\epsilon \epsilon_{kjt}, \tag{3.11}$$

where p_{kjt} is the household's plan premium (net of the employer contribution), $b^*_{jt}(l, \omega_k) - c^*_{jt}(l, \omega_k)$ is the net benefit of the optimal utilization choice measured in units of out-of-pocket dollars, $\delta^{f(j)}_{kj}$ are insurer fixed effects that control for brand and other insurer characteristics, $\mathbf{fl}^{inertia}_{kjt}$ are a set of fixed effects for both the plan and the insurer a household was enrolled in the previous year, and $\mathbf{X}_{kjt}$ is a set of additional covariates that can affect household utility.[44] The payoff x_{kjt} is measured in units of premium dollars. Out-of-pocket costs can

[43] Under our parameterization, $b^*_{jt}(l, \omega_k) = \frac{\omega_k}{2}(1 - c'_{jt}(m^*_{jt})^2)$. Because both b^*_{jt} and c^*_{jt} are increasing in ω, a larger ω will contribute to a less risky distribution of payoffs. All else equal, this would work to align willingness to pay and efficient coverage level. An important motivation for the inclusion of unobservable heterogeneity in risk aversion is to allow it to vary flexibly with respect to the amount of moral hazard.

[44] These are: HRA or HSA contributions HA_{kjt}, vision and dental plan contributions VD_{kjt}, and a fixed effect $\nu^{NarrowNet}_{jt}$ for the plan Moda gave a limited provider network in 2011 and 2012. The associated parameters for health account and vision/dental contributions are α^{HA} and α^{VD}, respectively.

be valued differently than premiums through parameter α^{OOP}.[45] Finally, ϵ_{kjt} represents a household-plan-year specific idiosyncratic preference shock, with magnitude σ_ϵ to be estimated. We assume that the shocks are independently and identically distributed Type 1 Extreme Value. In each year, households choose the plan j^*_{kt} that maximizes expected utility from among the set of plans $\mathcal{J}_{kt}$ available to them:

$$j^*_{kt} = \underset{j \in \mathcal{J}_{kt}}{\text{argmax}}\ U_{kjt}.$$

Distribution of Health States. We parameterize the distribution F_{kft} under the assumption that individuals face lognormal distributions of health states, and households face the sum of draws from individuals' distributions. We estimate the parameters of individuals' health state distributions, allowing parameters to vary with individual characteristics. We represent a household's distribution using a lognormal that approximates the sum of draws from independent lognormals.[46] This novel method of modeling the distribution of health states allows us to capture and exploit the large amount of heterogeneity in household composition that exists in our data. Importantly, it also allows us to closely fit observed spending distributions using a smaller number of parameters than would be required if covariates were measured at the household level.

An individual i faces uncertain health state $\tilde{l}_{it}$ that has a shifted lognormal distribution with parameters μ_{it} and σ_{it} and support $(-\kappa_{it}, \infty)$:

$$\log(\tilde{l}_{it} + \kappa_{it}) \sim N(\mu_{it}, \sigma^2_{it}).$$

The parameter κ_{it} is included to capture the mass of individuals with zero spending that are observed in the data. If κ_{it} is positive, then negative health states are permitted, which

[45] Our model cannot distinguish between potential reasons why premiums may be valued differently from out-of-pocket costs. For example, we expect the tax deductibility of premiums would push α^{OOP} up, while systematic underestimation of out-of-pocket spending would push α^{OOP} down.

[46] We calculate the parameters of the approximating distribution using the Fenton-Wilkinson method; additional details can be found in Appendix C.2.1

may imply zero spending.[47] Parameters μ_{it}, σ_{it}, and κ_{it} are parameterized to vary with individual demographics, including risk score, which can vary over time.

A household k faces an uncertain health state $\tilde{l}_{kt}$ that has a shifted lognormal distribution with parameters μ_{kt} and σ_{kt} and support $(-\kappa_{kt}, \infty)$. Under the approximation, household parameters μ_{kt}, σ_{kt}, and κ_{kt} can be calculated as functions of the individual parameters μ_{it}, σ_{it}, and κ_{it} of the individuals in the household. Variation in μ_{kt}, σ_{kt}, and κ_{kt} across households and within households over time arises from variation in household composition: the number of individuals and each individual's demographics. In addition to this observable heterogeneity, we also incorporate unobserved heterogeneity in household health though parameter μ_{kt}. In this way, adverse selection (on unobservables) is permitted because households can hold private information about their health that can drive both plan choice decisions and spending outcomes.

Finally, to account for the fact that there are multiple insurers in our empirical setting, we introduce an additional set of parameters ϕ_f to serve as exchange rates for monetary health states across insurers. These parameters are intended to capture differences in total healthcare spending that are driven by differences in provider prices across insurers. For example, an identical doctor's visit might lead to different amounts of total spending across insurers simply because each insurer paid the doctor a different price. We do not want such variation to be attributed to differences in underlying health or healthcare utilization. We therefore capture it in a structured way by estimating insurer-level parameters that multiply realized health states, transforming them from underlying "quantities" into the monetary spending amounts that we observe in the claims data.[48] A household's money-metric health state l is then the product of an insurer-level multiplier ϕ_f and the underlying

[47] If a household realizes a negative health state, this implies zero spending as long as ω_k is not too large that optimal spending becomes positive. Operationally, this entails amending the optimal spending policy to be: $m^*_{jt}(l, \omega_k) = \max(0, \omega_k(1 - c'_{jt}(m^*_{jt})) + l)$.

[48] In reality, ϕ_f will also capture other multiplicative differences across insurers such as care management protocols or provider practice patterns, but we find it likely that most of the variation in ϕ_f comes from differences in average provider prices across insurers. Our estimates of ϕ_f conform to our priors on provider price variation across insurers, most notably that Kaiser pays lower prices.

"quantity" health state $\tilde{l}$, where $\tilde{l}$ is lognormally distributed depending only on household characteristics. Taken together, the distribution F_{kft} is defined by

$$l = \phi_f \tilde{l},$$
$$\log(\tilde{l} + \kappa_{kt}) \sim N(\mu_{kt}, \sigma^2_{kt}).$$

3.4.2 Identification

We aim to recover the joint distribution across households of willingness to pay, risk protection, and the social cost of moral hazard associated with different levels of insurance. Variation in these objects arises from variation in either household preferences (risk aversion and moral hazard parameters) or in the ex ante distribution of health states. Our primary identification concerns are (i) distinguishing preferences from private information about health, (ii) distinguishing taste for mean out-of-pocket spending (α^{OOP}) from risk aversion, and (iii) identifying heterogeneity in the risk aversion and moral hazard parameters.

We first explain how ω, capturing moral hazard, is distinguished from unobserved variation in μ_{kt}, capturing adverse selection. In the data, there is a strong correlation between chosen plan generosity and total healthcare spending (see Figure 3.2a). A large part of this relationship can be explained by observable household characteristics[49]. However, conditional on observables, there is residual positive correlation between chosen coverage level and spending. This residual correlation could be attributable to either the effect of lower out-of-pocket prices driving utilization (moral hazard) or private information about health affecting both utilization and coverage choice (adverse selection). Just as in the instrumental variables analysis in Section 3.3.2, the key to distinguishing between these two explanations is the variation in plan menus.

We observe similar households facing different menus of plans.[50] As a result, some households are more likely to choose higher coverage only because of the plan menu they face. The level of moral hazard ω is identified by the extent to which households facing

[49], yielding the standard result that observably sicker consumers are willing to pay more for insurance

[50]Our identification argument for moral hazard is similar to that made in Cardon and Hendel (2001).

more generous plan menus also have higher healthcare spending. On the other hand, we also observe cases where similar households face similar menus of plans, but make different plan choices. This variation identifies the degree of private information about health, as well as the magnitude of the idiosyncratic preference shock ϵ. Conditional on observables, if households that choose more generous coverage also realize higher healthcare spending, this variation in plan choice will be attributed to private information about health. Otherwise, any residual unexplained variation in plan choice will be attributed to the idiosyncratic preference shock.

Both risk aversion and the relative valuation of premiums and out-of-pocket spending (α^{OOP}) affect households' preference for more or less generous insurance but do not affect their healthcare spending. To distinguish between these parameters, we use cases where observably different households face similar menus of plans. Risk aversion is identified by the degree to which households' taste for higher coverage is positively related to uncertainty in out-of-pocket spending, holding expected out-of-pocket spending fixed. α^{OOP} is identified by the rate at which households trade off premiums with expected out-of-pocket spending, holding uncertainty in out-of-pocket spending fixed.

Unlike the preceding arguments, identification of unobserved heterogeneity in risk aversion ψ and the moral hazard parameter ω relies on the panel nature of our data. Plan menus, household characteristics, and plan characteristics change over time. We therefore observe the same households making choices under different circumstances. If we had a large number of observations for each household and sufficient variation in circumstances, the preceding arguments could be applied household by household to identify a household-specific value of ψ and ω. In this case, heterogeneity in these parameters across households would be nonparametrically identified. In reality, we have at most five observations of each household. We ask less of this data by placing a parametric form on the distribution of types and estimating only the variance and covariance of types across households. As an example, if some households consistently make choices consistent with high risk aversion and others consistently make choices consistent with low risk aversion, this will show up as

a high variance in the unobserved component of the risk aversion parameter.

3.4.3 Estimation

We allow the parameters of the individual health state distributions μ_{it}, σ_{it}, and κ_{it} to vary by time-varying individual demographics:

$$\begin{aligned} \mu_{it} &= \boldsymbol{\beta}^{\mu} \mathbf{X}_{it}^{\mu}, \\ \sigma_{it} &= \boldsymbol{\beta}^{\sigma} \mathbf{X}_{it}^{\sigma}, \\ \kappa_{it} &= \boldsymbol{\beta}^{\kappa} \mathbf{X}_{it}^{\kappa}. \end{aligned} \tag{3.12}$$

$\mathbf{X}_{it}^{\mu}$, $\mathbf{X}_{it}^{\sigma}$, and $\mathbf{X}_{it}^{\kappa}$ contain indicators for the 0–30th, 30–60th, 60–90th, and 90–100th percentiles of individual risk scores.[51] $\mathbf{X}_{it}^{\mu}$ and $\mathbf{X}_{it}^{\kappa}$ also contain a linear term in risk score, which is estimated separately for the 0–90th risk score percentile group and the 90–100th percentile group. $\mathbf{X}_{it}^{\mu}$ also contains an indicator for whether the individual is under 18 years old and for whether the individual is a female between the ages of 18 and 30.

Using the derivations shown in Appendix C.2.1, household health state distribution parameters are calculated as a function of individual parameters:

$$\begin{aligned} \sigma_{kt}^2 &= \log[1 + \left[\sum_{i \in \mathcal{I}_k} \exp(\mu_{it} + \frac{\sigma_{it}^2}{2})\right]^{-2} \sum_{i \in \mathcal{I}_k} (\exp(\sigma_{it}^2) - 1) \exp(2\mu_{it} + \sigma_{it}^2)], \\ \bar{\mu}_{kt} &= -\frac{\sigma_{kt}^2}{2} + \log[\sum_{i \in \mathcal{I}_k} \exp(\mu_{it} + \frac{\sigma_{it}^2}{2})], \\ \kappa_{kt} &= \sum_{i \in \mathcal{I}_k} \kappa_{it}, \end{aligned} \tag{3.13}$$

where $\mathcal{I}_k$ represents the set of individuals in household k. We incorporate private information about health at the household level by adding normally distributed unobservable heterogeneity in μ_{kt}. The household-specific mean of μ_{kt} is given by $\bar{\mu}_{kt}$, and the variance is given by σ_{μ}^2. A large σ_{μ}^2 means that households have substantial private information about their health that cannot be explained by observables.

[51] As the distribution of risk score is highly right skewed, these groupings allow us to fit the data better than if we use true quartiles.

We similarly model the risk aversion (ψ_k) and moral hazard (ω_k) parameters with both observable and unobservable heterogeneity. Across parameters, we assume that μ_{kt}, ψ_k, and ω_k are jointly normally distributed, according to

$$\begin{bmatrix} \mu_{kt} \\ \omega_k \\ \log(\psi_k) \end{bmatrix} \sim N\left(\begin{bmatrix} \bar{\mu}_{kt} \\ \boldsymbol{\beta}^{\omega}\mathbf{X}_k^{\omega} \\ \boldsymbol{\beta}^{\psi}\mathbf{X}_k^{\psi} \end{bmatrix}, \begin{bmatrix} \sigma_{\mu}^2 & & \\ \sigma_{\omega,\mu}^2 & \sigma_{\omega}^2 & \\ \sigma_{\psi,\mu}^2 & \sigma_{\omega,\psi}^2 & \sigma_{\psi}^2 \end{bmatrix} \right). \tag{3.14}$$

Covariates $\mathbf{X}_k^{\omega}$ and $\mathbf{X}_k^{\psi}$ include an indicator for whether the household has children and a constant.[52]

We model inertia at both the plan and the insurer level: $\mathbf{fl}_{kjt}^{inertia} = \gamma_k^{plan}\mathbf{1}_{k,j=j(t-1)} + \gamma_k^{ins}\mathbf{1}_{k,f=f(t-1)}$. We allow γ_k^{plan} to vary linearly with household age and allow the intercept to vary by whether the household has children.[53] We allow γ_k^{ins} to vary linearly with household risk score. We include household risk score here to capture whether sicker households face higher barriers to switching insurers (and therefore provider networks). Additionally, in 2013, Moda rebranded and changed the names of all of its plans, and added a plan, in a way that did not result in a direct mapping between all 2012 and 2013 plans. To capture this flexibly, we estimate a separate insurer-level inertia parameter for Moda plans in 2013. We allow insurer fixed effects ($\delta_k^{f(j)}$) to vary by household age and whether a household has children, and allow the intercepts to vary by geographic region to capture the relative attractiveness of insurer provider networks across different parts of the state (as well as other sources of geographical heterogeneity in insurer preferences). We normalize the insurer fixed effect for Moda to be zero. As the parameters of the individual health state distributions are allowed to vary freely, the "provider price" parameters require normalization: ϕ_{Moda} is normalized to one.

We estimate the model via simulated maximum likelihood. Our estimation approach follows Revelt and Train (1998) and Train (2009), with the important distinction that we

[52] If a household changes whether they have children during the sample, we assign it to its modal status.

[53] Household age is calculated as the mean age of all adults in a household across all years.

model a discrete/continuous choice. Our construction of the discrete/continuous likelihood follows Dubin and McFadden (1984). The likelihood function for a given household is the conditional density of its observed sequence of total healthcare spending amounts, given its observed sequence of plan choices. We use Gaussian quadrature to integrate numerically over the distribution of unobserved heterogeneity as well as the distributions of household health states. Additional details on the estimation procedure are provided in Appendix C.2.2.

3.5 Results

3.5.1 Model estimates

Table 3.4 presents the estimated parameters of our empirical model. Column 3 presents our primary specification as described in Section 3.4. Columns 1 and 2 present simpler specifications that are useful in understanding and validating the model. The table excludes insurer fixed effects and health state distribution parameters; these can be found in Table C.10.

Column 1 presents a version of the model where there is no moral hazard and there is no heterogeneity in health across individuals. That is, we do not allow μ_{it}, σ_{it}, or κ_{it} to vary by observable individual characteristics. However, unobservable heterogeneity in household health (through σ_μ) is still permitted. In column 2, we introduce the full extent of observable individual heterogeneity in health. A key difference across columns 1 and 2 is in the magnitude of the adverse selection parameter σ_μ, which falls by more than half. When rich observable heterogeneity in health is introduced to the model, the estimated amount of unobservable heterogeneity in health falls substantially. Moral hazard is introduced in column 3. Here, an important difference is the increase in the estimated amount of risk aversion. When moral hazard is introduced, the model can explain a larger part of the dispersion in spending for observably similar households. This implies that households are facing less risk, and that more risk aversion is necessary to explain the same plan

Table 3.4: *Parameter Estimates*

	(1)		(2)		(3)	
Variable	Parameter	Std. Err.	Parameter	Std. Err.	Parameter	Std. Err.
Employee Premium ($000s)	$-1.000^{\dagger}$		$-1.000^{\dagger}$		$-1.000^{\dagger}$	
OOP spending, $-\alpha^{OOP}$	−1.504	0.024	−1.519	0.024	−1.348	0.028
HRA/HSA contrib., α^{HA}	0.292	0.023	0.293	0.023	0.250	0.023
Vision/dental contrib., α^{VD}	1.346	0.025	1.340	0.025	1.143	0.037
Plan inertia, γ^{plan}	4.272	0.095	5.009	0.059	4.265	0.098
Plan inertia * (Age−40), γ^{plan}	0.019	0.002	0.073	0.006	0.018	0.002
Plan inertia * **1**[Children], γ^{plan}	0.189	0.040	1.208	0.119	0.188	0.041
Insurer inertia,γ^{ins}	6.097	0.116	4.605	0.231	6.030	0.120
Insurer inertia * Risk score, γ^{ins}	0.182	0.026	0.501	0.074	0.117	0.026
Moda-specific inertia, 2013	1.824	0.196	1.924	0.199	1.555	0.198
Moda narrow net. plan	−2.662	0.165	−2.665	0.165	−2.459	0.169
Kaiser prov. price, ϕ_K	0.669	0.007	0.831	0.006	0.766	0.000
Providence prov. price, ϕ_P	1.038	0.017	1.096	0.017	1.061	0.006
Risk aversion intercept, β^{ψ}	−0.495	0.059	−0.597	0.065	0.313	0.049
Risk aversion * **1**[Children], β^{ψ}	−0.344	0.070	−0.221	0.062	−1.103	0.096
SD of risk aversion, σ_{ψ}	0.921	0.037	0.997	0.102	0.603	0.131
SD of mu, σ_{μ}	0.853	0.003	0.314	0.049	0.271	0.005
Moral hazard intercept, β^{ω}					1.133	0.000
Moral hazard * **1**[Children], β^{ω}					0.615	0.000
SD of moral hazard, σ_{ω}					0.145	0.073
Corr(μ, ψ), $\rho_{\mu,\psi}$	0.354	0.000	0.168	0.088	0.710	0.102
Corr(ψ, ω), $\rho_{\psi,\omega}$					−0.168	0.045
Corr(μ, ω), $\rho_{\mu,\omega}$					0.027	0.013
Scale of logit error, σ_{ϵ}	2.516	0.027	2.519	0.027	2.406	0.028
Insurer * {Region, Age, **1**[Child.]}	Yes		Yes		Yes	
Heterogeneity in spending dists.			Yes		Yes	
Number of observations	679,773		679,773		679,773	

Notes: This table presents parameter estimates from our empirical model. Column 3 presents our primary estimates, while columns 1 and 2 present alternative specifications. All models are estimated on an unbalanced panel of 44,562 households over five years. Coefficients of absolute risk aversion are relative to thousands of dollars. Estimates from column 3 are the inputs into the calculation in Section 3.5.2. To make non-interacted coefficients more readily interpretable, we use (Age−40). †By normalization.

choices. Because estimated risk aversion increases, the relative valuation of premiums and out-of-pocket costs (α^{OOP}), which had been compensating for low risk aversion, falls.

In column 3, we estimate an average moral hazard parameter (ω) of $1,115 among individuals and $1,542 among families.[54] Recall that ω represents the additional total spending that would be induced when moving a household from no insurance to full

[54]For comparison, the average ω estimated by Einav et al. (2013) is $1,330.

insurance. [55] For scale, we estimate an average household health state of $4,702 for individual households and $11,044 for families. These estimates imply that moving from a plan with a 50 percent coinsurance rate to full insurance would result in an increase in total healthcare spending equal to 11 percent of mean unavoidable spending for individuals, and 7 percent for families.

We estimate a large degree of risk aversion. Our estimates imply a mean (median) coefficient of absolute risk aversion of 1.12 (0.84) across households.[56] Put differently, to make households indifferent between (i) a payoff of zero, and (ii) an equal odds gamble between gaining $100 and losing $X, the mean (median) value of $X in our population is $90.17 ($92.94).[57] Our estimates of risk aversion are with respect to financial risk as well as health risk (through b^*_{jt}), and so are not directly comparable to estimates that consider only financial risk. The standard deviation of the uncertain portion of payoffs ($b^*_{jt} - c^*_{jt}$) with respect to the distribution of health states is $853 on average across households-plans-years. This corresponds to an average standard deviation of out-of-pocket costs of $1,358. To avoid a normally distributed lottery (in units of $b^*_{jt} - c^*_{jt}$) with mean zero and standard deviation $853, the median household would be willing to pay $305.

The importance of unobserved heterogeneity varies for health, risk aversion, and moral hazard.[58] The estimated amount of private information about health is fairly small once we account for the full set of household observables as well as moral hazard: unobserved heterogeneity in μ_{kt} accounts for 8 percent of the total variation in μ_{kt} across household-

[55] We find that the moral hazard parameter is increasing in age, to the extent that an additional 10 years of age increases ω by $XX.

[56] We measure monetary variables in thousands of dollars; dividing our estimated coefficients of absolute risk aversion by 1,000 makes them comparable to estimates that use risk is measured in dollars.

[57] In this example, a risk neutral household would have a value of $X equal to $100 and an infinitely risk averse household would have a value of $X equal to $0. Using the same example, Handel (2013) reports a mean $X of $91.0, Einav et al. (2013) report a mean $X of $84.0, and Cohen and Einav (2007) report a mean $X of $76.5.

[58] Following Revelt and Train (2001), we derive each household's posterior type distribution using Bayes' rule, conditioning on their observed choices and the population distribution. For the purposes of examining total variation in types across households (accounting for both observed and unobserved heterogeneity), we assign each household the expectation of their type with respect to their posterior distribution. This procedure is described in detail in Appendix C.2.3.

years.[59] Unobserved heterogeneity in the moral hazard parameter accounts for 9 percent of its total variation across households. On the other hand, unobserved heterogeneity in risk aversion accounts for 54 percent of its total variation.

Conditional on observables, we find that households that are idiosyncratically risk averse also have private information that they are unhealthy ($\rho_{\mu,\psi} > 0$) and are less prone to moral hazard than expected ($\rho_{\psi,\omega} < 0$). We find that households with private information that they are unhealthy are also more prone to moral hazard than expected ($\rho_{\mu,\omega} > 0$). Accounting for both unobservable *and* observable variation, we find that risk aversion and moral hazard have a strong negative correlation of -0.90. Among households with (without) children, expected health state $\mathbb{E}[\tilde{l}]$ has a correlation of 0.15 (0.13) with risk aversion, and a correlation of 0.05 (0.08) with the moral hazard parameter. Figure C.3 plots the unconditional joint distribution of these three key dimensions of household type.

Our estimates imply substantial disutility from switching insurers and plans. Average disutility across households from switching insurers is \$6,372, with a standard deviation of \$91. Average disutility from switching plans (but not insurers) is \$4,466, with a standard deviation of \$1,739. We estimate that insurer inertia is increasing in household risk score, and that plan inertia is increasing in household age and is on average \$188 higher for households with children.[60] The exceptionally large magnitudes of our inertia coefficients reflect in large part the infrequency with which households switch plans and insurers, as described in Table 3.1. Only 3.3 percent of household-years ever voluntarily switch insurers and 13.6 percent of household-years ever voluntarily switch plans.

Finally, the estimates in column 3 indicate that households weight out-of-pocket expenditures 34.8 percent more than plan premiums. We believe this could be driven by a variety of factors, including (i) household premiums are tax deductible, while out-of-pocket expenditures are not; and (ii) employee premiums are very low (at the median, zero), per-

[59] This finding is consistent with the minimal selection on unobservables found by Cardon and Hendel (2001).

[60] We do not investigate the micro-foundations of our estimates of household disutility from switching; see Handel (2013) for a full treatment of inertia in health insurance.

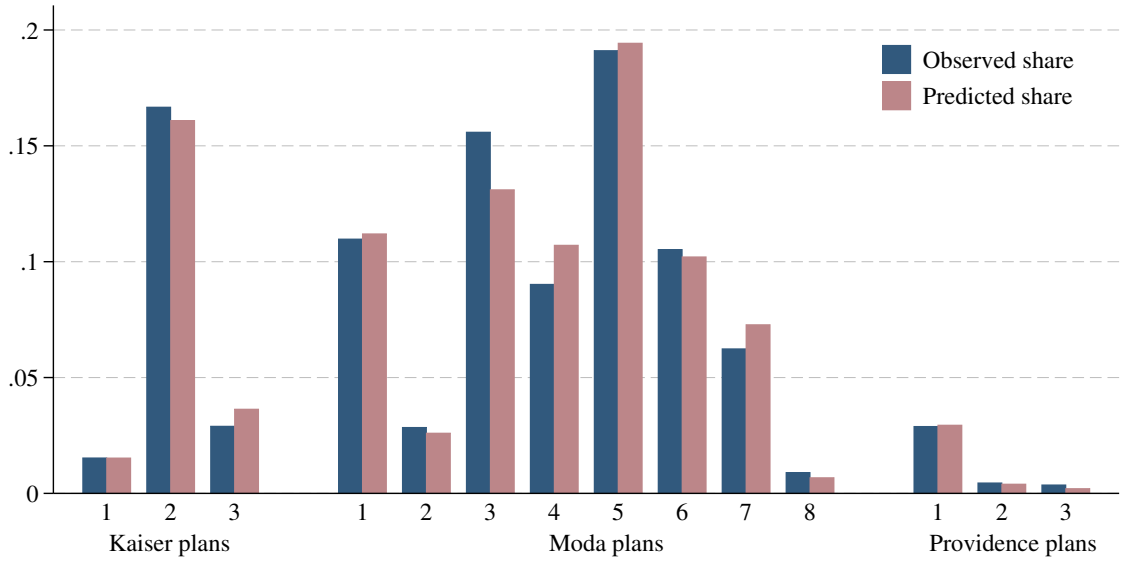
haps making potential out-of-pocket costs in the thousands of dollars seem relatively more salient. A single household in Oregon with income of $80,000 paid an effective state plus federal income tax rate of 28.9 percent in 2013. Using this tax rate, a dollar of out-of-pocket spending (after-tax) would be equivalent to 1.41 dollars of premiums (pre-tax). We also find that households value a dollar in HSA/HRA contributions on average 75 percent less than a dollar of premiums. This is consistent with substantial hassle costs associated with these types of accounts, as documented by Reed, Fung, Price, Brand, Benedetti, Derose, Newhouse and Hsu (2009) and McManus, Berman, McInerny and Tang (2006).

Model Fit. We conduct two procedures to evaluate model fit, corresponding to the two stages of the model. First, we compare households' predicted plan choices to those observed in the data. Figure 3.3 displays the predicted and observed market shares for each plan, pooled across all years in our sample.[61] Shares are matched exactly at the insurer level due to the presence of insurer fixed effects, but are not matched exactly plan by plan. Predicted choice probabilities over plans within an insurer are driven by plan prices, inertia, and households' valuation of different levels of coverage through their expectation of out-of-pocket spending, their value of risk protection, and their expectation of utility from the consumption of healthcare services. Given the relative inflexibility of the model with respect to household choice of coverage level within an insurer, the fit is quite good.

In our second exercise, we compare the predicted distributions of households' total healthcare spending to the distributions of total healthcare spending we observe in the data. In a given year, each household faces a predicted distribution of health states and a corresponding plan-specific distribution of total healthcare spending, as defined by our model and estimated parameters. To construct the predicted distribution of total spending in a population of households, we take a random draw from the predicted distribution of each household corresponding to the household's chosen plan. Figure 3.4 presents kernel

[61] Figure C.4 provides the corresponding comparisons separately for each year. As another metric, the model predicts 72 percent of household plan choices correctly (assigns the highest predicted probability to the correct plan). If households were modeled as choosing randomly from their plan choice set, 23 percent of plan choices would be predicted correctly (i.e., the average choice set size is approximately 4 plans).

Figure 3.3: *Model Fit: Plan Choices*



Notes: The figure shows predicted and observed market shares at the plan level. All years are pool together, so the observation is the household-year. Predicted shares are calculated using the estimates in column 3 of Table 3.4 and Table C.10.

density plots of the predicted and observed distribution of household total spending among household-years enrolled by each insurer.[62] The vertical lines in each plot represent the mean of the respective distribution. Overall across all household-year observations, average total healthcare spending is observed to be $10,754 and is predicted to be $10,738.

3.5.2 Willingness to Pay and Social Surplus

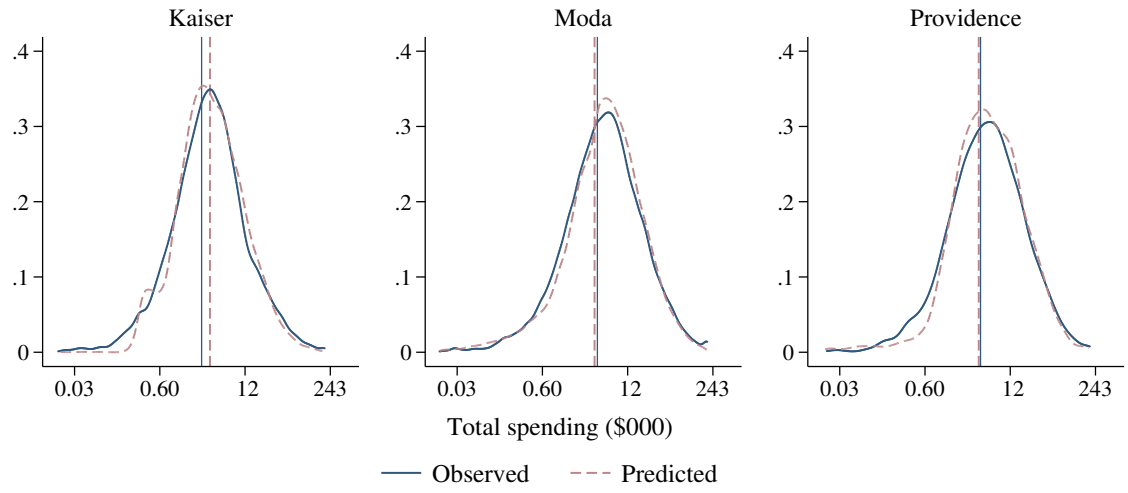
Using our estimates, we next construct each household's willingness to pay for insurance, as well as the social surplus generated by its allocation to different levels of insurance. [63] We conduct our remaining analyses on a set of vertically differentiated plans that roughly correspond to the types of coverage offered on Affordable Care Act (ACA) exchanges.[64]

[62]Figures C.5 and C.6 present similar comparisons by family size and quartile of household risk score.

[63]More precisely, we construct each household's *marginal* willingness to pay and *marginal* social surplus between any two contracts, assuming that all non-financial features of the contracts are held fixed.

[64]We use these "artificial" plans instead of the set of Moda plans in our data because the Moda plans are densely packed in coverage space and are also not perfectly vertically differentiated. The plans we do consider

Figure 3.4: *Model Fit: Healthcare Spending*



Notes: The figure shows kernel density plots of the predicted and observed distribution of total healthcare spending on a log scale among households enrolled with each of the three insurers. All years are pooled together, so the observation is the household-year. The vertical lines represent the mean of the respective distribution. Predicted distributions are estimated using parameter estimates from column 3 in Table 3.4 and Table C.10.

Our candidate plans are 'Full insurance', 'Gold', 'Silver', 'Bronze', and 'Catastrophic', corresponding to an actuarial value of 1.00, 0.85, 0.70, 0.60, and 0.50.[65] The out-of-pocket cost functions of these plans are depicted in Figure C.7.

Willingness to Pay. For the purposes of our remaining analyses, we put aside intertemporal variation in households' estimated distribution of health states and focus on the first year that each household appears in the data. We also use the provider price parameter $\phi = 1$, corresponding to that used for Moda. This leaves us with one type for each household: $\{F_k, \psi_k, \omega_k\}$, just as in our theoretical model.[66] We first express utility in certainty equivalent

span the full range of Moda plans offered, but are evenly distributed in coverage space and are truly vertically differentiated.

[65]These actuarial values are calculated with respect to the population in our data. The exact deductible, coinsurance rate, and out-of-pocket maximum of the plans are $1,000, 15%, $2,000 for Gold; $3,500, 20%, $4,500 for Silver; $7,000, 30%, $7,500 for Bronze; and $10,000, 30%, $10,000 for Catastrophic.

[66]To account for unobservable heterogeneity, we assign household types by integrating over each household's posterior distribution of types. This caveat likewise applies to the calculations of certainty equivalent and social surplus that follow. We omit these steps in this section because the notation is cumbersome, but it is provided in Appendix C.2.3.

units:

$$\begin{aligned} CE_{kj} &= -\psi_k^{-1} \log(-U_{kj}) \\ &= \bar{x}_{kj} - \psi_k^{-1} \log\left(\int_0^\infty \exp\left(-\psi_k (x_{kj}(l) - \bar{x}_{kj}\right) dF_k(l) \right), \end{aligned}$$

where $x_{kj}(l)$ is the payoff associated with health state l in plan j (equation (3.11)), and $\bar{x}_{kj}$ is the expectation of $x_{kj}(l)$ with respect to the distribution of l. Willingness to pay for marginally more generous insurance is equal to the difference in certainty equivalent between a (higher coverage) focal plan and the (lowest coverage) reference plan (j_0), when both plans have zero premium. We make comparisons over plans holding all non-financial features fixed, so inertia terms and insurer fixed effects cancel. We set α^{OOP} to one so that premiums and out-of-pocket costs are valued one-for-one.[67] With attention restricted to the dimension of coverage level, willingness to pay depends only on the benefit of healthcare spending, out-of-pocket costs, and riskiness in both:

$$\begin{aligned} WTP_{kj} &= CE_{kj} - CE_{k,j_0} \\ &= \bar{c}_{k,j_0} - \bar{c}_{kj} + \bar{b}_{kj} - \bar{b}_{k,j_0} + \Psi_{kj}, \end{aligned}$$

where $\bar{c}_{kj}$ is the expectation of out-of-pocket costs $c_j(m_j^*(l, \omega_k))$ with respect to the distribution of l, and $\bar{b}_{kj}$ is defined similarly. As in our theoretical expression for WTP, we pull out the mean and leave deviations from the mean lumped into Ψ_{kj}, which measures the value of risk protection. If consumers are risk averse and plan j provides a less risky distribution of out-of-pocket spending than does plan j_0, Ψ_{kj} will be positive. Whereas our theoretical reference plan was the null contract x_0, our empirical reference plan j_0 is the Catastrophic plan. We hereinafter refer to "willingness to pay" for a given plan, but bear in mind that this is *marginal* willingness to pay with respect to this particular reference point.

Figure 3.5 presents the distribution of willingness to pay among family households.[68]

[67] We do this because otherwise welfare could be created simply by moving a dollar of spending between premiums and out-of-pocket, which we find undesirable. If we leave α^{OOP} as estimated, optimal levels of insurance increase as out-of-pocket costs are so disliked.

[68] We focus our analysis of results on the set of family households because families make up 75 percent of

Households are ordered on the horizontal axis according to their willingness to pay. The highest willingness to pay households are on the left, as in a demand curve. Figure 3.5, as well as the figures that follow, is composed of connected binscatter plots. For each percentile of willingness to pay, households in that percentile are grouped together and the average value of the vertical axis variable (in this case, willingness to pay itself) is plotted for each plan. These 100 points for each plan are then connected with a line.[69] As the plans are vertically differentiated, all households are willing to pay more for higher coverage. The highest willingness to pay households are willing to pay $10,000 more for the full insurance plan rather than the Catastrophic plan.

As in equation (3.2), we can decompose willingness to pay for each plan into its three component parts: mean reduced out-of-pocket costs from unavoidable medical spending, mean benefit from moral hazard spending, and the value of risk protection. Recall that only the latter two components are relevant to social welfare. Figure C.8 presents this decomposition of willingness to pay for the Gold plan (the shape of the breakdown is similar for all plans). We find that mean reduced out-of-pocket costs for unavoidable medical spending represents the majority of willingness to pay for most households, but there is substantial variation across the distributions of willingness to pay. The highest willingness to pay households have nearly 100 percent of their willingness to pay driven by mean reduced out-of-pocket costs, while for the lowest willingness to pay households it is only about 25 percent. Importantly, this means that the highest willingness to pay households are not generating *any* social surplus from having more comprehensive insurance.

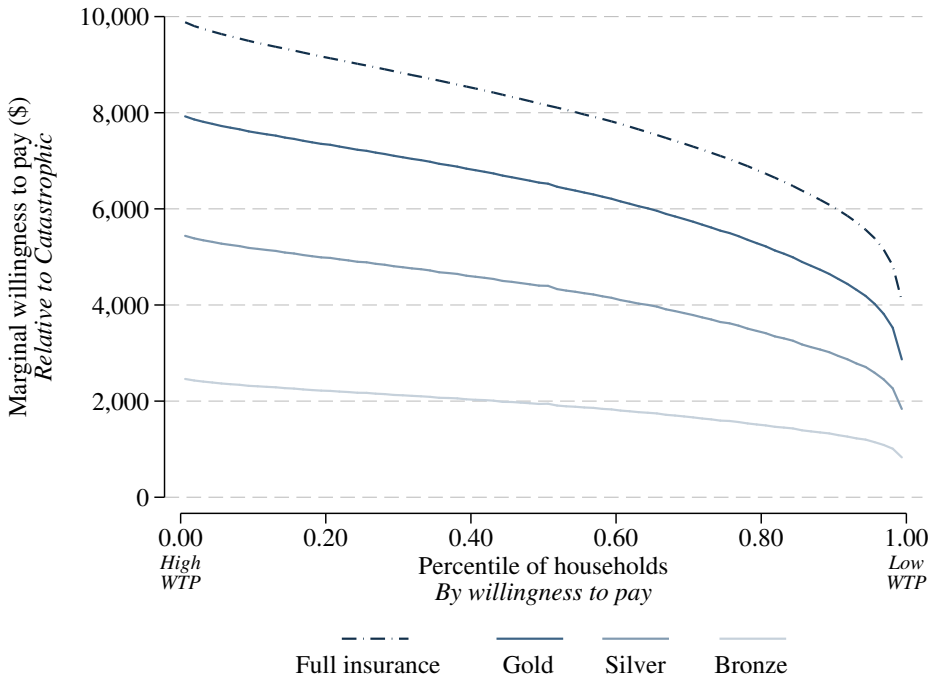
Social Surplus. Using willingness to pay, we can determine households' *privately* optimal plan choices given any premiums. We next specify *socially* optimal plan choices. As in Section 3.2, we calculate the social surplus generated by allocating a household to a given

---

our sample and our set of candidate plans is chosen to mimic the coverage levels offered to families.

[69]The households are in fact ordered by their willingness to pay for the Silver plan, but because the ordering is nearly identical across plans, the lines in this plot are monotonically decreasing and appear smooth (if it were not the case, the connected binscatter plot would have a "jagged" look). The fact that the ordering of households is the same across plans is important because it permits a graphical analysis on multiple plans analogous to that used in the two plan example in Figure 3.1.

Figure 3.5: *Willingness to Pay*



Notes: The figure shows the distribution of willingness to pay. Households are arranged on the horizontal axis according to their willingness to pay. The plot consists of four connected binscatter plots with respect to 100 bins of households ordered by willingness to pay.

plan as the difference between willingness to pay and expected insurer cost:

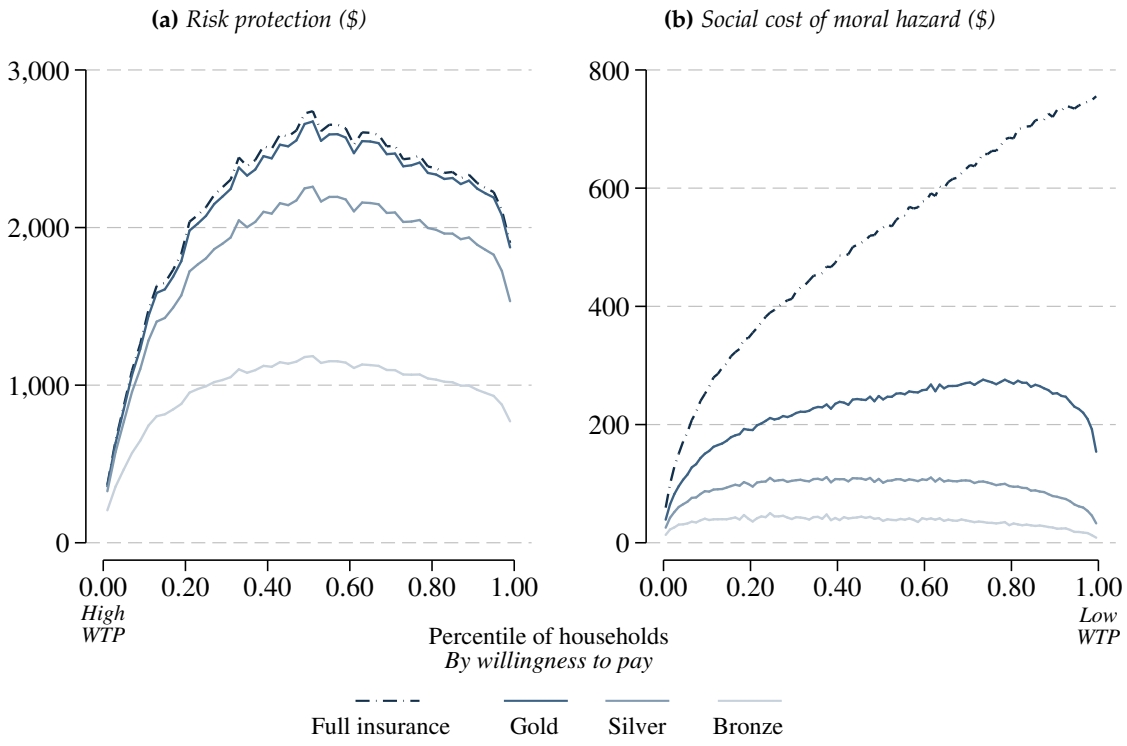
$$SS_{kj} = \underbrace{\Psi_{kj}}_{\substack{Value\ of\ risk \\ protection}} - \underbrace{\left((\bar{k}_{kj} - \bar{k}_{k,j_0}) - (\bar{c}_{k,j_0} - \bar{c}_{kj} + \bar{b}_{kj} - \bar{b}_{k,j_0})\right)}_{\substack{Social\ cost \\ of\ moral\ hazard}},$$

where $\bar{k}_{kj}$ is the expectation of insured spending $k_j(m_j^*(l, \omega_k))$ with respect to the distribution of l. The value of risk protection will vary in the population to the extent that there is variation in risk aversion and in the probability that households realize health states that would result in different levels of out-of-pocket cost across plans. The social cost of moral hazard will vary in the population to the extent that there is variation in the moral hazard parameter and in the probability that households realize health states that would result in different marginal out-of-pocket cost across plans.

To understand the contribution of each of these components to the overall relationship

between willingness to pay and social surplus, we first plot them separately. Figure 3.6a presents households' value of risk protection for each plan across the distribution of willingness to pay. We find that the majority of the social welfare gains from more generous insurance are driven by households with intermediate levels of willingness to pay. This "shape" of risk protection could be driven either by the distribution of risk aversion or the distribution of risk in the population. We investigate by examining the joint distribution of risk aversion and willingness to pay (see Figure C.9a). While there is substantial variation in the risk aversion parameter, average risk aversion is monotonically increasing in willingness to pay. The inverted U-shape in Figure 3.6a must therefore be driven by the shape of household *risk*.

Figure 3.6: *Value of Risk Protection and Social Cost of Moral Hazard*



Notes: The figure shows the marginal value of risk protection and the marginal social cost of moral hazard. Households are arranged on the horizontal axis according to their marginal willingness to pay. The left panel shows the marginal value of risk protection in the given plan relative to the Catastrophic plan. The right panel shows the marginal social cost of moral hazard in the given plan relative to the Catastrophic plan. Both panels are composed of connected binscatter plots with respect to 50 bins of households ordered by willingness to pay.

The inverted U-shape of risk makes sense given the nonlinear nature of the plans we consider. Very sick households are overwhelmingly likely to realize health states above the out-of-pocket maximum of every plan, leaving essentially no uncertainty in out-of-pocket spending. On the other hand, very healthy households are overwhelmingly likely to realize health states below the deductible of all plans, rendering the plans roughly identical for them. The households that do face substantial uncertainty in their out-of-pocket spending across plans are those for which much of the density of their health state distribution lies in the range of total spending where out-of-pocket costs vary both across plans and across health states.[70]

Figure 3.6b shows the distribution of the social cost of moral hazard. The figure provides two important insights. First, high willingness to pay households on average do not change their behavior across the range of plans we consider. [71] While they may have *already* been consuming more healthcare in the Catastrophic plan than they would have done absent any insurance at all, the difference between the full insurance plan and the Catastrophic plan is minimal. On the other hand, households with low willingness to pay on average do change their behavior substantially over this range of coverage levels. This pattern is driven by the interaction of health state distributions and the nonlinear contracts (treatment intensity), as well as by the fact that the household moral hazard parameter is decreasing in willingness to pay (treatment effect).[73] The second insight is that the Gold plan can recover more than half of the social cost of moral hazard induced by the full insurance plan. The $1,000 deductible is enough to undo the majority of the social cost of moral hazard under full insurance, while, as seen in Figure 3.6a, giving up only a small amount of risk protection.

Finally, we construct the social surplus curve for each plan by vertically summing Figure 3.6a and (the negative of) Figure 3.6b. Figure 3.7 presents the social surplus generated by

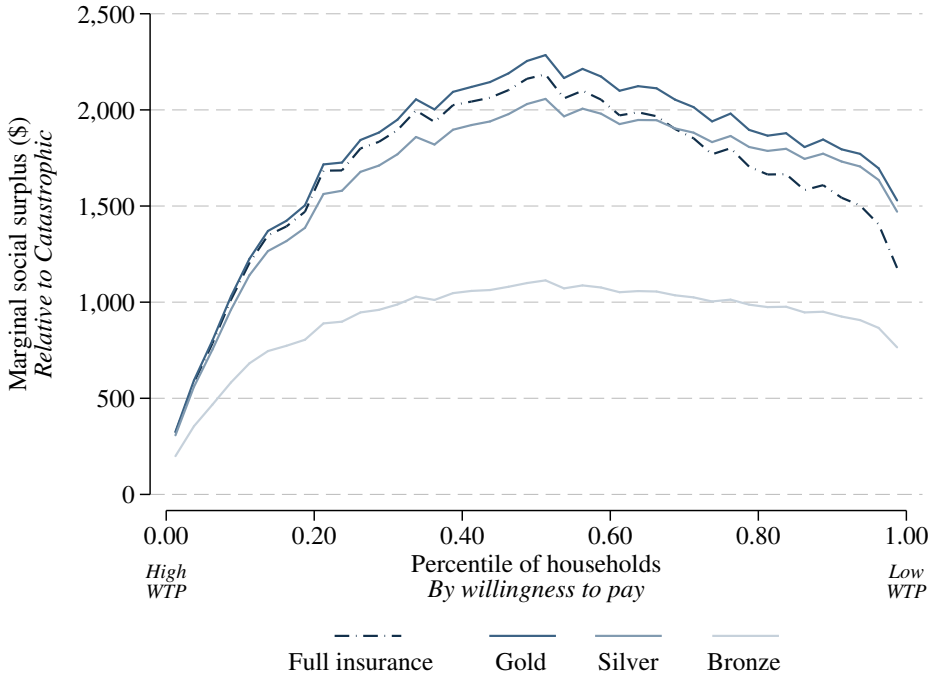
[70]Figure C.10 shows the distributions of health states faced by households, by willingness to pay.

7172

[73]Variation in treatment intensity can be inferred from the health state distributions at different levels of willingness to pay, shown in Figure C.10. Variation in treatment effect can be seen in the distribution of moral hazard parameter by willingness to pay, shown in Figure C.9b.

allocating households to a given plan relative to the Catastrophic plan. The plot consists of a connected binscatter for each plan, with respect to 50 (to reduce noise) quantiles of willingness to pay. At a given quantile of willingness to pay, the social surplus curves measure the average social surplus generated if all households at that quantile were allocated to a given plan.

Figure 3.7: *Social Surplus*



Notes: The figure shows the marginal value of social surplus among family households. Households are arranged on the horizontal axis according to their willingness to pay. The figure is composed of connected binscatter plots with respect to 50 bins of households ordered by willingness to pay.

The social surplus curves for all plans are above zero, meaning that the Catastrophic plan is the worst plan, from a social welfare perspective, at any level of willingness to pay. The Bronze plan is strictly second worst. Among the other plans, we find that the Gold plan generates weakly greater average surplus than any other plan at every level of willingness to pay. This figure is the empirical analog of the theoretical examples in the two-contract setting in Section 3.2.2. The Catastrophic plan is the "low" contract and the four others are potential "high" contracts. Vertical choice should only be offered if the high-willingness to

pay consumers should have more insurance than the low-willingness to pay consumers. As in the theoretical example, this statement corresponds to a "crossing" of upper-envelope social surplus curves, with the higher coverage plan to the left. Here, the upper envelope of social surplus curves is composed of a single plan. A regulator facing this population of consumers would find it optimal to forgo vertical choice and offer only the Gold plan.

While the Gold plan is the efficient plan *on average* at every level of willingness to pay, it is not the efficient plan for every household. Figure C.11 displays the heterogeneity in households' efficient plans. It shows that the Silver plan is the efficient level of coverage for 30 percent of households, full insurance is efficient for 1 percent of households, and the Gold plan is efficient for 69 percent of households. While the efficient coverage level does vary, it is not correlated with willingness to pay. The optimal feasible allocation under community rated prices would achieve social surplus equal to the integral of the Gold plan's social surplus curve in Figure 3.7. In the next section, we quantify welfare and compare outcomes under alternative pricing policies.

3.6 Counterfactual Pricing Policies

We compare outcomes under five pricing policies: (i) regulated pricing with community rating (our baseline), (ii) regulated pricing with type-specific prices, (iii) competitive pricing with community rating, (iv) competitive pricing with type-specific prices, and (v) subsidies to support full vertical choice. Regulated pricing is the baseline policy considered in this paper, in which the regulator can observe the distribution of consumer types and can set premiums (equivalent to a competitive market with taxes/subsidies). Competitive pricing is the case in which competition among private firms drives premiums to equal average costs on a plan by plan basis, rendering the market susceptible to unraveling due to adverse selection.[74] Subsidies to support full vertical choice is a policy of subsidies set with the

[74]Under this policy, there is still some regulation because the only plans that can be offered are the set of give potential plans we consider. This can be thought of as perfect competition with a mandate that all consumers purchase at least Catastrophic coverage.

intention of supporting the availability of (read: enrollment in) every plan.

We consider two scenarios in which prices can vary by consumer attributes: (ii) and (iv). If observable dimensions of consumer type are predictive of their efficient coverage level, allowing plan menus to be tailored to specific types may improve allocations. We divide consumers into four groups: childless households under age 55, childless households over age 55, households under age 40 with children, and households over age 40 with children. The age cutoffs are chosen to divide households in half within each group (childless or not). We choose age and whether the household has children because these are used in ACA exchanges and are also the important observables on which parameters of our model may vary.

Welfare Outcomes. Table 3.5 summarizes outcomes under each of our five pricing policies. It shows the percent of households Q enrolled in each plan at the optimal feasible allocation under the policy, the percent of first best social surplus that is achieved, and the average expected insurer cost AC among households that would choose to enroll in each plan at the prevailing premiums. Costs are measured in thousands of dollars. Social surplus is still normalized to zero for the Catastrophic plan. We benchmark outcomes against the first-best allocation of households to plans (as depicted in Figure C.11). This allocation cannot be supported by prices unless prices can vary by all aspects of consumer type, including risk aversion and the moral hazard parameter. The first-best allocation generates \$1,796 in social surplus per household relative to the counterfactual of allocating all households to the Catastrophic plan. Expected total healthcare spending per household under this allocation is \$12,140, and average expected insurer cost is \$10,067.

Alternative (i) is our baseline pricing policy where the regulator can design the market but is restricted to community-rated pricing. As indicated by Figure 3.7, under this scenario it is welfare maximizing to offer only Gold. The average expected insurer cost of all households in the Gold plan is \$10,706. In order to break even, the regulator sets a premium equal to \$10,706 per household.[75] Interestingly, though 31 percent of households are

[75]Given there is only one possible option, this mandatory premium can just as well be thought of as a tax on

Table 3.5: *Outcomes of Alternative Pricing Policies*

	Policy	% of First Best Surplus		Potential Plans: Full	Gold	Silver	Bronze	Ctstr.
*	First best	1.000	*Q:*	0.01	0.69	0.30	–	–
			AC:	5.55	9.37	11.67	–	–
(i)	Regulated pricing with community rating	0.965	*Q:*	–	1.00	–	–	–
			AC:	–	10.71	–	–	–
(ii)	Regulated pricing with type-specific prices	0.965	*Q:*	–	1.00	–	–	–
			AC:	–	10.71	–	–	–
(iii)	Competitive pricing with community rating	0.000	*Q:*	–	–	–	–	1.00
			AC:	–	–	–	–	6.21
(iv)	Competitive pricing with type-specific prices	0.230	*Q:*	–	–	–	0.33	0.67
			AC:	–	–	–	7.58	5.99
(v)	Subsidies to support vertical choice	0.797	*Q:*	0.01	0.07	0.63	0.28	0.01
			AC:	59.53	32.41	8.39	1.89	0.28

Notes: This table summarizes outcomes under the five pricing policies we consider as well as the first best outcome, among the 32,377 family households. At the first-best allocation, per-household social surplus is $1,796 and average expected insurer cost is $10,067. *Q* represents the percent of households enrolled in each plan, and *AC* represents the average expected insurer cost (in thousands of dollars) among households enrolled in a given plan.

misallocated under this policy, it generates 96.5 percent of the welfare generated under the first-best allocation. Among the households for whom the Gold plan was not optimal, there is little variation in social surplus between the three most generous plans. In fact among all households, the welfare gains from more generous insurance are concave in coverage level and fairly flat among the top plans. If the regulator were to offer only a single plan, the percent of first best surplus generated by allocating all households to the Bronze plan is 51 percent, to the Silver plan is 92 percent, to the Gold plan is 96 percent, and to full insurance is 91 percent.

Because pricing policy (i) is almost as efficient as the first best outcome, there is little scope for improvement by varying prices by consumer types in alternative (ii). Even so, we find that allowing the regulator to price discriminate does not improve allocational efficiency at all. Age and whether or not a household has children does not predict households'

incomes and a zero premium.

efficient level of insurance. Within each of the four household subgroups, the Gold plan is again the most efficient plan across the distribution of willingness to pay. The regulator therefore finds it optimal to only offer the Gold plan within each subgroup.

Alternative (iii) considers competitive pricing with community rating. We calculate the competitive equilibrium using the algorithm proposed in Azevedo and Gottlieb (2017).[76] We find that in this population, the market fully unravels to the Catastrophic plan. The premium and expected insurer cost per household at the Catastrophic plan is $6,210. While choice is permitted under this policy, the market cannot deliver it. Alternative (iv) considers which allocations could be supported under competitive pricing if prices could vary by consumer subgroup. We find that both populations of households without children (both above and below age 55) can support a pooling equilibrium at the Bronze plan. A higher coverage level can be supported within these subpopulations because there is less variation in willingness to pay. On the other hand, both markets for households with children still unravel to the Catastrophic plan.

The first four policies are natural benchmarks, but none turn out to feature vertical choice. The regulator bans vertical choice under regulated pricing, and adverse selection prevents the availability of choice under competitive pricing. In reality, vertical choice does exist. It is sustained in U.S. health insurance markets in part (if not all) by a variety of subsidies and tax policies. To mimic this status quo outcome, alternative (v) considers premiums that support enrollment shares matching the true metal-tier shares observed on ACA exchanges in 2018.[77] The targeted shares are those shown in Table 3.5. The premiums that can support these shares and break even in aggregate are $13,492 for full insurance, $11,536 for Gold, $9,102 for Silver, $6,992 for Bronze, and $6,085 for Catastrophic. Because households with mid-range willingness to pay (for whom social surplus increases steeply at

[76] Like the authors, we use a mass of behavioral consumers equal to 1 percent of the population of households. See Azevedo and Gottlieb (2017) for additional details.

[77] Shares are pulled from Kaiser Family Foundation "Marketplace Plan Selections by Metal Level," available at https://www.kff.org/health-reform/state-indicator/marketplace-plan-selections-by-metal-level. We map the Platinum coverage level to full insurance.

low coverage levels, see Figure 3.7) now choose the Silver plan, this allocation recovers 80 percent of first-best welfare.

Distributional Outcomes. The population faces an unavoidable healthcare spending bill of $11,455 per household. It is unavoidable because it arises even if all households have the least generous insurance (Catastrophic). While full insurance offers the benefit of additional risk protection, it would also raise the spending bill to $12,497 per household due to moral hazard. The spending bill is funded by a combination of out-of-pocket costs and insured costs. Insured costs are in turn funded by premiums or by taxes. We do not distinguish between the two: an increase in premiums on all plans by $5 is equivalent to a tax of $5. If all households had Catastrophic coverage, in expectation 49 percent of spending would be paid out-of-pocket and 51 percent of spending would be insured. If all households had full insurance, 100 percent of spending would be insured. There are therefore large differences among the policies in the source of funding for the population healthcare spending bill, and in turn, how evenly the spending bill is shared across households. If all households had full insurance, the spending bill would be split perfectly evenly in the population.[78] If all households had no insurance, each household would pay their own expected cost.[79]

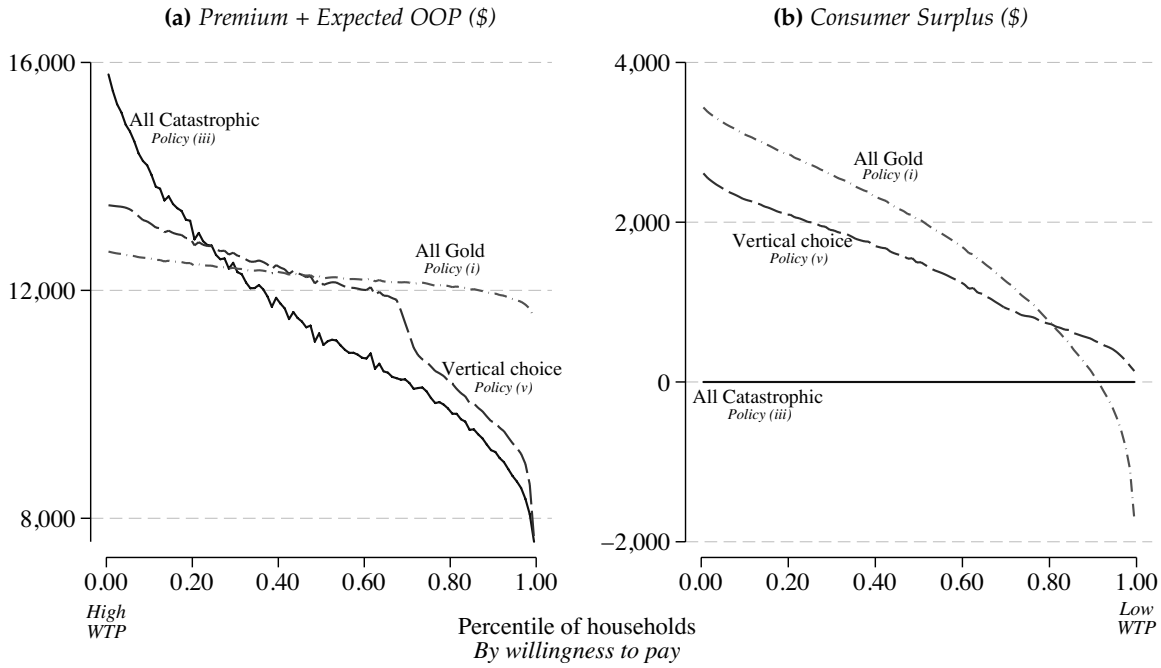
Figure 3.8 shows the distribution of premium plus out-of-pocket spending and of consumer surplus across the distribution of willingness to pay under three of our candidate policies: (i) regulated pricing ("All Gold"), (iii) competitive pricing ("All Catastrophic"), and (v) subsidies to support vertical choice ("Vertical Choice"). In Panel (a), for each policy we show the sum of the premium of households' chosen plan and households' expected out-of-pocket cost given their chosen utilization in their chosen plan. This is each household's healthcare spending bill under a given policy, and we show it across the distribution of

[78] In reality, if there were a single required premium, this would be assessed as a tax (as in countries that run a national health insurance scheme). In that case, premiums would not be split evenly, but according to the prevailing income tax system.

[79] In this case, the top 10 percentile of households by willingness to pay would pay 30 percent of the population spending bill, while the bottom 10 percentile of households by willingness to pay would pay 3 percent. Many households could not afford to pay their full expected cost, which is one reason why we do not consider a null contract.

willingness to pay.[80] For example, if all households enroll in Catastrophic, each pays a premium of $6,210. The top one percent of households by willingness-to-pay have expected out-of-pocket costs of $9,580 (for total spending of $15,790), while the bottom one percent have expected out-of-pocket costs of $1,379 (for total spending of $7,589). The population healthcare spending bill is split more evenly in the population when households have higher coverage.

Figure 3.8: *Distributional Outcomes*



Notes: The figure shows the distribution of premiums plus expected out-of-pocket (OOP) costs and of (marginal) consumer surplus in the population under three of the policies considered in Table 3.5. Because willingness to pay is calculated relative to the Catastrophic plan, consumer surplus is normalized to zero under the Catastrophic plan. The premium for the single plan is $6,210 under "All Catastrophic" and $10,706 under "All Gold." Premiums under "Vertical Choice" are $13,492 for full insurance, $11,536 for Gold, $9,102 for Silver, $6,992 for Bronze, and $6,085 for Catastrophic. Premiums break even in aggregate. Consumer surplus is equal to marginal willingness to pay less marginal premium. Households are arranged on the horizontal axis according to their willingness to pay.

Panel (b) of Figure 3.8 shows the distribution of consumer surplus under the three policies across the distribution of willingness to pay. For a given focal policy, we calculate

[80] All households at a particular level of willingness to pay choose the same plan and thus have the same premium, but there is variation in expected out-of-pocket cost at a given level of willingness to pay.

marginal consumer surplus for each household as their marginal willingness to pay for their chosen plan (relative the Catastrophic plan) less the marginal premium for their chosen plan. Marginal premium is equal to the difference between the premium of the household's chosen plan under the focal policy and the premium of the Catastrophic plan when all households are allocated to it ($6,210). The sum of consumer surplus across all households under a given policy equals the total social surplus generated by that policy. [81] We find that 91 percent of households prefer optimal regulation under policy (i) to the alternative of an unregulated (and unraveled) market. We find that *all* households prefer vertical choice under policy (v) to the alternative of an unraveled market. Strikingly, we also find that 81 percent of households prefer optimal regulation to vertical choice. While a shift to optimal regulation from vertical choice would make 19 percent of households worse off, only 9 percent of households would be at least $500 worse off. The shift would raise welfare by $302 per household per year.

3.7 Conclusion

This paper presents a framework for examining how policymakers can evaluate whether or not to offer a choice over coverage levels in health insurance markets. Our framework incorporates consumer heterogeneity along numerous dimensions, endogenous healthcare utilization through moral hazard, and permits menus of multiple nonlinear insurance contracts among which traded contracts are endogenous. Our analysis emphasizes the importance of distinguishing between the components of willingness to pay for insurance that are transfers and the components that are relevant to social welfare. Transfers play a large role in health insurance markets. Health status is persistent and contracts (at least in the U.S.) often span only a short, one-year time horizon.[82] The implication is that a

[81] The difference between the "All Gold" consumer surplus curve in Figure 3.8 and the Gold plan social surplus curve in Figure 3.7 is that the former shows who receives the surplus while the latter shows who generates the surplus. The integrals of the two curves are the same.

[82] Handel, Hendel and Whinston (2017) consider long-term contracts in health insurance markets. It would be interested to consider the welfare effects of vertical choice in that setting.

large part of healthcare spending can be foreseen, so it may not be possible to align the private incentive to maximize personal transfer and the social incentive to mitigate financial uncertainty. The presence of moral hazard means the problem is more complex than simply mandating full insurance for everyone.

We show that the key condition for vertical choice to be desirable is whether high willingness to pay consumers have higher efficient levels of coverage. In reverse, this implies that a lowest-coverage plan should only be offered if the lowest-willingness-to-pay consumers should have it. In our empirical setting, the lowest coverage plan we consider is a high deductible health plan. We find that low willingness to pay consumers are sufficiently risk averse to warrant higher coverage, and thus that a high deductible health plan should not be offered in the market. On the other hand, a highest-coverage plan should only be offered if the highest-willingness-to-pay consumers should have it. The highest coverage plan we consider is full insurance, and we find that it would more efficient for the high willingness-to-pay consumers to have less coverage. Between these extremes, we find that private values for coverage level are not positively correlated with social values, and thus that choice over coverage level should not be offered. We find that the best single plan to offer (among those we consider) has an actuarial value of 85 percent, but also that the social welfare stakes with respect to the exact plan design are low in the range of 80 percent to 90 percent actuarial value.

We limit our attention to a range of coverage levels over which uncertainty about healthcare utilization represents a purely financial gamble. Important considerations that our model does not address arise when consumers face liquidity constraints (Ericson and Sydnor, 2018) and when consumers are protected from large losses by limited liability in addition to by insurance (Gross and Notowidigdo, 2011). These distortions would become more pronounced outside the range of coverage levels we consider, and it would be interesting to explore their effects on our conclusions. In addition, the socially optimal level of healthcare utilization in our model is that which a consumer would choose if she were enrolled in the least generous insurance contract. If healthcare providers charge

supracompetitive prices or if there are externalities with respect to the consumption of healthcare services, it may be the case that inducing additional health spending with insurance is desirable. Such distortions would likely push up efficient coverage levels. Finally, an important simplification of our model is that healthcare is a homogenous good over which consumers must only choose the quantity to consume. In reality, healthcare is multidimensional and the time and space over which utilization decisions are made is complex. We see the extension of our model to capture other dimensions of healthcare utilization to be an important direction for future research.

References

AARP Public Policy Institute, "Caregivers of older adults: A focused look at those caring for someone age 50 +," *Research Report: Caregiving in the U.S. 2015*, 2015, (June), 1–53.

Abaluck, Jason and Jonathan Gruber, "Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance," *American Economic Review*, August 2016, *106* (8), 2145–84.

Agha, Leila, Brigham Frandsen, and James B Rebitzer, "Causes and Consequences of Fragmented Care Delivery: Theory, Evidence, and Public Policy," *NBER Working Paper 23078*, 2017.

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Alsan, Marcella, Owen Garrick, and Grant Graziani, "Does diversity matter for health? Experimental evidence from Oakland," *American Economic Review*, 2019, *109* (12), 4071–4111.

Aron-Dine, Aviva, Liran Einav, Amy Finkelstein, and Mark Cullen, "Moral Hazard in Health Insurance: Do Dynamic Incentives Matter?," *The Review of Economics and Statistics*, 2015, *97* (4), 725–741.

Arrow, Kenneth J., "Uncertainty and the Welfare Economics of Medical Care: Reply (The Implications of Transaction Costs and Adjustment Lags)," *The American Economic Review*, 1965, *55* (1/2), 154–158.

Artiga, Samantha, Anthony Damico, Katherine Young, Elizabeth Cornachione, and Rachel Garfield, "Health Coverage and Care for Immigrants," *Kaiser Family Foundation: Issue Brief*, 2016, *January*.

Azevedo, Eduardo M. and Daniel Gottlieb, "Perfect Competition in Markets With Adverse Selection," *Econometrica*, 2017, *85* (1), 67–105.

Azoulay, Pierre, Joshua S Graff Zivin, and Jialan Wang, "Superstar Extinction," *Quarterly Journal of Economics*, 2010, *125.2*, 549–589.

Barcellos, Silvia Helena, Dana P. Goldman, and James P. Smith, "Undiagnosed disease, especially diabetes, casts doubt on some of reported health 'advantage' of recent Mexican immigrants," *Health Affairs*, 2012, *31* (12), 2727–2737.

Barnett, Michael L., Zirui Song, Sherri Rose, Asaf Bitton, Michael E. Chernew, and Bruce E. Landon, "Insurance Transitions and Changes in Physician and Emergency Department Utilization: An Observational Study," *Journal of General Internal Medicine*, 2017, *32* (10), 1146–1155.

Bayard, Joanne M., Carol Calianno, and Cheryl L. Mee, "Care Coordinator-Blending Roles to Improve Patient Outcomes," *Nursing Management*, 1997, *28* (8), 49–52.

Billings, John, Nina Parikh, and Tod Mijanovich, "Emergency Room Use: The New York Story," *The Commonwealth Fund. Issue Brief*, 2000, (November), 1–11.

Bloomberg, Michael R, Amanda Burden, and Fatima Shama, "The Newest New Yorkers: Characteristics of the City's Foreign-born Population," *NYCPlanning*, 2013.

Boccuti, Cristina, Christa Fields, Giselle Casillas, and Liz Hamel, "Primary Care Physicians Accepting Medicare: A Snapshot," *The Kaiser Family Foundation*, 2015, (Issue Briefs).

Bogin, Alexander, William Doerner, and William Larson, "Local House Price Dynamics: New Indices and Stylized Facts," *Real Estate Economics*, 2019, *47* (2), 365–398.

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad, "What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics*," *The Quarterly Journal of Economics*, 04 2017, *132* (3), 1261–1318.

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney, "Pricing and welfare in health plan choice," *American Economic Review*, 2012, *102* (7), 3214–3248.

Capps, Randy, James D Bachmeier, Michael Fix, and Jennifer Van Hook, "A Demographic, Socioeconomic, and Health Coverage Profile of Unauthorized Immigrants in the United States," *Migration Policy Institute, Issue Brief*, 2013.

Cardon, James H. and Igal Hendel, "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey," *The RAND Journal of Economics*, 2001, *32* (3), 408–427.

Carlin, Caroline and Robert Town, "Sponsored Health Plans," *ReVision*, 2008, pp. 0–67.

CDC, "Estimates of Influenza Vaccination Coverage among Adults–United States, 2017-2018 Flu Season," 2018.

Chernew, Michael E., Lindsay Sabik, Amitabh Chandra, and Joseph P. Newhouse, "Would having more primary care doctors cut health spending growth?," *Health Affairs*, 2009, *28* (5), 1327–1335.

Cobb, Barry, Rafael Rumí, and Antonio Salmerón, "Approximating the Distribution of a Sum of Log-normal Random Variables," *Proceedings of the 6th European Workshop on Probabilistic Graphical Models, PGM 2012*, 09 2012.

Cohen, Alma and Liran Einav, "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, June 2007, *97* (3), 745–788.

Cox, Robynn, Seva Rodnyansky, Benjamin Henwood, and Suzanne L. Wenzel, "Measuring Population Estimates of Housing Insecurity in the United States: A Comprehensive Approach," *SSRN Electronic Journal*, 2017.

Cutts, Diana Becker, Alan F. Meyers, Maureen M. Black, Patrick H. Casey, Mariana Chilton, John T. Cook, Joni Geppert, Stephanie Ettinger De Cuba, Timothy Heeren, Sharon Coleman, Ruth Rose-Jacobs, and Deborah A. Frank, "US housing insecurity and the health of very young children," *American Journal of Public Health*, 2011, *101* (8), 1508–1514.

Dalton, Christina M, Gautam Gowrisankaran, and Robert Town, "Salience, Myopia, and Complex Dynamic Incentives: Evidence from Medicare Part D," Working Paper 21104, National Bureau of Economic Research April 2015.

David, Guy and Kunhee Lucy Kim, "The Effect of Workforce Assignment on Performance: Evidence from Home Health Care," *Journal of Health Economics*, 2018, *59*, 26–45.

Dehmer, Steven P., Michael V. Maciosek, Amy B. LaFrance, and Thomas J. Flottemesch, "Health benefits and cost-effectiveness of asymptomatic screening for hypertension and high cholesterol and aspirin counseling for primary prevention," *Annals of Family Medicine*, 2017, *15* (1), 23–36.

Dixit, Avinash K. and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *The American Economic Review*, 1977, *67* (3), 297–308.

Dubin, Jeffrey A. and Daniel L. McFadden, "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica*, 1984, *52* (2), 345–362.

Einav, Liran, Amy Finkelstein, and Jonathan Levin, "Beyond Testing: Empirical Models of Insurance Markets," *Annual Review of Economics*, 2010, *2* (1), 311–336.

_ , _ , Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen, "Selection on moral hazard in health insurance," *American Economic Review*, 2013, *103* (1), 178–219.

_ , Amy N. Finkelstein, and Mark Cullen, "Estimating Welfare in Insurance Markets Using Variation in Prices," *The Quarterly Journal of Economics*, 2010, *CXV* (3).

_ and Amy Finkelstein, "Moral Hazard in Health Insurance: What We Know and How We Know It," *Journal of the European Economic Association*, 2018, *16* (4), 957–982.

Ericson, Keith Marzilli and Justin R Sydnor, "Liquidity Constraints and the Value of Insurance," Working Paper 24993, National Bureau of Economic Research September 2018.

_ and Justin Sydnor, "The Questionable Value of Having a Choice of Levels of Health Insurance Coverage," *Journal of Economic Perspectives*, November 2017, *31* (4), 51–72.

Fadlon, Itzik and Jessica Van Parys, "Primary Care Physician Practice Styles and Patient Care: Evidence from Physician Exits in Medicare," *NBER Working Paper 26269*, 2019.

Fenton, L. F., "The sum of log-normal probability distributions in scatter transmission systems," *IRE Transactions on Communication Systems*, 1960, *8*, 57–67.

Finkelstein, Amy and Kathleen McGarry, "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *American Economic Review*, September 2006, *96* (4), 938–958.

_ , Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *Quarterly Journal of Economics*, 2016, (November), 1681–1726.

Finkelstein, Amy N., Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan H. Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group, "The Oregon Health Insurance Experiment: Evidence from the First Year," *The Quarterly Journal of Economics*, 2012, *127* (3), 1057–1106.

Friedberg, Mark W., Peter S. Hussey, and Eric C. Schneider, "Primary care: a critical review of the evidence on quality and costs of health care.," *Health Affairs*, 2010, *29* (5), 766–772.

Geruso, Michael, "Demand heterogeneity in insurance markets: Implications for equity and efficiency," *Quantitative Economics*, 2017, *8* (3), 929–975.

Goldin, Claudia and Lawrence F. Katz, "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation," *Journal of Labor Economics*, 2016, *34* (3), 705–746.

Goldman, Dana P, James P Smith, and Neeraj Sood, "Immigrants And The Cost Of Medical Care," *Health Affairs*, 2006, *6* (6), 1700–1711.

Goldstein, Amy, "New York City Mayor Vows Health Care for all — Including Undocumented Immigrants," jan 2019.

Goodman-Bacon, Andrew, "Public insurance and mortality: Evidence from medicaid implementation," *Journal of Political Economy*, 2018, *126* (1), 216–262.

Gross, Tal and Matthew J. Notowidigdo, "Health insurance and the consumer bankruptcy decision: Evidence from expansions of Medicaid," *Journal of Public Economics*, 2011, *95* (7), 767 – 778.

Handel, Benjamin, Igal Hendel, and Michael Whinston, "Equilibria in Health Exchanges : Adverse Selection vs . Reclassification Risk," *Econometrica*, 2015, *83* (4), 1261–1313.

Handel, Benjamin R, "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, 2013, *103* (7), 2643–2682.

Handel, Benjamin R. and Jonathan T. Kolstad, "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, August 2015, *105* (8), 2449–2500.

Handel, Benjamin R, Igal Hendel, and Michael D Whinston, "The Welfare Effects of Long-Term Health Insurance Contracts," Working Paper 23624, National Bureau of Economic Research July 2017.

Hausman, Naomi and Kurt Lavetti, "Physician Concentration and Negotiated Prices: Evidence from State Law Changes," *Working Paper*, 2016.

Ho, Kate and Robin Lee, "Health Insurance Menu Design: Managing the Spending Coverage Tradeoff," in "in" Presentation. Conference Celebrating the Scholarly Career of Mark Satterthwaite Kellogg School of Management 2019.

Iacus, Stefano M., Gary King, and Giuseppe Porro, "Causal inference without balance checking: Coarsened exact matching," *Political Analysis*, 2012, *20* (1), 1–24.

Jäger, Simon, "How Substitutable Are Workers? Evidence from Worker Deaths," *Ssrn*, 2017.

Jaravel, Xavier, Neviana Petkova, and Alex Bell, "Team-Specific Capital and Innovation," *American Economic Review*, 2018, (108(4-5)), 1034–1073.

Johnson, Erin, M. Marit Rehavi, David C. Chan, and Daniela Carusi, "A Doctor Will See You Now: Physician-Patient Relationships and Clinical Decisions," *NBER Working Paper 22666*, 2016.

Johnston, Kenton J., Lindsay Allen, Taylor A. Melanson, and Stephen R. Pitts, "A "Patch" to the NYU Emergency Department Visit Algorithm," *Health Services Research*, 2017, *52* (4), 1264–1276.

Jovanovic, Boyan, "Firm-specific Capital and Turnover," *Journal of Political Economy*, 1979, *87* (6), 1246–1260.

_ , "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 1979, *87* (5), 972–990.

Kahn, Richard, Peter Alperin, David Eddy, Knut Borch-Johnsen, John Buse, Justin Feigelman, Edward Gregg, Rury R. Holman, M. Sue Kirkman, Michael Stern, Jaakko Tuomilehto, and Nick J. Wareham, "Age at initiation and frequency of screening to detect type 2 diabetes: a cost-effectiveness analysis," *The Lancet*, 2010, *375* (9723), 1365–1374.

King, Gary and Richard Nielsen, "Why Propensity Scores Should Not Be Used for Matching," *Political Analysis*, 2016, pp. 1–20.

Kwok, Jennifer H, "How Do Primary Care Physicians Influence Healthcare ? Evidence on Practice Styles and Switching Costs from Medicare," 2019.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *The American Economic Review*, 1987, *77* (3), 251–277.

McManus, Margaret A., Stephen Berman, Thomas McInerny, and Suk fong Tang, "Weighing the Risks of Consumer-Driven Health Plans for Families," *Pediatrics*, 2006, *117* (4), 1420–1424.

McWilliams, J. Michael, Bruce E. Landon, Vinay K. Rathi, and Michael E. Chernew, "Getting more savings from ACOs - Can the pace be pushed?," *New England Journal of Medicine*, 2019, *380* (23), 2190–2192.

_ , Laura A. Hatfield, Michael E. Chernew, Bruce E. Landon, and Aaron L. Schwartz, "Early performance of accountable care organizations in medicare," *New England Journal of Medicine*, 2016, *374* (24), 2357–2366.

Miller, Sarah, "The effect of insurance on emergency room visits: An analysis of the 2006 Massachusetts health reform," *Journal of Public Economics*, dec 2012, *96* (11-12), 893–908.

Miranda, Mario J. and Paul L. Fackler, *Applied Computational Economics and Finance*, MIT Press, 2002.

Newhouse, Joseph, "Free for All? Lessons from the RAND Health Insurance Experiment," *Cambridge, MA. Harvard University Press*, 1993.

Passel, Jeffrey and D'Vera Cohn, "20 metro areas are home to six in ten unauthorized immigrants in U.S.," *Pew Research*, 2017, pp. 23–26.

Pauly, Mark V., "The Economics of Moral Hazard: Comment," *American Economic Review*, 1968, *58* (3), 531–537.

Pauly, Mark V, "Overinsurance and Public Provision of Insurance : The Roles of Moral Hazard and Adverse Selection," *The Quaterly Journal of Economics*, 1974, *88* (1), 44–62.

Pham, Hoangmai H., Deborah Schrag, Ann S. O'Malley, Beny Wu, and Peter B. Bach, "Care Patterns in Medicare and Their Implications for Pay for Performance," *New England Journal of Medicine*, 2007, *356* (11), 1130–1139.

Pollack, Craig E., Peter S. Hussey, Robert S. Rudin, D. Steven Fox, Julie Lai, and Eric C. Schneider, "Measuring care continuity : A comparison of claims-based methods," *Medical Care*, 2016, *54* (5), e30–e34.

Reed, Mary, Vicki Fung, Mary Price, Richard Brand, Nancy Benedetti, Stephen F. Derose, Joseph P. Newhouse, and John Hsu, "High-Deductible Health Insurance Plans: Efforts To Sharpen A Blunt Instrument," *Health Affairs*, 2009, *28* (4), 1145–1154. PMID: 19597214.

Reinhard, Susan C, Lynn Friss Feinberg, Ari Houser, Rita Choula, and Molly Evans, "Valuing the Invaluable: 2019 Update," *AARP Public Policy Institute*, 2019, (November).

Revelt, David and Kenneth Train, "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level," *The Review of Economics and Statistics*, 1998, *80* (4), 647–657.

_ and _ , "Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier," January 2001, (0012001).

Rodríguez, Michael A, Arturo Vargas Bustamante, and Alfonso Ang, "Perceived Quality of Care, Receipt of Preventive Care, and Usual Source of Health Care Among Undocumented and Other Latinos," *Journal of General Internal Medicine*, 2009, *24* (Suppl 3), 508–513.

Rothschild, Michael and Joseph Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Source: The Quarterly Journal of Economics*, 1976, *90* (4), 629–649.

Sabety, Adrienne, "Association of Primary Care Physician Turnover with Utilization and Outcomes Among Medicare Patients: Observational Study," *Working Paper*, 2020.

Sarsons, Heather, "Interpreting Signals in the Labor Market: Evidence from Medical Referrals," *Working Paper*, 2017, pp. 1–71.

Saultz, John W. and Waleed Albedaiwi, "Interpersonal continuity of care and patient satisfaction: A critical review," *Annals of Family Medicine*, 2005, *2* (5), 445–451.

Sessums, Laura L., Sanjay Basu, and Bruce E. Landon, "Primary Care First — Is It a Step Back?," *The New England journal of medicine*, 2019, pp. 1–3.

Sinaiko, Anna D., Mary Beth Landrum, David J. Meyers, Shehnaz Alidina, Daniel D. Maeng, Mark W. Friedberg, Lisa M. Kern, Alison M. Edwards, Signe Peterson Flieger, Patricia R. Houck, Pamela Peele, Robert J. Reid, Katharine McGraves-Lloyd, Karl Finison, and Meredith B. Rosenthal, "Synthesis Of Research On Patient-Centered Medical Homes Brings Systematic Differences Into Relief," *Health Affairs*, 2017, *36* (3), 500–508.

Slichter, S.H., D. Appleton, 1919.

Sommers, Benjamin D. and Kosali Simon, "Health Insurance and Emergency Department Use: A Complex Relationship," *New England Journal of Medicine*, 2017, *376* (18), 1708–1711.

Sommers, Benjamin D and Wendy E Parmet, "Health Care for Immigrants — Implications of Obama's Executive Action," 2015, pp. 1187–1189.

_ , Robert J Blendon, and E John Orav, "Both The 'Private Option' And Traditional Medicaid Expansions Improved Access To Care For Low-Income Adults," *Health Affairs*, 2016.

Starfield, Barbara, "Is primary care essential?," *The Lancet*, 1994, *344* (8930), 1129–1133.

Stole, Lars A and Jeffrey Zwiebel, "Intra-firm Bargaining under Non-binding Contracts," 1996, *63* (3), 375–410.

Tarraf, Wassim, William Vega, and Hector M Gonza, "Emergency Department Services Use Among Immigrant and Non-immigrant Groups in the United States," 2014, pp. 595–606.

Tolbert, Jennifer, Kendal Orgera, Natalie Singer, and Anthony Damico, "Key Facts about the Uninsured Population," 2019.

Train, Kenneth, *Discrete Choice Methods with Simulation: Second Edition*, Cambridge University Press, 2009.

Watson, Tara, "Inside the Refrigerator: Immigration Enforcement and Chilling Effects in Medicaid Participation," *American Economic Journal: Economic Policy*, 2014, *6* (3), 313–338.

Whitcomb, Michael E. and Jordan J. Cohen, "The Future of Primary Care Medicine," *New England Journal of Medicine*, 2004, *351* (7), 710–712.

Zeckhauser, Richard, "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, 1970, *2* (1), 10–26.

Appendix A

Supplementary Material for Chapter 1

This supplement contains proofs and additional results for the paper "Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends." Section **??** provides proofs for the results in the main text. Section **??** introduces corrections to parametric approaches that have good properties conditional on surviving a test for pre-trends. Section **??** states and proves asymptotic results. Section **??** provides additional simulation results in which the treatment and control group receive stochastic common shocks. Finally, Section **??** contains additional figures.

A.1 Additional Data Details

Table A.1: *Prioritization of Data Sources*

Outcome	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}
NP, PA, or MD/DO	NPPES	MD-PPAS	PC	Doximity	Claims
Physician Specialty	Doximity	MD-PPAS	Claims		
Age	MD-PPAS	Doximity			
Gender	Modal Response				

Notes: The National Plan and Provider Enumeration System (NPPES), Doximity, Medicare's MD-PPAS, Physician Compare (PC), and Medicare's Part B Carrier file (Claims) were combined and prioritized based on reliability to determine the type of provider, specialty, sub-specialty age and gender based on health providers' national provider identifiers (NPI). MD/DO, PA, and NP stands for medical doctors/doctors of osteopathy, physician assistants, and nurse practitioners, respectively. "Modal Response" means that the modal gender across all five data sources was considered the correct sex.

A.1.1 Additional Definitions

Each PCP is assigned a unique identifier combining their NPI and associated clinic identifier (NPI-TIN-ZIP). NPIs belonging to organizations, and not individuals, are dropped from the sample because it is not possible to observe individual provider exits.

Clinic Closures. I define clinic closures to occur when a PCP departs and (i) the TIN is the last TIN observed at the ZIP,[1] (ii) all PCPs at the clinic completely disappear from the data, or (iii) the TIN disappears from the data in month $t+1$ and the number of NPIs at that nine digit zip decreases by the exact number of NPIs affiliated with that clinic as of month t.

Utilization. "Office settings" are tagged using place of service codes equal to 11 in the carrier file.Medicare Provider Analysis and Review files identify hospitalizations. I do not use years 2002-2007 to define departures because in 2008 provider identifiers switch from UPINs to NPIs. The crosswalk between UPINs and NPIs is imperfect, so to avoid misclassifying departures I focus on later years. See Section 1.4.1 for how I identified whether a provider's NPI belonged to a PCP or specialist.

Total spending follows recent literature and aggregates patients' carrier, inpatient, outpatient, urgent care, and ED charges (Finkelstein, Gentzkow and Williams, 2016). Out-of-pocket costs aggregate the coinsurance and deductibles paid by patients for these services. Prescription drug claims are obtained from the Part D Event and Plan Characteristic files. Medicare Beneficiary Summary files provide patients' date of death, demographic characteristics, and enrollment information. All files are linked using beneficiary identification numbers and claim dates.

Medicare Advantage Patients. Patients were coded as being in Medicare Advantage (MA) according to the Master Beneficiary Summary file.

Medications. The number of medications as well as chronic medications are classified. Medications are counted based on filled prescriptions, so prescriptions that are written

[1]This would not include clinics that switched TINs due to an ownership change because I would observe a new clinic at that ZIP in that case.

but not filled are missed. The chronic medication category includes Selective Serotonin Reuptake Inhibitors (SSRIs), antihypertensives, antidiabetics, and statins. In order to derive the classification, prescriptions are aggregated into categories and classes using a crosswalk between RedBook data and the generic names from Medicare's Part D plan characteristics file.

Statins were classified as antihyperlipidemic drugs. Opioids were classified as opiate agonists, opiate part agonists, and opiate antagonists. Antihypertensives include NEC cardiac drugs (e.g. losartan and olmesartan), ACE inhibitors, alpha-beta blockers, beta blockers, and calcium channel blockers. Antidiabetics include insulins, sulfonylureas, and other antidiabetic agents. Supplies used by diabetics, such as lancets and blood sugar diagnostic materials, were also included. Antidepressants included prescriptions like fluoxetine, escitalopram, and sertraline. Benzodiazepines included prescriptions like lorazepam, alprazolam, and diazepam. NPIs on prescription scripts were used to identify prescribers. Whether a prescriber was a PCP or specialist was determined from merging in the NPI data set described in 1.4.1.

Patient Risk Score and High Risk Patients. Elixhauser scores were used to create a risk index based on patients' entire set of International Classification of Disease 9th and 10th edition (ICD-9/10) diagnosis codes from the carrier file. The Elixhauser Index scores patients based on commorbities and pre-existing conditions that are predictive of death. Patients' Elixhauser scores were derived using the Stata function "Elixhauser."[2] To define high risk patients, yearly risk scores were derived based on all diagnosis codes recorded over the year. Patients who had no claims in a specific year were given a risk score of zero. The score was used to stratify the population into low and high risk patients, within a PCP's pool of patients.[3] The top quartile of scores were defined to be high risk, the bottom $\frac{3}{4}$ were defined to be not high risk.

---

[2]Vicki Stagg, 2015. "Elixhauser: Stata module to calculate Elixhauser index of comorbidity," Statistical Software Components S458077, Boston College Department of Economics.

[3]Elixhauser, Anne, Claudia Steiner, D. Robert Harris, and Rosanna M. Coffey. "Comorbidity measures for use with administrative data." Medical care (1998): 8-27.

Clinic Rural or Urban. A clinic's zip code was determined to be urban or rural using the Center for Medicare and Medicaid Services 2019 fee schedule.

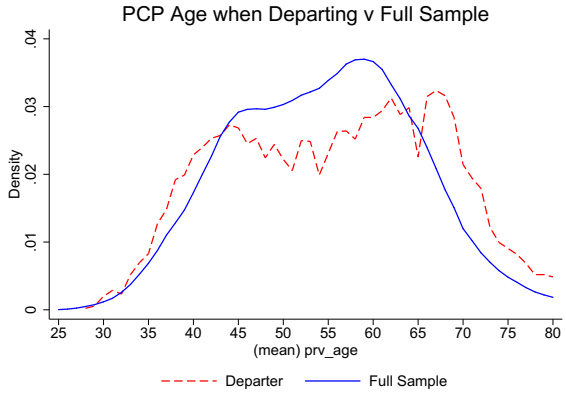
Additional Patient Sample Restrictions. Patients who switched in and out of being enrolled in Medicare Advantage (MA) were dropped from the sample to avoid missing data on the patient while they were in MA. All remaining patients who enrolled in MA from $-36 \leq t < -24$, or over the assignment period, were dropped from the main sample. They were only dropped once it was determined that patients do not differentially switch to MA at the threshold. This is checked and quantified in Table A.16. Patients that died over the assignment period were also dropped from the sample. This was done to precisely define a PCP's pool of patients from $-36 \leq t < -24$.

A.1.2 Additional Description of Departures

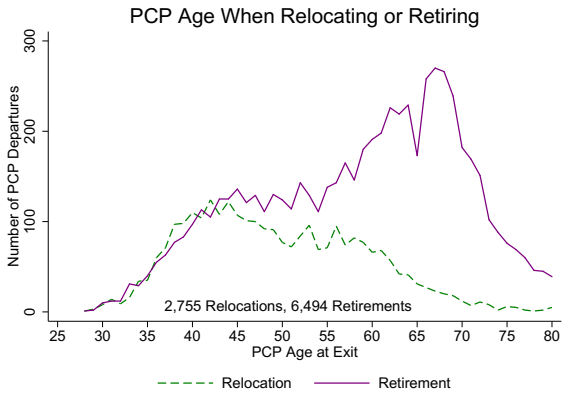
Table A.2: *Breakdown of Departures*

2011	2,673	*2.32%*
2012	3,075	*2.63%*
2013	3,374	*2.87%*
2014	3,653	*3.1%*
Total Departures:	12,761	*9%*

Figure A.1: *Histograms of PCP Age*



(a) *Departing versus the Full Sample, by PCP Age*



(b) *PCP Age When Moving versus Full Departures*

Table A.3: *Match Attrition for Each Matching Variable*

Dropping	Treatment Match Rate
–	87%
PCP Age	89%
Female	88%
Patients Seen at t=-24	92%
NP, PA, or MD	94%

A.2 Clinic Level Sample

Treated Clinics: The same departure definition was used as in Section 1.4.2. Treatment clinics are defined to be clinics that (i) existed for 49 months, (ii) see an average of $\geq$2 patients per month over the period, and (iii) see $\geq$30 patients from $-36 \leq t \leq -24$. This algorithm drops clinic closures and solo clinics. Further, clinics with $>$100 PCPs were excluded.

Control Clinics: If a PCP departed a clinic but the conditions laid out in Section 1.4.2 were not met, the PCP would be considered a control and, as such, the PCP's clinic would be called a "control clinic." Further, control clinics had to (i) exist for 49 months, (ii)

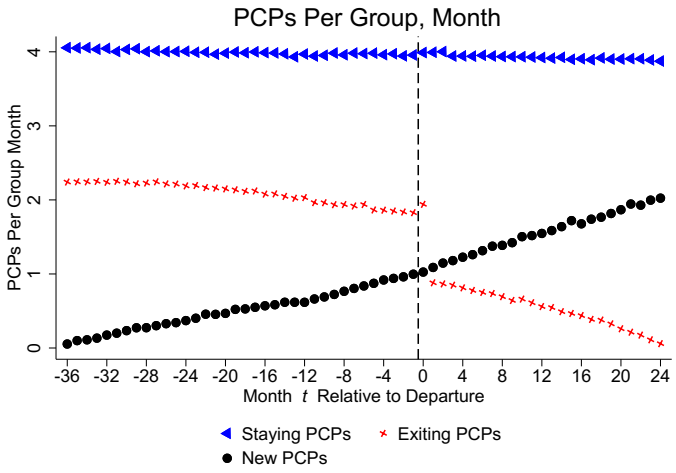
see an average of $\geq$2 patients per month over the period, and (iii) see $\geq$30 patients from $-36 \leq t \leq -24$.

Matching: Three coarsened bins of average PCP age at the clinic, seven coarsened bins of the number of PCPs per practice, nine coarsened bins of the number of patients seen in $t = -36$, whether or not the clinic was on an individual or shared model, and month and year of calendar time were matched on. Balance for clinics is illustrated in Table A.4 and balance for indirectly affected patients is shown in Table A.5

Indirectly Affected Patients: I call a patient indirectly affected if they *never* saw a strongly departing PCP. I assign patients to clinics in $-36 \leq t \leq -24$ and then I assign patients to PCPs within the clinic based on their modal PCP.

Additional Sample Restrictions: Clinics where the total number of PCPs changed by more than 2 standard deviations in a given year were dropped. This additional restriction was imposed on the clinic sample because being able to accurately draw clinic barriers is more important for clinic level analyses, opposed to PCP level analyses. This drops 30% of treated clinics. Analyses without this restriction were run for robustness and showed that the number of visits at non-focal clinics decreased in response to a focal-PCP departure, which is unintuitive and suggests that clinic boundaries were likely incorrect in this case.

Figure A.2: *PCPs per Group Over Time*



Notes: Graphs rely on a sample that matches *clinics*, instead of PCPs, which is described in Appendix A.2. Figure 1.9b shows the number of PCPs per group over time. The average number of exiting PCPs does not go to zero at $t = 1$, instead showing that there are subsequent departures at the group. PCPs that exit after the focal PCP departure in $t = 0$ gradually exit over the two year post period.

Table A.4: *Balance Table for Clinics and Patients Following Groups 2 Years post-departure*

	Treatment	Control	P-Value
% Match	82%	24%	
Number of Clinics	1,579	1,579	
Number of Patients	225,625	213,951	
Clinic Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	50.3	49.9	0.16
Caseload per Clinic/Month	54.0	49.9	0.04
PCPs Per Practice	5.2	4.6	p< 0.001
Patient Covariates That Were Not Matched On, 3 Years Prior to Exit			
Patient Demographics, 3 Years Prior to Exit			
Patient Age (yr)	71.1	71.5	0.02
White (%)	85.6	87.2	0.02
Urban (%)	76.6	81.2	0.001
Patient Clinical Characteristics, 3 Years Prior to Exit			
Elixhauser Risk Score	2.5	2.5	0.51
Also Enrolled in Medicaid (%)	19.6	16.8	p< 0.001
Average Annual Rate per Patient, 36-24 Months Before Exit			
No. of Primary Care Visits	6.1	6.1	0.84
No. of Specialist Visits	10.3	10.4	0.07
No. of Emergency Department Visits	0.87	0.84	p< 0.001
No. of Inpatient Visits	0.42	0.43	0.005
Annual Spending ($)	8728.85	8905.83	p< 0.001
Prob. of Death (%)	2.8	2.9	p< 0.001
No. of Visits with Departing PCP	8.3	–	p< 0.001
No. of Visits at Clinic	23.3	22.9	0.02

Notes: See Table 1.1 for description of overlapping outcomes. PCPs per practice categorizes the total number of PCPs at the clinic. Share of NPs/PAs (females) are the number of NPs/PAs (females) per clinic over the total number of PCPs at the clinic.

Table A.5: *Balance Table for Clinics and Patients Indirectly Affected by Departure*

	Treatment	Control	P-Value
% Match	82%	24%	
Number of Clinics	1,558	1,558	
Number of Patients	137,047	212,605	
Clinic Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	50.2	49.9	0.20
Caseload per Clinic/Month	54.4	50.3	0.04
PCPs Per Practice	5.2	4.6	$p< 0.001$
Clinic Covariates That Were Not Matched On, 3 Years Prior to Exit			
Share of NPs/PAs	0.17	0.13	$p< 0.001$
Share Female	0.38	0.36	0.02
Avg Pop. in Zip	25623	26645	0.07
Patient Covariates That Were Not Matched On			
Patient Demographics, 3 Years Prior to Exit			
Patient Age (yr)	71.0	71.5	0.003
White (%)	85.9	87.2	0.05
Urban (%)	76.5	81.1	0.002
Patient Clinical Characteristics, 3 Years Prior to Exit			
Elixhauser Risk Score	2.4	2.5	0.001
Also Enrolled in Medicaid (%)	19.1	16.8	$p< 0.001$
Average Annual Rate per Patient, 36-24 Months Before Exit			
No. of Primary Care Visits	5.4	6.1	$p< 0.001$
No. of Specialist Visits	9.9	10.4	$p< 0.001$
No. of Emergency Department Visits	0.82	0.84	$p< 0.001$
No. of Inpatient Visits	0.40	0.43	$p< 0.001$
Annual Spending ($)	8410.41	8893.87	$p< 0.001$
Prob. of Death (%)	3.0	2.9	0.005
No. of Visits with Departing PCP	–	–	–
No. of Visits at Clinic	18.3	22.9	$p< 0.001$

Notes: See Table 1.1 for description of overlapping outcomes. PCPs per practice categorizes the total number of PCPs at the clinic. Share of NPs/PAs (females) are the number of NPs/PAs (females) per clinic over the total number of PCPs at the clinic.

A.3 Balance Tables

Table A.6: *Balance Table for Patients and PCPs, with Practice Size Match*

	Treatment	Control	P-Value
% Match	71%	23%	
No. of PCPs	8,307	8,307	
No. of Patients	237,273	261,242	
No. of Observations	23,003,605	23,003,605	
PCP Matching Covariates 3 Years Prior to PCP Exit			
PCP Age (yr)	53.8	53.0	p< 0.001
Caseload per PCP/Month	11.9	11.9	1.00
Patient Covariates That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	71.1	71.8	p< 0.001
White (%)	84.9	84.5	0.32
Female (%)	37.1	36.2	0.002
Urban (%)	79.5	84.9	p< 0.001
Patient Clinical Characteristics			
Elixhauser Risk Score	2.5	2.5	0.30
End Stage Renal Disease (%)	0.93	1.0	0.06
Also Enrolled in Medicaid (%)	20.0	17.4	p< 0.001
Average Annual Rate per Patient, 36-24 Months Before Exit			
No. of Primary Care Visits	17.1	17.5	p< 0.001
No. of PCPs Seen	5.5	4.4	p< 0.001
No. of Specialty Care Visits	29.4	29.8	p< 0.001
No. of Specialists Seen	11.9	10.3	p< 0.001
No. of Emergency Department Visits	2.3	2.0	p< 0.001
No. of Urgent Care Visits	0.056	0.046	p< 0.001
No. of Inpatient Department Visits	1.3	1.3	p< 0.001
Annual Spending ($)	28396.43	28421.79	0.53
Prob. of Death (%)	9.9	9.8	p< 0.001

Notes: See Table 1.1 for description of outcomes. Values are as of 3 years before a PCP's exit.

Table A.7: *Balance Table for Treated PCPs and Patients by Whether Clinic Closed or Remained Open*

	Only Treated PCPs and Patients		
	Closed	Open	P-Value
PCP Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	60.3	51.4	p< 0.001
Caseload per PCP/Month	13.1	11.8	p< 0.001
PCP Covariates That Were Not Matched On			
PCPs Per Practice	1.5	12.7	p< 0.001
Avg Pop. in Zip	27782	25804	p< 0.001
Median Income in Zip ($)	54270.84	52774.15	0.002
Panel Model	0.69	0.24	p< 0.001
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	71.8	70.8	p< 0.001
PCP-Patient Bond (yr)	6.5	5.2	p< 0.001
White (%)	82.1	86.9	p< 0.001
Female (%)	38.2	36.5	p< 0.001
Urban (%)	80.1	78.8	0.15
Patient Clinical Characteristics			
Elixhauser Score	2.6	2.5	p< 0.001
End Stage Renal Disease (%)	0.93	0.91	0.74
Also Enrolled in Medicaid (%)	20.6	19.0	p< 0.001
Average Annual Rate per Patient, 36-24 Months Before Exit			
No. of Primary Care Visits	6.3	6.0	p< 0.001
No. of PCPs Seen	1.0	1.8	p< 0.001
No. of Specialty Care Visits	9.9	9.6	p< 0.001
No. of Specialists Seen	2.6	3.6	p< 0.001
No. of Emergency Department Visits	0.67	0.71	p< 0.001
No. of Urgent Care Visits	0.016	0.019	p< 0.001
No. of Inpatient Department Visits	0.37	0.39	p< 0.001
Annual Spending ($)	8687.84	8338.57	p< 0.001
Prob. of Death (%)	0.0036	0.0031	p< 0.001

Notes: See Table 1.1 for description of outcomes. Values are as of 3 years before a PCP's exit.

Table A.8: *Balance Table for Control Clinics by Individual or Shared*

	Only Treated PCPs and Patients		
	Panel	Shared	P-Value
Clinic Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	56.4	52.4	p< 0.001
Caseload per PCP/Month	13.9	11.2	p< 0.001
Clinic Covariates That Were Not Matched On			
Share of Practice that is Female	0.28	0.42	p< 0.001
Share of Practice that are APs	0.073	0.18	p< 0.001
PCPs Per Practice	4.4	12.6	p< 0.001
Avg Pop. in Zip	26791	26091	0.04
Median Income in Zip ($)	52753.43	53425.48	0.14
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	71.5	70.8	p< 0.001
PCP-Patient Bond (yr)	6.0	5.3	p< 0.001
White (%)	83.3	86.9	p< 0.001
Female (%)	39.1	35.7	p< 0.001
Urban (%)	77.8	80.0	0.007
Patient Clinical Characteristics			
Elixhauser Score	2.7	2.4	p< 0.001
End Stage Renal Disease (%)	0.90	0.92	0.78
Also Enrolled in Medicaid (%)	21.3	18.4	p< 0.001
Average Annual Rate per Patient, 36-24 Months Before Exit			
No. of Primary Care Visits	6.8	5.6	p< 0.001
No. of PCPs Seen	1.1	1.8	p< 0.001
No. of Specialty Care Visits	9.7	9.6	p< 0.001
No. of Specialists Seen	2.6	3.8	p< 0.001
No. of Emergency Department Visits	0.71	0.69	p< 0.001
No. of Urgent Care Visits	0.015	0.020	p< 0.001
No. of Inpatient Department Visits	0.40	0.38	p< 0.001
Annual Spending ($)	8914.25	8162.86	p< 0.001
Prob. of Death (%)	0.0037	0.0030	p< 0.001

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' practice locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending included provider charges from the carrier file, inpatient charges, and outpatient charges.

Table A.9: *Balance Table for Control PCPs and Patients by Local Density of PCPs*

	Only Treated PCPs and Patients		
	High PCP Density	Low PCP Density	P-Value
PCP Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	53.9	53.6	0.47
Caseload per PCP/Month	12.3	11.4	p< 0.001
PCP Covariates That Were Not Matched On			
PCPs Per Practice	9.7	9.0	0.02
Avg Pop. in Zip	30166	9379	p< 0.001
Median Income in Zip ($)	53247.01	52898.79	0.54
Panel Model	0.36	0.37	0.44
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	71.1	71.0	0.36
PCP-Patient Bond (yr)	5.6	5.6	0.58
White (%)	85.5	86.0	0.37
Female (%)	36.8	37.5	0.14
Urban (%)	80.0	75.4	p< 0.001
Patient Clinical Characteristics			
Elixhauser Score	2.5	2.4	p< 0.001
End Stage Renal Disease (%)	0.90	1.0	0.41
Also Enrolled in Medicaid (%)	19.7	18.4	0.02
Average Annual Rate per Patient, 2-3 Years Before Exit			
No. of Primary Care Visits	6.0	6.1	0.01
No. of PCPs Seen	1.5	1.7	p< 0.001
No. of Specialty Care Visits	9.8	9.2	p< 0.001
No. of Specialists Seen	3.4	3.4	0.75
No. of Emergency Department Visits	0.70	0.69	0.17
No. of Urgent Care Visits	0.017	0.024	p< 0.001
No. of Inpatient Department Visits	0.39	0.38	0.05
Annual Spending ($)	8485.75	8204.19	p< 0.001
Prob. of Death (%)	0.0033	0.0031	0.04

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' practice locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending included provider charges from the carrier file, inpatient charges, and outpatient charges.

Table A.10: *Balance Table for Control PCPs and Patients by Length of Relationship*

	Only Treated PCPs and Patients		
	4-13 Yrs	3 Yrs	P-Value
PCP Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	57.0	54.8	p< 0.001
Caseload per PCP/Month	15.3	16.2	p< 0.001
PCP Covariates That Were Not Matched On			
PCPs Per Practice	2.3	2.4	0.03
Avg Pop. in Zip	28918	28615	0.38
Median Income in Zip ($)	56704.04	57155.58	0.34
Panel Model	0.62	0.60	0.02
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	81.8	81.7	0.48
PCP-Patient Bond (yr)	9.5	3.0	p< 0.001
Max SP Relationship Length (yr)	4.9	3.8	p< 0.001
Patient Clinical Characteristics			
Elixhauser Score	2.9	2.9	0.07
End Stage Renal Disease (%)	0.50	0.60	0.19
Also Enrolled in Medicaid (%)	15.5	16.4	0.10
Average Annual Rate per Patient, 2-3 Years Before Exit			
No. of Primary Care Visits	6.6	6.1	p< 0.001
No. of PCPs Seen	2.8	6.1	p< 0.001
No. of Specialty Care Visits	10.3	10.6	p< 0.001
No. of Specialists Seen	8.0	17.0	p< 0.001
No. of Emergency Department Visits	0.60	0.74	p< 0.001
No. of Inpatient Department Visits	0.43	0.55	p< 0.001
Annual Spending ($)	9380.53	11019.21	p< 0.001
Prob. of Death (%)	0.0057	0.0058	0.41

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' practice locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending included provider charges from the carrier file, inpatient charges, and outpatient charges.

Table A.11: *Balance Table for Control PCPs and Patients by Clinic Size*

	Only Treated PCPs and Patients		
	1-3 PCPs	4-100 PCPs	P-Value
PCP Matching Covariates 2 Years Prior to PCP Exit			
Avg PCP Age (yr)	58.0	51.4	p< 0.001
Caseload per PCP/Month	12.5	12.0	0.010
PCP Covariates That Were Not Matched On			
PCPs Per Practice	1.3	14.4	p< 0.001
Avg Pop. in Zip	27352	25763	p< 0.001
Median Income in Zip ($)	53646.50	52915.73	0.11
Panel Model	0.66	0.19	p< 0.001
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	71.3	71.0	0.005
PCP-Patient Bond (yr)	6.1	5.3	p< 0.001
White (%)	82.2	87.6	p< 0.001
Female (%)	38.0	36.4	p< 0.001
Urban (%)	79.3	79.1	0.83
Patient Clinical Characteristics			
Elixhauser Score	2.6	2.4	p< 0.001
End Stage Renal Disease (%)	1.0	0.88	0.12
Also Enrolled in Medicaid (%)	21.7	18.1	p< 0.001
Average Annual Rate per Patient, 2-3 Years Before Exit			
No. of Primary Care Visits	6.3	5.9	p< 0.001
No. of PCPs Seen	1.1	1.8	p< 0.001
No. of Specialty Care Visits	10.0	9.5	p< 0.001
No. of Specialists Seen	2.8	3.6	p< 0.001
No. of Emergency Department Visits	0.70	0.69	0.12
No. of Urgent Care Visits	0.016	0.020	p< 0.001
No. of Inpatient Department Visits	0.38	0.39	p< 0.001
Annual Spending ($)	8719.56	8269.32	p< 0.001
Prob. of Death (%)	0.0035	0.0031	p< 0.001

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' practice locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending included provider charges from the carrier file, inpatient charges, and outpatient charges.

Table A.12: *Balance Table for Control PCPs and Patients by Patient Risk Status*

	Only Treated PCPs and Patients		
	High Risk	Not High Risk	P-Value
PCP Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	53.4	53.4	0.95
Caseload per PCP/Month	12.3	12.3	0.94
PCP Covariates That Were Not Matched On			
PCPs Per Practice	2.6	2.5	0.78
Avg Pop. in Zip	28672	28711	0.87
Median Income in Zip ($)	56291.65	56412.63	0.72
Panel Model	0.57	0.57	0.93
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	73.3	70.9	$p< 0.001$
PCP-Patient Bond (yr)	6.2	5.8	$p< 0.001$
White (%)	81.3	81.9	0.11
Female (%)	37.8	36.7	$p< 0.001$
Urban (%)	83.8	84.0	0.82
Patient Clinical Characteristics			
Elixhauser Score	4.9	1.8	$p< 0.001$
End Stage Renal Disease (%)	2.5	0.61	$p< 0.001$
Also Enrolled in Medicaid (%)	22.6	19.9	$p< 0.001$
Average Annual Rate per Patient, 2-3 Years Before Exit			
No. of Primary Care Visits	8.1	5.5	$p< 0.001$
No. of PCPs Seen	4.1	1.4	$p< 0.001$
No. of Specialty Care Visits	15.1	8.1	$p< 0.001$
No. of Specialists Seen	12.9	3.0	$p< 0.001$
No. of Emergency Department Visits	1.1	0.52	$p< 0.001$
No. of Urgent Care Visits	0.015	0.013	$p< 0.001$
No. of Inpatient Department Visits	0.71	0.28	$p< 0.001$
Annual Spending ($)	14918.82	6643.82	$p< 0.001$
Prob. of Death (%)	0.0063	0.0026	$p< 0.001$

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' practice locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending included provider charges from the carrier file, inpatient charges, and outpatient charges.

Table A.13: *Balance Table for Control PCPs and Patients by Disability Status*

	Only Treated PCPs and Patients		
	Disabled	Not Disabled	P-Value
PCP Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	53.6	53.4	0.32
Caseload per PCP/Month	13.0	12.3	p< 0.001
PCP Covariates That Were Not Matched On			
PCPs Per Practice	2.5	2.6	0.34
Avg Pop. in Zip	28511	28671	0.53
Median Income in Zip ($)	56351.25	56320.34	0.93
Panel Model	0.58	0.57	0.09
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	71.6	71.6	0.89
PCP-Patient Bond (yr)	6.3	5.7	p< 0.001
White (%)	82.3	81.5	0.05
Female (%)	31.6	39.3	p< 0.001
Urban (%)	83.8	83.8	0.96
Patient Clinical Characteristics			
Elixhauser Score	3.1	2.4	p< 0.001
End Stage Renal Disease (%)	1.4	1.0	p< 0.001
Also Enrolled in Medicaid (%)	25.1	18.6	p< 0.001
Average Annual Rate per Patient, 2-3 Years Before Exit			
No. of Primary Care Visits	7.4	5.6	p< 0.001
No. of PCPs Seen	3.9	1.5	p< 0.001
No. of Specialty Care Visits	13.4	8.6	p< 0.001
No. of Specialists Seen	11.7	3.5	p< 0.001
No. of Emergency Department Visits	1.1	0.47	p< 0.001
No. of Urgent Care Visits	0.020	0.011	p< 0.001
No. of Inpatient Department Visits	0.66	0.29	p< 0.001
Annual Spending ($)	13528.81	7001.06	p< 0.001
Prob. of Death (%)	0.0050	0.0030	p< 0.001

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' practice locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending included provider charges from the carrier file, inpatient charges, and outpatient charges.

Table A.14: *Balance Table for Control PCPs and Patients by Patient Race*

	Only Treated PCPs and Patients		
	Minority	White	P-Value
PCP Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	54.4	53.3	p< 0.001
Caseload per PCP/Month	14.9	12.3	p< 0.001
PCP Covariates That Were Not Matched On			
PCPs Per Practice	2.4	2.6	p< 0.001
Avg Pop. in Zip	30275	28402	p< 0.001
Median Income in Zip ($)	56615.08	56519.93	0.82
Panel Model	0.63	0.57	p< 0.001
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	69.0	71.8	p< 0.001
PCP-Patient Bond (yr)	5.8	5.9	0.22
White (%)	–	100.0	p< 0.001
Female (%)	37.9	37.5	0.33
Urban (%)	87.1	83.6	p< 0.001
Patient Clinical Characteristics			
Elixhauser Score	2.8	2.6	p< 0.001
End Stage Renal Disease (%)	2.9	0.67	p< 0.001
Also Enrolled in Medicaid (%)	35.9	17.1	p< 0.001
Average Annual Rate per Patient, 2-3 Years Before Exit			
No. of Primary Care Visits	6.0	6.2	p< 0.001
No. of PCPs Seen	5.7	1.6	p< 0.001
No. of Specialty Care Visits	8.6	10.2	p< 0.001
No. of Specialists Seen	13.7	3.9	p< 0.001
No. of Emergency Department Visits	0.78	0.64	p< 0.001
No. of Urgent Care Visits	0.010	0.014	p< 0.001
No. of Inpatient Department Visits	0.41	0.40	p< 0.001
Annual Spending ($)	9122.15	8871.90	p< 0.001
Prob. of Death (%)	0.0026	0.0036	p< 0.001

Notes: P-values were estimated with the use of two-sample Student's t-tests. NPIs identified PCPs and patients were assigned to PCPs based on their modal number of pre-period evaluation and management visits. The PCP sample included high volume PCPs, control PCPs who did not practice with exiting PCPs, and PCPs practicing in clinics with fewer than 100 PCPs. PCP age was determined from a secondary data set encompassing the NPPES, Doximity, Medicare's MD-PPAS, and Physician Compare. In addition to the PCP level matching covariates reported, gender, type of PCP (i.e. MD/DO, PA, NP), and month year of calendar time were exactly matched on. The caseload per PCP/month captured the number of patients seen by each PCP for any type of visit, regardless of whether the PCP was assigned as the patient's PCP. Patient age, gender, and race were determined according to the Medicare Beneficiary Summary File. "Urban" refers to patients' practice locations, which were determined using the Center for Medicare and Medicaid Services rural/urban fee schedule. The Elixhauser Risk Index scores patients based on commorbities and pre-existing conditions that are predictive of death; scores range from 0 to 12, with higher scores indicating more coexisting conditions and that patients are of "higher risk." Whether a patient was also enrolled in the end stage renal disease program and/or Medicaid was determined according to the Medicare Beneficiary Summary File. Spending included provider charges from the carrier file, inpatient charges, and outpatient charges.

Table A.15: *Balance Table for Treated PCPs and Patients by Female Patients versus Male Patients Matched with Female PCPs*

	Only Treated PCPs and Patients		
	Male	Female	P-Value
PCP Matching Covariates 3 Years Prior to PCP Exit			
Avg PCP Age (yr)	48.1	48.1	1.00
Caseload per PCP/Month	9.9	9.9	1.00
PCP Covariates That Were Not Matched On			
PCPs Per Practice	2.9	2.9	1.00
Avg Pop. in Zip	27236	27236	1.00
Median Income in Zip ($)	55754.04	55754.04	1.00
Panel Model	0.53	0.53	1.00
Outcomes That Were Not Matched On			
Patient Demographics			
Patient Age (yr)	71.7	69.5	p< 0.001
PCP-Patient Bond (yr)	5.2	5.0	p< 0.001
White (%)	82.8	83.1	0.69
Female (%)	–	100.0	–
Urban (%)	79.6	79.6	1.00
Patient Clinical Characteristics			
Elixhauser Score	2.6	2.7	p< 0.001
End Stage Renal Disease (%)	0.81	1.3	p< 0.001
Also Enrolled in Medicaid (%)	22.4	20.5	0.007
Average Annual Rate per Patient, 36-24 Months Before Exit			
No. of Primary Care Visits	6.4	5.9	p< 0.001
No. of PCPs Seen	1.8	3.9	p< 0.001
No. of Specialty Care Visits	9.9	9.9	0.59
No. of Specialists Seen	4.1	9.7	p< 0.001
No. of Emergency Department Visits	0.64	0.67	p< 0.001
No. of Inpatient Department Visits	0.36	0.41	p< 0.001
Annual Spending ($)	8064.97	9447.20	p< 0.001
Prob. of Death (%)	0.0029	0.0042	p< 0.001

Notes: See Table 1.1 for description of outcomes. Values are as of 3 years before a PCP's exit.

A.4 Details of Health Valuation Calculation

A.4.1 Number of Cardiovascular Deaths Averted

Cholesterol:

$$0.039 * (1010/0.90) = 44 \text{ deaths per } 100,000$$

0.039 is the decrease in the number of cholesterol screens after a PCP's exit (Table A.20). According to Dehmer et al. (2017), 1,080 is the number of deaths averted through cholesterol screening (per 100,000) and 0.9 scales the estimate to the entire population (or a screening rate of 90%).

Diabetes:

$$0.016 * (300) = 5 \text{ deaths per } 100,000$$

0.016 is the decrease in the number of diabetes screens after a PCP's exit (Table A.20). According to Kahn et al. (2010), diabetes screens prevent 3 deaths per 1000 people for those over age 60, of 300 per 100,000.

A.4.2 Number of Influenza and Pneumonia Deaths Averted by Vaccination

Flu Vaccination:

$$.031 * ((5,637/.49)/44,574,166) = .8 \text{ deaths per } 100,000$$

where 0.031 percentage points represents the decrease in the probability a patient receives a flu vaccine after losing a PCP (Table 1.2), 5,637 is the number of older adult ($\geq$ 65) deaths averted due to flu vaccinations in 2015-2016,[4] .49 is the rate of flu vaccination among control patients at baseline (Table 1.2), and 44,574,166 scales the estimate per 100,000 adults in the United States.[5]

[4] https://www.cdc.gov/flu/about/burden-averted/2015-16.htm

[5] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_1YR_S0101&prodType=table;%20https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_1YR_B01001B&prodType=table

A.5 Additional Results

Table A.16: *Additional Outcomes*

Type	Mean	Impact	Type	Mean	Impact
Count of PCP Departures	1.3	0.50***	**Type of Visit**		
		(0.046)	Prob. of Any PCP EM Visit	3.2	-0.57***
		38.7%			(0.015)
Count of Strong Departures	0.061	0.083***			-17.8%
		(0.0078)	Prob. Visit Pre-Existing PCP	3.0	-0.73***
		136.0%			(0.014)
Prob. of Enrolling in MA	0.028	0.0019*			-24.1%
		(0.00092)	Prob. Form New PCP Relationship	0.14	0.17***
		6.9%			(0.0038)
Total Spending on ED and IP	5262.9	124.0**			115.9%
		(59.8)	Prob. of Any SP EM Visit	3.1	0.23***
		2.4%			(0.0093)
ED Classification Outcomes					7.4%
No. of ED Visits, Preventable	0.042	0.0016	Prob. Visit Pre-Existing SP	2.6	0.17***
		(0.0011)			(0.0083)
		3.8%			6.6%
No. of ED Visits, Not Preventable	0.19	0.011***	Prob. Form New SP Relationship	0.57	0.065***
		(0.0032)			(0.0033)
		5.7%			11.4%
No. of ED Visits, Primary Care Treatable	0.18	0.0048*	**Timing of Visits**		
		(0.0028)	Mo. Since Visited Any PCP	2.8	0.11***
		2.6%			(0.011)
No. of ED for Non-Emergent	0.36	0.0097**			3.9%
		(0.0043)	Mo. Since Visited Any SP	2.1	-0.036***
		2.7%			(0.0059)
No. of ED for Injury	0.29	0.0070*			-1.7%
		(0.0039)			
		2.4%			

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure.

Table A.17: *Treatment Effect of PCP Unexpectedly Leaving Practice*

Type	Mean	Impact Year 1 Post Exit	Impact Year 2 Post Exit	Impact Year 3 Post Exit	Impact Year 4 Post Exit
Total PCP, SP, UC, and ED Utilization	16.0	-0.36*** (0.070) -2.2%	-0.49*** (0.076) -3.0%	-0.60*** (0.081) -3.7%	-0.59*** (0.085) -3.7%
Utilization of Services					
No. of Primary Care Visits	5.5	-0.87*** (0.046) -15.9%	-0.88*** (0.048) -16.0%	-0.91*** (0.050) -16.6%	-0.91*** (0.051) -16.7%
Prob. Form New PCP Relationship	0.040	0.057*** (0.0018) 143.1%	0.020*** (0.0013) 49.8%	0.017*** (0.0013) 42.9%	0.014*** (0.0013) 35.6%
No. of Patients Visting PCP at Least Once	0.86	-0.10*** (0.0036) -12.1%	-0.11*** (0.0037) -12.3%	-0.10*** (0.0037) -12.2%	-0.10*** (0.0037) -12.0%
No. of PCP Visits, Intensive Margin	6.2	-0.84*** (0.15) -13.5%	-0.86*** (0.15) -13.8%	-0.95*** (0.15) -15.2%	-0.76*** (0.15) -12.2%
No. of Specialist Visits	9.4	0.47*** (0.049) 5.1%	0.37*** (0.054) 4.0%	0.32*** (0.057) 3.5%	0.32*** (0.061) 3.5%
No. of Urgent Care Visits	0.0088	0.00018 (0.00079) 2.0%	-0.0013 (0.00079) -14.2%	-0.00066 (0.00091) -7.5%	0.000069 (0.00093) 0.78%
Tot. Spending	8653.1	290.3*** (97.5) 3.4%	176.0* (98.9) 2.0%	-27.1 (103.7) -0.31%	-21.7 (107.5) -0.25%
Tot. Out of Pocket	1294.7	36.1** (15.9) 2.8%	28.1* (16.3) 2.2%	-1.7 (17.2) -0.13%	4.1 (17.8) 0.32%
Aggregate Markers for Poor Care					
No. of Emergency Department Visits	0.74	0.021*** (0.0071) 2.9%	0.0066 (0.0075) 0.90%	-0.013 (0.0078) -1.7%	-0.0030 (0.0085) -0.41%
No. of Inpatient Visits	0.39	0.013** (0.0049) 3.4%	0.0090* (0.0050) 2.3%	0.0010 (0.0052) 0.27%	0.00099 (0.0054) 0.25%
Prob. of Death	0.050	0.00070 (0.0014) 1.4%	0.0011 (0.0013) 2.2%	0.00090 (0.0014) 1.8%	-0.00024 (0.0015) -0.48%
Medications					
No. of Filled Prescriptions	14.0	0.011 (0.075) 0.081%	0.075 (0.085) 0.53%	0.19* (0.11) 1.4%	0.26* (0.13) 1.9%
No. of Chronic Med RX Fills	5.4	0.034 (0.031) 0.62%	0.060 (0.035) 1.1%	0.11** (0.041) 1.9%	0.096* (0.048) 1.8%
Preventive Care					
Prob. of a Flu Vaccine	0.47	-0.036*** (0.0036) -7.8%	-0.029*** (0.0034) -6.3%	-0.027*** (0.0034) -5.7%	-0.026*** (0.0036) -5.4%
No. of Annual Exams	0.040	-0.022*** (0.0042) -54.5%	-0.016*** (0.0044) -39.4%	-0.013** (0.0048) -33.8%	-0.013** (0.0052) -32.8%
No. of Preventive Screens	1.5	-0.048*** (0.017) -3.1%	-0.065*** (0.020) -4.2%	-0.11*** (0.022) -7.2%	-0.13*** (0.024) -8.3%

Table A.18: *Treatment Effect of a PCP Leaving a Clinic with Practice Size Match*

Type	Mean	Impact
Utilization of Clinic Based Services		
No. of Specialist and Primary Care Visits	14.9	-0.35***
		(0.049)
		-2.4%
No. of Primary Care Visits	5.4	-0.90***
		(0.033)
		-16.7%
No. of Specialist Visits	9.5	0.55***
		(0.036)
		5.7%
No. of Urgent Care Visits	0.012	0.00095
		(0.00063)
		8.0%
Tot. Spending	8576.8	145.3*
		(70.8)
		1.7%
Preventive Care		
Tot. Amount of Preventive Care	2.2	-0.074***
		(0.015)
		-3.4%
Prob. of a Flu Vaccine	0.48	-0.032***
		(0.0024)
		-6.7%
No. of Annual Exams	0.091	-0.022***
		(0.0033)
		-23.7%
No. of Preventive Screens	1.6	-0.021
		(0.014)
		-1.3%
Medications		
No. of Filled Prescriptions	18.9	0.36***
		(0.079)
		1.9%
No. of Chronic Med RX Fills	7.0	0.16***
		(0.031)
		2.2%
Aggregate Markers for Poor Care		
No. of Emergency Department Visits	0.75	0.036***
		(0.0052)
		4.8%
No. of ED Visits, Not Preventable	0.11	0.0081***
		(0.0019)
		7.3%
No. of ED Visits, Primary Care Treatable	0.37	0.018***
		(0.0034)
		4.8%
No. of Inpatient Visits	0.36	0.0069*
		(0.0034)
		1.9%
Prob. of Death	0.045	0.00097
		(0.00092)
		2.2%
Treated PCP Sample Size	9596	
Control PCP Sample Size	9596	

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. See Section A.1.1 for how medications were defined.

Table A.19: *Treatment Effect of a PCP Leaving a Clinic with Individual/Team Match*

Type	Mean	Impact
Utilization of Clinic Based Services		
No. of Specialist and Primary Care Visits	14.8	-0.41***
		(0.044)
		-2.8%
No. of Primary Care Visits	5.3	-0.84***
		(0.028)
		-15.8%
No. of Specialist Visits	9.5	0.44***
		(0.032)
		4.6%
No. of Urgent Care Visits	0.012	0.0011*
		(0.00054)
		9.6%
Tot. Spending	8443.8	116.5*
		(64.6)
		1.4%
Preventive Care		
Tot. Amount of Preventive Care	2.2	-0.076***
		(0.015)
		-3.4%
Prob. of a Flu Vaccine	0.49	-0.030***
		(0.0021)
		-6.0%
No. of Annual Exams	0.099	-0.020***
		(0.0029)
		-20.1%
No. of Preventive Screens	1.7	-0.028*
		(0.014)
		-1.7%
Medications		
No. of Filled Prescriptions	18.7	0.23***
		(0.072)
		1.2%
No. of Chronic Med RX Fills	7.0	0.12***
		(0.028)
		1.7%
Aggregate Markers for Poor Care		
No. of Emergency Department Visits	0.73	0.031***
		(0.0048)
		4.2%
No. of ED Visits, Not Preventable	0.11	0.0069***
		(0.0017)
		6.3%
No. of ED Visits, Primary Care Treatable	0.36	0.013***
		(0.0031)
		3.7%
No. of Inpatient Visits	0.36	0.0084**
		(0.0031)
		2.3%
Prob. of Death	0.045	0.0015*
		(0.00083)
		3.4%
Treated PCP Sample Size	11906	
Control PCP Sample Size	11906	

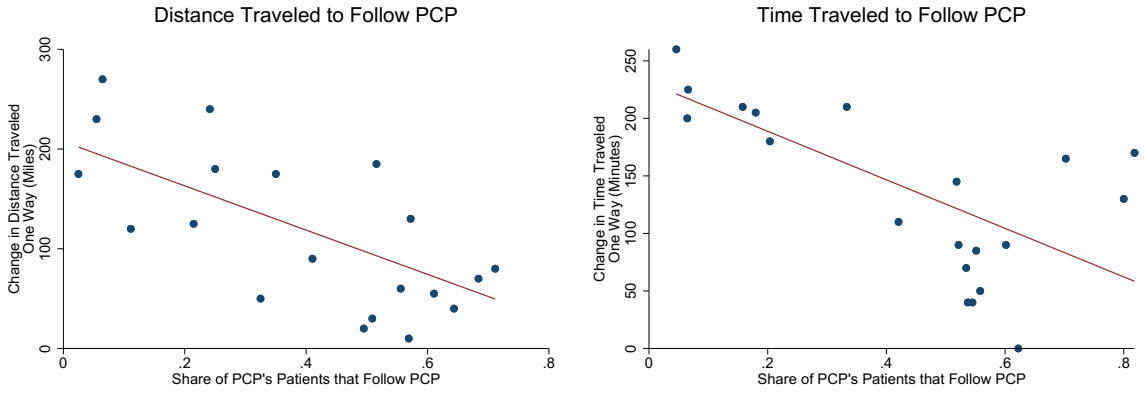
Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. See Section A.1.1 for how medications were defined.

Table A.20: *Additional Results*

Type	Mean	Impact
Additional Medication Results		
No. of RX Starts	2.1	0.14*** (0.013) 6.6%
No. of RX Ends	1.9	0.033** (0.014) 1.7%
No. of RX Classes Started	1.6	0.100*** (0.012) 6.2%
No. of RX Classes Ended	1.3	0.0047 (0.012) 0.35%
No. of RX Switches	0.30	0.026*** (0.0028) 8.7%
No. of Opioid RX	1.5	-0.040*** (0.0098) -2.7%
No. of Benzo RX	0.24	-0.011 (0.0076) -4.5%
No. of SSRI RX	1.6	0.010 (0.0096) 0.65%
No. of Antihypertensive RX	4.6	0.042** (0.018) 0.90%
No. of Antidiabetic RX	1.4	-0.0072 (0.0096) -0.52%
No. of Statin RX	2.1	-0.0093 (0.010) -0.45%
Sampling of Sub-Specialties		
No. of Nephrologist Visits	0.14	0.023*** (0.0035) 15.8%
No. of Cardiologist Visits	0.92	0.057*** (0.0078) 6.2%
No. of Gastroenterologists Visits	0.19	0.021*** (0.0030) 11.1%
No. of Surgeon Visits	0.73	0.0092 (0.0066) 1.2%
Preventive Care in Retail Setting		
Prob. of a Flu Vaccine by Retail	0.18	0.012*** (0.0017) 6.6%
No. of Preventive Screens by Retail	0.017	0.0034** (0.0012) 20.3%
Additional Screens		
No. of Depression Screens	0.018	0.0064 (0.0049) 35.2%
No. of Mammography Screens	0.77	-0.0030 (0.0067) -0.39%
No. of Colorectal Cancer Screens	0.17	0.0058 (0.0046) 3.3%
No. of Diabetes Screens	0.092	-0.016*** (0.0033) -17.8%
No. of BMI Screens	0.029	-0.0073** (0.0035) -25.5%
No. of Tobacco Screens	0.035	-0.0014 (0.0023) -3.9%
No. of Bone Density Screens	0.090	0.012*** (0.0012) 13.5%
No. of Cholesterol Screens	0.72	-0.039*** (0.0062) -5.4%

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. "RX Start" signifies that the patient had not been on that particular prescription before. "RX End" signifies that the prescription fill was the last prescription of that particular drug. "RX Classes Started" tags a new drug in a class of drugs that the patient was not previously prescribed. "RX Classes Ended" tages the end of a class of drugs. "RX Switches" is a different drug within the same class of drugs.

Figure A.3: *Distance Traveled by Patients to Visit a PCP Among Patients Who Move With PCPs*



Notes: Binscatter shows the share of a PCP's original pool of patients that follow her when she moves clinics by distance and time traveled.

Table A.21: *Treatment Effect of a PCP Moving to a Nearby Clinic*

Type	Mean	Impact
Utilization of Clinic Based Services		
No. of Specialist and Primary Care Visits	18.9	0.44 (0.87) 2.3%
No. of Primary Care Visits	7.0	-0.11 (0.58) -1.6%
No. of PCP Visits at Clinic	5.8	-3.4*** (0.98) -59.8%
No. of PCP Visits at Other Clinics	1.2	3.3*** (0.92) 271.1%
	5.3	-0.44 (0.58) -8.3%
No. of Specialist Visits	11.9	0.55 (0.66) 4.6%
Medications		
No. of Filled Prescriptions	26.0	-0.94 (1.7) -3.6%
No. of Chronic Med RX Fills	9.1	-0.62 (0.80) -6.8%
Preventive Care		
No. of Preventive Screens	1.5	-0.044 (0.15) -3.0%
Prob. of a Flu Vaccine	0.59	-0.038 (0.039) -6.5%
Treated PCP Sample Size	73	
Control PCP Sample Size	73	

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability. "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. Preventive screens include mammography screens, colorectal cancer screens, cholesterol screens, and diabetes screens. See Section A.1.1 for how medications were defined. See Section 1.4.2 for other variable definitions.

Table A.22: *Treatment Effect of a PCP Leaving a Clinic by Heterogeneity*
Test for Importance of Costs of Starting a New Relationship

	Clinic Also Closed		Clinic Stayed Open				
			Individual		Team		
Type	Mean	Impact	Mean	Impact	Mean	Impact	P-Value
No. of PCP Visits at Clinic	4.4	-3.9***	4.7	**-2.1***	4.1	**-1.7***	p< 0.001
		(0.047)		(0.072)		(0.042)	
		-88.2%		-45.2%		-40.9%	
No. of PCP Visits at Other Clinics	1.2	2.6***	1.1	**1.4***	1.070	**0.98***	p< 0.001
		(0.046)		(0.057)		(0.033)	
		219.7%		121.7%		91.3%	
Prob. Visit Pre-Existing SP	2.6	0.24***	2.6	**0.15***	2.5	**0.088***	0.003
		(0.013)		(0.017)		(0.011)	
		9.3%		6.0%		3.5%	
Prob. of a Flu Vaccine by Specialist	0.037	0.024***	0.039	**0.020***	0.044	**0.012***	0.003
		(0.0021)		(0.0022)		(0.0014)	
		64.4%		50.2%		26.6%	
Treated PCP Sample Size	3523		293		7293		
Control PCP Sample Size	12497		2301		8136		

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. Bolded estimates indicate that the groups are significantly different at the 5% level. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. See Section 1.4.2 for variable definitions and Section 1.4.2 for more details on how heterogeneity was defined.

Table A.23: *Treatment Effect of a PCP Leaving a Clinic by Heterogeneity*
Test for Importance of Costs of Starting a New Relationship
Matching on Team v. Individual Model

	Clinic Also Closed		Clinic Stayed Open				
			Individual		Team		
Type	Mean	Impact	Mean	Impact	Mean	Impact	P-Value
No. of Specialist and Primary Care Visits	15.4	-0.51***	15.5	-0.42***	14.4	-0.37***	0.61
		(0.088)		(0.081)		(0.054)	
		-3.3%		-2.7%		-2.6%	
No. of Primary Care Visits	5.7	-1.3***	5.7	**-0.76***	5.1	**-0.63***	0.04
		(0.062)		(0.060)		(0.032)	
		-22.5%		-13.3%		-12.3%	
Prob. Form New PCP Relationship	0.043	0.12***	0.043	**0.041***	0.046	**0.031***	p< 0.001
		(0.0021)		(0.0020)		(0.0013)	
		272.0%		94.7%		66.8%	
No. of Specialist Visits	9.7	0.76***	9.7	0.35***	9.3	0.26***	0.22
		(0.059)		(0.061)		(0.041)	
		7.8%		3.6%		2.8%	
Tot. Spending	9850.50	342.90***	8881.6	37.9	8059.2	94.3	0.53
		(126.30)		(114.8)		(81.6)	
		3.5%		0.43%		1.2%	
Medications							
No. of Filled Prescriptions	26.6	0.30**	19.9	0.25	17.9	0.27***	0.85
		(0.13)		(0.16)		(0.094)	
		1.1%		1.2%		1.5%	
No. of Chronic Med RX Fills	9.7	0.24***	7.4	0.12**	6.8	0.076**	0.38
		(0.053)		(0.063)		(0.037)	
		2.5%		1.7%		1.1%	
Preventive Care							
Tot. Amount of Preventive Care	2.1	-0.10***	2.4	-0.082***	2.2	-0.059***	0.26
		(0.024)		(0.026)		(0.020)	
		-4.7%		-3.5%		-2.7%	
Prob. of a Flu Vaccine	0.47	-0.032***	0.49	-0.025***	0.51	-0.028***	0.46
		(0.0045)		(0.0042)		(0.0027)	
		-6.8%		-5.1%		-5.5%	
Aggregate Markers for Poor Care							
No. of Emergency Department Visits	0.72	0.049***	0.79	0.032***	0.72	0.022***	0.42
		(0.0083)		(0.0086)		(0.0060)	
		6.8%		4.0%		3.1%	
No. of ED Visits, Not Preventable	0.11	0.011***	0.12	**0.011***	0.11	**0.0028**	0.04
		(0.0029)		(0.0029)		(0.0023)	
		10.2%		9.6%		2.6%	
No. of Inpatient Visits	0.43	0.018***	0.38	0.0048	0.35	0.0085**	0.59
		(0.0068)		(0.0056)		(0.0039)	
		4.2%		1.3%		2.4%	
Prob. of Death	0.045	0.000013	0.046	0.0026*	0.044	0.0013	0.62
		(0.0014)		(0.0015)		(0.0011)	
		0.028%		5.6%		2.9%	
Not Preventable ED, Inpatient Use, and Death	0.59	0.029***	0.54	0.018***	0.50	0.013***	0.61
		(0.0083)		(0.0072)		(0.0053)	
		5.0%		3.4%		2.5%	
Treated PCP Sample Size	2802		2128		6333		
Control PCP Sample Size	8318		4578		7328		

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure.

Table A.24: *Treatment Effect of a PCP Leaving a Clinic by Patient Status*

	25th Percentile Threshold					50th Percentile Threshold					75th Percentile Threshold				
	Individual		Team			Individual		Team			Individual		Team		
Type	Mean	Impact	Mean	Impact	P-Value	Mean	Impact	Mean	Impact	P-Value	Mean	Impact	Mean	Impact	P-Value
No. of Primary Care Visits	5.6	-0.66***	5.0	-0.72***	0.37	6.0	-0.81***	5.1	-0.68***	0.16	7.3	**-1.2***	5.2	**-0.67***	0.03
		(0.042)		(0.047)			(0.086)		(0.033)			(0.23)		(0.032)	
		-11.9%		-14.3%			-13.5%		-13.2%			-15.7%		-12.9%	
No. of Specialist Visits	9.5	**0.40***	9.2	**0.13****	p< 0.001	9.7	**0.50***	9.3	**0.30***	0.03	9.6	0.60***	9.4	0.30***	0.10
		(0.046)		(0.059)			(0.082)		(0.043)			(0.17)		(0.039)	
		4.2%		1.4%			5.2%		3.2%			6.2%		3.2%	
No. of Emergency Department Visits	0.76	0.030***	0.74	0.017*	0.30	0.79	0.036***	0.73	0.017***	0.13	0.90	0.048*	0.75	0.025***	0.37
		(0.0068)		(0.0094)			(0.012)		(0.0062)			(0.027)		(0.0059)	
		3.9%		2.3%			4.6%		2.3%			5.4%		3.3%	
No. of ED Visits, Primary Care Treatable	0.38	0.012***	0.37	0.017***	0.44	0.42	0.0099	0.36	0.0054	0.60	0.47	0.0085	0.37	0.013***	0.89
		(0.0043)		(0.0059)			(0.0079)		(0.0040)			(0.018)		(0.0037)	
		3.0%		4.6%			2.4%		1.5%			1.8%		3.6%	
Treated PCP Sample Size	3871		3715			1298		7676			293		7293		
Control PCP Sample Size	7878		2559			6158		6339			2301		8136		

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure.

Table A.25: *Treatment Effect of a PCP Leaving a Clinic by Length of PCP-Patient Relationship*

Age 75+:	3 Year	4-9 Year	10-11 Year	12-13 Year
No. of Primary Care Visits	-0.69***	-0.88***	-1.083***	-1.060***
	(0.076)	(0.077)	(0.093)	(0.091)
	-13.2%	-15.5%	-17.6%	-17.7%
No. of Specialist Visits	0.23*	0.54***	0.67***	0.77***
	(0.12)	(0.10)	(0.12)	(0.13)
	2.4%	5.9%	6.8%	7.9%
No. of Emergency Department Visits	-0.011	0.028*	0.037*	0.050**
	(0.020)	(0.015)	(0.019)	(0.020)
	-1.4%	4.1%	5.1%	6.8%
Not Preventable ED, Inpatient Use, and Death	0.015	-0.015	0.054**	0.040*
	(0.023)	(0.018)	(0.023)	(0.023)
	2.1%	-2.2%	7.7%	5.8%

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure.

Table A.26: *Treatment Effect of a PCP Leaving a Clinic by Patient Status*
Matching on Practice Size

	High Risk		Not High Risk		Func.		No Func.		Minority		White		Male-Female Match		Female-Female Match	
Type	Mean	Impact	Mean	Impact	Mean	Impact	Mean	Impact	Mean	Impact	Mean	Impact	Mean	Impact	Mean	Impact
No. of Primary Care Visits	7.0	**-1.086*** **	4.9	**-0.84*** **	5.0	**-0.80*** **	6.5	**-1.1*** **	5.3	-0.98***	5.4	-0.89***	5.4	-0.65***	5.0	-0.57***
		(0.052)		(0.033)		(0.032)		(0.057)		(0.084)		(0.035)		(0.038)		(0.089)
		-15.5%		-17.3%		-16.1%		-17.1%		-18.5%		-16.4%		-12.2%		-11.3%
p-value			p< 0.001				p< 0.001				0.31				0.34	
Prob. Form New PCP Relationship	0.030	**0.044*** **	0.043	**0.067*** **	0.041	**0.066*** **	0.035	**0.048*** **	0.025	**0.057*** **	0.043	**0.064*** **	0.044	**0.034*** **	0.030	**0.020*** **
		(0.0013)		(0.0015)		(0.0015)		(0.0015)		(0.0026)		(0.0014)		(0.0014)		(0.0028)
		146.7%		154.2%		162.6%		138.5%		231.5%		147.9%		78.8%		66.3%
p-value			p< 0.001				p< 0.001				0.02				p< 0.001	
No. of Specialist Visits	13.6	0.58***	8.1	0.41***	8.4	0.43***	12.6	0.56***	8.1	0.59***	9.7	0.44***	9.4	**0.29*** **	9.7	**-0.065**
		(0.077)		(0.039)		(0.040)		(0.072)		(0.092)		(0.041)		(0.044)		(0.15)
		4.3%		5.1%		5.1%		4.5%		7.3%		4.6%		3.0%		-0.68%
p-value			0.17				0.06				0.12				0.02	
Tot. Amount of Preventive Care	1.9	-0.071***	2.1	-0.080***	2.0	-0.083***	1.9	-0.062**	1.5	-0.069*	2.1	-0.066***	1.2	-0.044**	1.6	-0.083**
		(0.023)		(0.017)		(0.016)		(0.024)		(0.035)		(0.017)		(0.016)		(0.033)
		-3.7%		-3.9%		-4.1%		-3.3%		-4.6%		-3.1%		-3.7%		-5.3%
p-value			0.65				0.32				0.95				0.27	
No. of ED and Inpatient Visits	1.5	**0.057*** **	0.77	**0.029*** **	0.74	**0.024*** **	1.5	**0.067*** **	1.056	0.029	0.94	0.038***	0.98	0.025***	1.00	0.013
		(0.017)		(0.0070)		(0.0070)		(0.016)		(0.022)		(0.0073)		(0.0082)		(0.031)
		3.8%		3.7%		3.2%		4.4%		2.7%		4.0%		2.6%		1.3%
p-value			0.002				0.010				0.69				0.70	
Treated PCPs (N)	8125		8108		8203		7395		4032		7996		5467		1467	
Control PCPs (N)	8125		8108		8203		7395		4032		7996		8271		1971	

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. Relationship specifications drop 12% of patients contained in the main sample because their PCP's NPI did not match to a UPIN, the PCP identification scheme used from 2002-2016. Further, patients with ages < 75 as of $t = -36$ were dropped from analyses to avoid censoring due to only observing patients once they become Medicare eligible (typically at age 65).

Table A.27: *Treatment Effect of a PCP Leaving a Clinic by Market Heterogeneity Test for Importance of Local Availability*

		Low PCP Density		High PCP Density		
Type		Mean	Impact	Mean	Impact	P-Value
No. of Primary Care Visits		5.3	-0.61***	5.4	-0.55***	0.35
			(0.032)		(0.066)	
			-11.6%		-10.2%	
Prob. Form New PCP Relationship		0.046	**0.033*** **	0.042	**0.024*** **	0.001
			(0.0011)		(0.0024)	
			71.9%		58.0%	
No. of Specialist Visits		9.5	0.30***	8.8	0.28***	0.77
			(0.036)		(0.077)	
			3.2%		3.2%	
No. of Emergency Department Visits		0.75	0.028***	0.74	0.0068	0.10
			(0.0053)		(0.012)	
			3.7%		0.92%	
No. of ED Visits, Primary Care Treatable		0.38	0.012***	0.37	0.0036	0.32
			(0.0034)		(0.0079)	
			3.2%		0.96%	
No. of ED Visits, Not Preventable		0.12	0.0082***	0.12	0.0029	0.24
			(0.0019)		(0.0040)	
			7.0%		2.4%	
Treated PCP Sample Size	6172	1414				
Control PCP Sample Size	8619	1818				

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. Bolded estimates indicate that the groups are significantly different at the 5% level. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. See Section 1.4.2 for more details on how heterogeneity and variables were defined. Density is defined by the number of PCPs within a 30 mile radius divided by the population . Above average areas are high PCP density and below average areas are low PCP density. Figure 1.8 shows monthly estimates.

Table A.28: *Treatment Effect of a PCP Leaving a Clinic With More Than One Similar Clinic Option*

	Similar TIN Surrounding Clinic that Closed				
	No Similar Clinic		At Least One Similar Clinic		
Type	Mean	Impact	Mean	Impact	P-Value
No. of Primary Care Visits	5.7	-1.2***	5.7	-1.3***	0.20
		(0.072)		(0.093)	
		-20.9%		-23.7%	
No. of UC and Specialist Visits	9.8	0.85***	9.7	0.74***	0.28
		(0.065)		(0.077)	
		8.8%		7.7%	
No. of ED and Inpatient Visits	0.92	0.059***	0.93	0.043***	0.38
		(0.012)		(0.013)	
		6.4%		4.7%	
	Similar TIN Surrounding Open Clinics				
	No Similar Clinic		At Least One Similar Clinic		
Type	Mean	Impact	Mean	Impact	P-Value
No. of Primary Care Visits	5.4	-0.51***	5.3	-0.65***	0.05
		(0.065)		(0.033)	
		-9.3%		-12.3%	
No. of UC and Specialist Visits	9.3	**0.44*** **	9.4	**0.25*** **	0.010
		(0.064)		(0.039)	
		4.7%		2.6%	
No. of ED and Inpatient Visits	0.96	0.042***	0.96	0.021***	0.14
		(0.012)		(0.0073)	
		4.4%		2.2%	

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Data only includes patients who lost a PCP and clinic. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure. Table **??** contains additional results. The bins were determined as above average and below average, where average clinic is not surrounded by a clinic with the same TIN. TIN density is calculated as whether there exists a non-focal clinic with the same TIN as the focal clinic, what I term a "sister clinic." The average clinic does not have a sister clinic, so I compare clinics with zero to those with at least one sister clinic.

Table A.29: *Treatment Effect of a PCP Leaving a Clinic by Urban/Rural Area*

	Rural Area		Urban Area		
Type	Mean	Impact	Mean	Impact	P-Value
No. of Primary Care Visits	5.5	**-0.88*** **	5.3	**-0.52*** **	p$<$ 0.001
		(0.066)		(0.032)	
		-16.0%		-9.9%	
No. of UC and Specialist Visits	7.6	0.37***	9.9	0.30***	0.34
		(0.065)		(0.037)	
		4.8%		3.0%	
No. of ED and Inpatient Visits	1.059	0.040***	0.94	0.022***	0.26
		(0.015)		(0.0067)	
		3.8%		2.3%	
No. of ED Visits, Primary Care Treatable	0.43	0.022***	0.36	0.0069**	0.09
		(0.0079)		(0.0034)	
		5.0%		1.9%	
Treated PCP Sample Size	1606		5981		
Control PCP Sample Size	1688		8749		

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3, which only follows patients for one year post-departure. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level.

Table A.30: *Treatment Effect of a PCP Leaving a Clinic by Heterogeneity Test for Importance of Local Availability*

Matching on Practice Size					
	Low PCP Density		High PCP Density		
No. of Primary Care Visits	5.3	**-0.64*** **	5.3	**-0.44*** **	0.010
		(0.037)		(0.070)	
		-12.0%		-8.3%	
No. of Patients Visting PCP at Least Once	0.86	**-0.083*** **	0.86	**-0.068*** **	0.04
		(0.0035)		(0.0064)	
		-9.6%		-7.9%	
Prob. Form New PCP Relationship	0.14	0.079***	0.13	0.066***	0.14
		(0.0041)		(0.0080)	
		56.2%		50.2%	
Prob. Visit Pre-Existing PCP	3.0	**-0.45*** **	2.9	**-0.38*** **	0.04
		(0.018)		(0.033)	
		-15.3%		-12.8%	
No. of Specialist Visits	9.5	0.32***	8.8	0.20**	0.21
		(0.041)		(0.086)	
		3.3%		2.2%	
Tot. Amount of Preventive Care	2.2	-0.053***	2.1	-0.048	0.89
		(0.018)		(0.035)	
		-2.5%		-2.2%	
No. of Emergency Department Visits	0.77	0.025***	0.75	0.023	0.87
		(0.0058)		(0.014)	
		3.3%		3.0%	
Matching on Individual v. Shared					
No. of Primary Care Visits	5.3	**-0.65*** **	5.4	**-0.49*** **	0.04
		(0.037)		(0.070)	
		-12.3%		-9.1%	
No. of Patients Visting PCP at Least Once	0.86	-0.084***	0.86	-0.071***	0.09
		(0.0035)		(0.0066)	
		-9.7%		-8.3%	
Prob. Form New PCP Relationship	0.14	0.077***	0.13	0.067***	0.24
		(0.0041)		(0.0080)	
		54.8%		49.9%	
Prob. Visit Pre-Existing PCP	2.9	-0.46***	3.0	-0.40***	0.09
		(0.018)		(0.034)	
		-15.7%		-13.5%	
No. of Specialist Visits	9.5	0.31***	8.9	0.16*	0.12
		(0.042)		(0.087)	
		3.3%		1.8%	
Tot. Amount of Preventive Care	2.1	-0.045***	2.2	-0.039	0.89
		(0.018)		(0.035)	
		-2.1%		-1.8%	
No. of Emergency Department Visits	0.77	0.028***	0.74	0.023*	0.70
		(0.0060)		(0.014)	
		3.7%		3.1%	

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3. "Prob." indicates that the outcome is the yearly probability and "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post-departure.

Table A.31: *Treatment Effect of a PCP Leaving a Clinic*
Utilization of Clinic Based Services by Closing Clinic Practice Size

Within Closed Clinics	1 PCP		2+ PCPs		
Type	Mean	Impact	Mean	Impact	P-Value
No. of Primary Care Visits	5.7	**-1.2***	5.4	**-2.0***	p< 0.001
		(0.065)		(0.17)	
		-20.9%		-37.2%	
No. of Specialist Visits	9.9	**0.95***	9.0	**0.29***	p< 0.001
		(0.063)		(0.12)	
		9.6%		3.2%	
No. of ED and Inpatient Visits	0.93	0.056***	0.88	0.058***	0.91
		(0.011)		(0.020)	
		6.0%		6.6%	
Tot. Spending	10092.6	246.7*	8337.7	432.3	0.52
		(135.5)		(267.5)	
		2.4%		5.2%	
Treated PCP Sample Size	2416		435		
Control PCP Sample Size	6129		4308		

Notes: The table displays results from the difference-in-differences specification outlined in Section 1.5.3, which only follows patients for one year post-departure. "No." indicates that the outcome is the yearly number. Regressions are at the PCP-year level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Clinics with 1-3 PCPs were compared to those with 4-100 PCPs because 3 PCPs was the median practice size (7 is the mean). Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level.

Table A.32: *Treatment Effect of PCP Leaving Practice by Whether Open Clinic Practices on Team or Individual Model by Clinic Size*

	1-3 PCPs	4-100 PCPs
No. Actual Clinic Visits		
Shared	**-1.7***	**-1.7***
	(0.12)	(0.047)
	3.1	3.9
	-54.0%	-43.5%
Individual	**-2.4***	**-2.1***
	(0.12)	(0.087)
	4.4	4.4
	-53.5%	-48.4%
No. Other PCP Clinic Visits		
Shared	**0.99***	**1.0052***
	(0.078)	(0.036)
	1.6	1.2
	61.9%	86.7%
Individual	**1.4***	**1.4***
	(0.095)	(0.070)
	1.4	1.2
	103.2%	116.0%
Treated Shared	599	5183
Control Shared	2307	2198
Treated Individual	592	1212
Control Individual	5322	610

Notes: The table displays results from the difference-in-differences specification outlined in 1.5.3. Regressions are at the PCP level and contain pre-departure PCP fixed effects. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. Bolded estimates mean that shared and panel groups are significantly different at the 5% level and † means that small and large groups are significantly different at the 5% level. Data relies on the main sample shown in Table 1.1, which only follows patients for one year post departure.

Table A.33: *Effects by Clinic Size*

	All Departures & All Clinics		
	1 PCPs	2-3 PCPs	4+ PCPs
No. of Primary Care Visits	-1.1***†	-0.61***†	-0.74***
	(0.17)	(0.062)	(0.050)
	5.4	5.3	5.3
	-20.5%	-11.5%	-13.9%
No. of PCP Visits at Clinic	-2.3***†	-1.8***†	-1.8***
	(0.18)	(0.067)	(0.056)
	4.0	4.1	4.3
	-58.1%	-44.7%	-42.1%
No. of PCP Visits at Other Clinics	1.2***	1.2***◇	1.078***◇
	(0.11)	(0.049)	(0.042)
	1.5	1.2	0.99
	80.7%	98.2%	108.9%
Prob. Form New PCP Relationship	0.040***	0.036***	0.035***
	(0.0047)	(0.0023)	(0.0021)
	0.051	0.044	0.045
	77.7%	82.3%	76.8%
No. of Specialist Visits	0.59***	0.37***	0.19**
	(0.16)	(0.072)	(0.070)
	9.9	9.6	9.3
	5.9%	3.9%	2.0%
No. of Emergency Department Visits	0.054**	0.034***	0.024**
	(0.025)	(0.011)	(0.011)
	0.93	0.78	0.73
	5.8%	4.3%	3.3%
No. of Urgent Care Visits	-0.0028	0.0038**	0.0019
	(0.0036)	(0.0015)	(0.0017)
	0.025	0.015	0.021
	-11.3%	25.8%	8.8%
No. of Inpatient Visits	-0.0097	0.028***◇	-0.0054◇
	(0.018)	(0.0091)	(0.0090)
	0.48	0.44	0.43
	-2.0%	6.3%	-1.2%
Treated PCP Sample Size	364	2224	4998
Control PCP Sample Size	6129	2833	1475

Notes: Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. † (◇) indicates that estimates from clinics with 1-3 PCPs (10+ PCPs) and 4-9 PCPs are significantly different at the 5% level.

A.6 Event Study Plots

The following plots are from the matched difference-in-difference model described in the methods. Point estimates for every months interacted for whether the observation was a treatment or control observation are plotted. Relative time is relative to the last month the PCP is observed practicing in $t = 0$.

Figure A.4: *Pre-Period and Event Plots per Patient, Relative to $t = -12$*

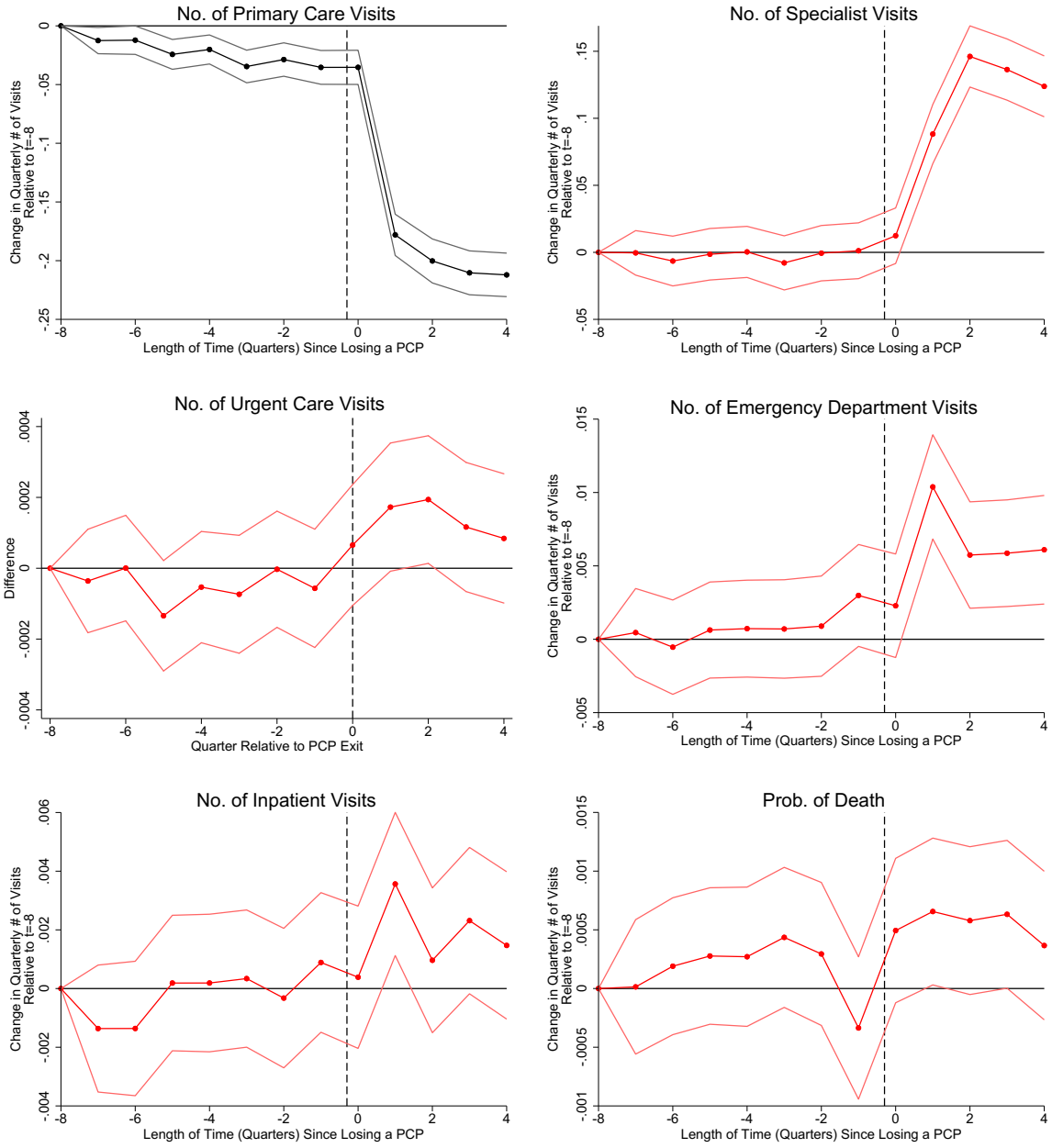
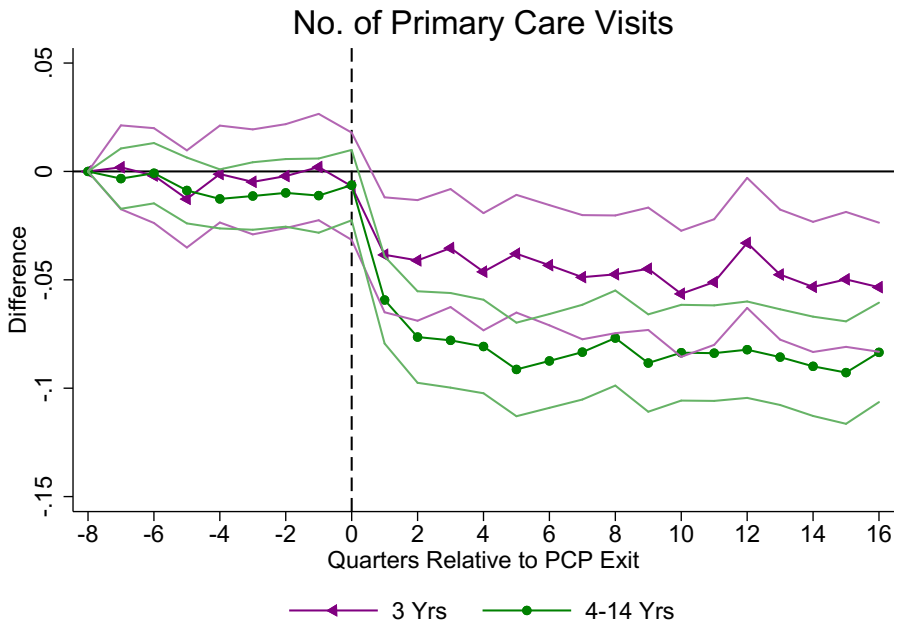
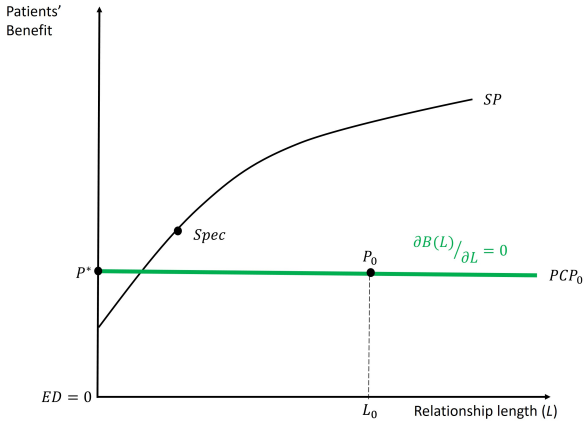


Figure A.6: *Effects of a PCP Leaving on PCP Visits by Length of Relationship*

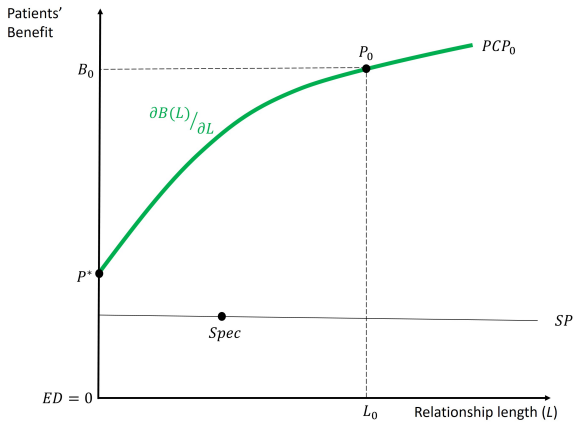


Notes: Event study graphs plot each coefficient from the difference-in-differences specification outlined in Section 1.5.3. "No." indicates that the outcome is the monthly number. Regressions are at the PCP-month level, contain pre-departure PCP fixed effects, and cluster at the PCP-match level. Event studies rely on the a sample that follows patients 4 years post-departure. See Table A.25 for pooled estimates across outcomes.

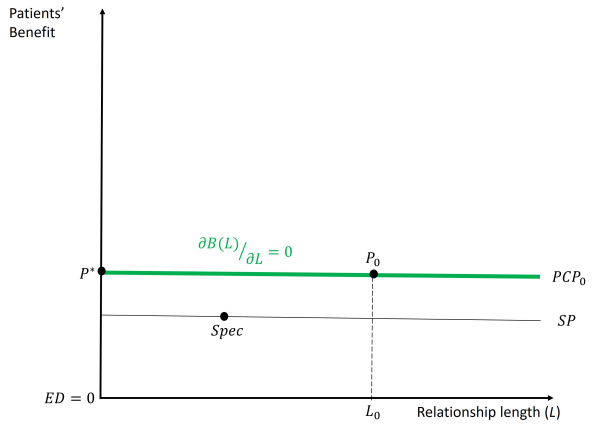
Figure A.7: *Different Information Models*



(a) *Model A*



(b) *Model B*



(c) *Model C*

Appendix B

Supplementary Material for Section 2

Figure B.1: *Timeline*

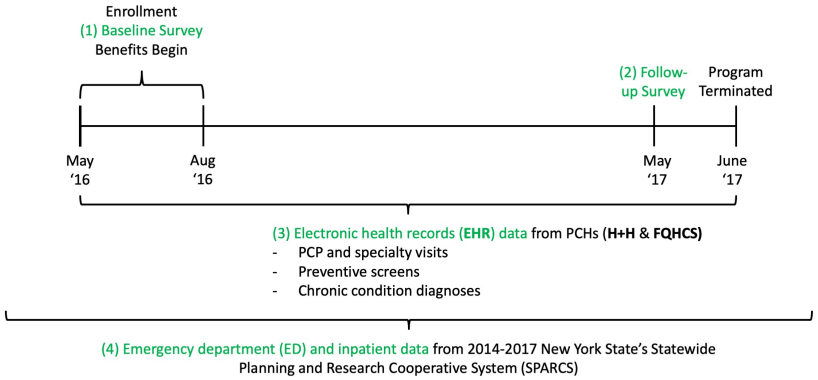


Figure B.2: *H+H and FQHC Network*

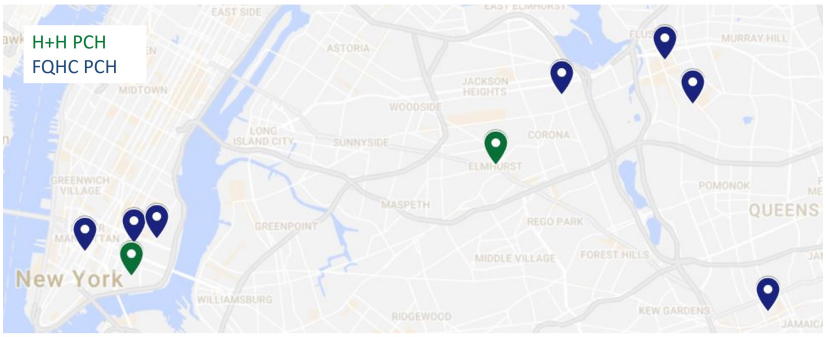


Table B.1: *Fee scale across sites*

	Treatment 0-150% FPL AHNYC	Control 0-150% FPL H+H	Control 0-150% FPL CHN	Control 0-150% FPL CBW	Control 0-150% FPL WR	Control 0-150% FPL UHP
Out-of-Pocket Costs						
Doctors office visit	$15	$15	$15	~$30	$59	$54
Dental	$15	$15	$50	$15	$59	$54
Emergency room visit**	$3	$3	$3	$3	$3	$3
Outpatient surgery	$150	$150	N/A	N/A	N/A	N/A
Inpatient stay	$150	$150	N/A	N/A	N/A	N/A
RX	N/A	$2	N/A	N/A	N/A	N/A

Notes : ActionHealthNYC (AHNYC); Health+Hospitals (H+H); Community Health Network (CHN); Charles B. Wong (CBW); Ryan-NENA Community Health Center (WR); Urban Health Plan (UHP). The costs that the control group sees are equivalent to the cost-sharing schedule that all individuals saw before ActionHealthNYC existed. Cost-sharing linearly increases once individuals are over 150% of FPL, but since 94% of our population is below 150% FPL it was chosen as the relevant benchmark. **These costs only apply to those from 0%-138% of FPL who qualified for Emergency Medicaid. For those between 138%-150% of FPL they saw ED costs of $15.

Table B.2: *Utilization of Care Using Baseline and Follow-up Survey Additional Results*

	All Individuals		Low Risk Individuals		High Risk Individuals	
Type	Mean	ITT	Mean	Impact	Mean	ITT
Utilization of Care						
Prob. IP Admit Survey	0.050	0.0012	0.036	-0.0096	0.069	0.016
		(0.024)		(0.025)		(0.045)
		2.5%		-26.9%		23.0%
No. IP Admit Survey	1.5	-0.80	1.3	—	1.7	-1.00
		(3.8)		(.)		(5.1)
		-52.7%		—%		-58.6%
Self Reported Costs						
Not Fill RX Bc of Cost	0.14	0.014	0.072	0.015	0.23	0.011
		(0.038)		(0.043)		(0.066)
		9.8%		21.4%		4.9%
Borrow, Skip Other Bills, Bills Late?	0.13	-0.028	0.078	-0.037	0.21	-0.015
		(0.034)		(0.039)		(0.059)
		-20.7%		-47.9%		-7.2%

Notes: Estimates compare self-reported outcomes on the baseline to follow-up survey. Only individuals who completed the follow-up survey are included, or 44% of individuals. Regressions include member fixed effects and are clustered at the couple level because couples were randomized together. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level.

Table B.3: *SPARCS Administrative Data*

	All Individuals		Low Risk Individuals		High Risk Individuals	
	Mean	ITT	Mean	ITT	Mean	ITT
Utilization						
Prob. IP Visit	0.051	0.00034 (0.010) 0.65%	0.045	-0.0048 (0.012) -10.7%	0.060	0.0071 (0.017) 11.9%
No. IP Visits	1.2	0.29 (1.020) 23.8%	1.1	0.67 (0.71) 59.6%	1.3	0.15 (1.4) 11.4%
Charges						
ED Tot. Charges, Winsorized	45235.8	-13845.1* (7026.7) -30.6%	28548.0	**5012.7** (7330.8) 17.6%	66790.8	**-38761.5***** (13011.2) -58.0%
IP Tot. Charges	160623.5	-35523.6 (114180.4) -22.1%	129494.0	-129367.5 (175109.4) -99.9%	200832.4	88471.0 (129183.9) 44.1%
IP Tot. Charges, Winsorized	126701.6	3915.3 (37923.4) 3.1%	100867.7	-2194.2 (29237.0) -2.2%	160070.3	11987.8 (79132.9) 7.5%
Treated Sample Size	1265		713		552	
Control Sample Size	1139		654		485	

Notes: Estimates are relative to the pre-period, which occurs 14 months before May 2016, the date when the intervention began. The intervention window spans May 2016 through June 2017. Emergency department (ED) and inpatient (IP) results use two years of additional pre-period data to estimate regressions with a linear time trend. "Prob." captures the extensive margin change in utilization across the period and "No." represents the number of visits conditional on non-zero utilization, or intensive margin change. The high risk patient category includes patients with one or more chronic condition, so low risk patients did not have a chronic condition diagnosis at baseline by definition. Bolded estimates indicate that High Risk and ECC Group estimates are significantly different at the 5% level. Regressions include individual fixed effects and are clustered at the couple level because couples were randomized together. Stars indicate significance at the 10% (*), 5% (**), and 1% (***) level. SPARCS administrative data captures any ED or IP visit in New York State, whereas EHR data from PCHs only contains information on individuals who used our PCH facilities. Winsorized estimates top code the top 0.1% of charges.

B.1 Details of Health Valuation Calculation

To calculate the value of screening individuals for chronic diseases, we rely on estimates from the literature to do a back-of-the-envelope calculation on how screening affects long-run mortality.

B.1.1 Number of Cardiovascular Deaths Averted

Blood Pressure:

$$4.3 * (1080/0.90) = 5160 \text{ deaths per } 100,000$$

4.3 percentage points (pp) is the increase in the probability an individual is given a

blood pressure screen in response to being enrolled in the program (Table 2.4). According to Dehmer et al. (2017), 1,080 is the number of deaths averted through blood pressure screening (per 100,000) and 0.9 scales the estimate to the entire population (or a screening rate of 90%).

Diabetes:

$$4.9 * (400) = 1960 \text{ deaths per } 100,000$$

4.9 pp is the increase in the probability an individual is given a diabetes screen in response to being enrolled in the program (Table 2.4). According to Kahn et al. (2010), diabetes screens prevent 4 deaths per 1000 people.

Appendix C

Supplementary Material for Chapter 3

C.1

C.1.1 Calculation of willingness to pay for insurance

The expected utility of a consumer with income y of type θ for contract x at premium p is given by $U(x,p,\theta)$, as defined in equation (3.1) and repeated here:

$$U(x,p,\theta) = \mathbb{E}\left[\, u_\psi(y-p+\ b^*(l,x,0) - c^*(l,x,0) + v(l,x,\omega))|\ l \sim F\right].$$

We can express the corresponding certainty equivalent $CE(x,p,\theta)$ as that which solves $u(CE(x,p,\theta)) = U(x,p,\theta)$. We can further write:

$$\begin{aligned} CE(x,p,\theta) &= u_\psi^{-1}(U(x,p,\theta)) \\ &= EV(x,\theta) + y - p + u_\psi^{-1}(U(x,p,\theta)) - EV(x,\theta) + p - y \\ &= EV(x,\theta) + y - p - RP(x,p,\theta), \end{aligned}$$

where $EV(x,\theta) + y - p$ is the expected payoff and $RP(x,p,\theta)$ is the risk premium associated with the lottery. In particular,

$$\begin{aligned} EV(x,\theta) &= \bar{b}(F,x,0) - \bar{c}(F,x,0) + \bar{v}(F,x,\omega) \\ RP(x,p,\theta) &= EV(x,\theta) + y - p - u_\psi^{-1}(U(x,p,\theta)), \end{aligned} \tag{C.1}$$

where $\bar{b}(F, x, \omega)$ is the expected value of $b^*(l, x, \omega)$ with respect to l, and $\bar{c}(F, x, \omega)$ and $\bar{v}(F, x, \omega)$ are similarly defined. A consumer's willingness to pay for contract x relative to the null contract x_0 is equal to $\tilde{p}$ that solves:

$$\begin{aligned} CE(x, \tilde{p}, \theta) &= CE(x_o, p_o, \theta) \\ EV(x, \theta) + y - \tilde{p} - RP(x, \tilde{p}, \theta) &= EV(x_0, \theta) + y - p_0 - RP(x_0, p_0, \theta) \\ \tilde{p} - p_0 &= EV(x, \theta) - EV(x_0, \theta) + RP(x_0, p_0, \theta) - RP(x, \tilde{p}, \theta), \end{aligned}$$

To obtain a closed form expression for willingness to pay, we assume constant absolute risk aversion, and thus that the risk premium RP does not depend on residual income $y - p$.[1] In this case, marginal willingness to pay for contract x relative to the null contract is given by:

$$\begin{aligned} WTP(x, \theta) &= EV(x, \theta) - EV(x_0, \theta) + RP(x_0, \theta) - RP(x, \theta) \\ &= \bar{c}(F, x_0, \omega) - \bar{c}(F, x, 0) + \bar{v}(F, x, \omega) + \Psi(x, \theta), \end{aligned}$$

where $\Psi(x, \theta) = RP(x_0, \theta) - RP(x, \theta)$. The last step uses the facts that (i) $\bar{b}(F, x, 0) = \bar{b}(F, x_0, 0)$ because the choice of optimal healthcare utilization is the same across contracts if there is not moral hazard, and (ii) $\bar{v}(F, x_0, \omega) = 0$ because there is not spending due to moral hazard in the null contract.

C.1.2 Estimation of plan cost sharing features

A key input to our empirical model is the cost sharing function of each plan that maps healthcare utilization into out-of-pocket costs. While Table 3.2 describes plans using the deductible and in-network out-of-pocket maximum, they are in reality characterized by a much more complex set of payment rules, including co-payments, specialist visit coinsurance, out-of-network fees, and fixed charges for emergency room visits. To structurally model moral hazard, we make the important simplification that healthcare is a homogenous good over which the consumer must choose only the quantity to consume. As described in Section

[1] In equation (C.1), $y - p$ cancels out completely. This assumption is often reasonable given that marginal premiums between relevant plans are small relative to income.

3.2, consumers decide how much healthcare to consume based in part on out-of-pocket cost. To that end, our empirical model requires as an input a univariate function that maps total healthcare spending into out-of-pocket spending.

A natural choice for such a function might be to use the deductible, non-specialist coinsurance rate, and in-network out-of-pocket maximum. However, in our setting, the out-of-pocket cost function described by these features does not correspond well to the shape of the relationship between out-of-pocket spending and total spending that we observe in the claims data. In particular, we often observe out-of-pocket spending amounts that exceed plans' in-network out-of-pocket maximum. Because of this, we take a different approach.

We define plan cost sharing functions by three parameters: a deductible, a coinsurance rate, and an out-of-pocket maximum. Taking the true deductibles as given (since these correspond well to the data), we estimate a coinsurance rate and an out-of-pocket maximum that minimizes the sum of squared residuals between predicted and observed out-of-pocket spending. We observe realized total healthcare spending for each household in the claims data. Predicted out-of-pocket spending is calculated by applying the deductible and supposed coinsurance rate and out-of-pocket maximum. Observed out-of-pocket spending is either observed directly in the claims data (if a household chose that plan) or else calculated counterfactually. We calculate counterfactual out-of-pocket spending using the "claims calculator" developed for this setting by Abaluck and Gruber (2016). We carry out this procedure separately for each plan, year, and family status (individual and family).[2]

Figure C.1 shows the data used to estimate the cost sharing features of a particular plan (Moda Plan 3 for individual households in 2012). Each open circle indicates a household; total healthcare spending is on the horizontal axis and out-of-pocket spending is on the vertical axis.[3] The dark dots are a binscatter plot of the gray open circles data, using 100 data points. The observed, basic cost sharing features of the plan are a deductible of $300, non-

[2]So that the cost-sharing estimates are not affected by large outliers, we drop observations where out-of-pocket spending was above $20,000 or total healthcare spending was above $100,000.

[3]Because there are thousands of individual households in 2012, the plot only shows the dots for a 20 percent random sample.

specialist coinsurance rate of 20 percent, and in-network out-of-pocket maximum of $2,000. It is clear that the data do not correspond well to a $2,000 out-of-pocket maximum. The red line shows the "estimated" cost sharing function of the plan: the estimated coinsurance rate is 20.5 percent and the estimated out-of-pocket maximum is $3,218. Table C.3 presents the estimated cost sharing features for all plans in all years.

C.1.3 Descriptive Evidence: Additional Details

Explaining Variation in Plan Menu Generosity. We replicate the analysis comparing plan menu generosity to observed household health risk in 2009–2013; these estimates are presented in Table C.6. The logit model (equation (3.7)) that produces predicted actuarial value is estimated separately for each year; these estimates are presented in Table C.5. We can consistently reject the hypothesis that household risk scores are correlated with plan menu generosity, conditional on family structure. We also consistently find that plan menus are most generous for single employee coverage and least generous for employee plus family coverage. This is consistent with our understanding of OEBB's benefit structure and is common in employer-sponsored health insurance.

We further explore what covariates, in addition to family structure, *do* seem to explain variation in plan menu generosity. Table C.7 presents three additional regressions using the 2008 sample of predicted actuarial value on employee-level covariates (part-time versus full-time status, occupation type, and union affiliation) as well as school district-level covariates (home price index and percent of Republicans). Employees are either part-time or full-time. There are eight mutually exclusive employee occupation types; the regressions omit the type "Licensed Administrator."[4] There are five mutually exclusive union affiliations and employees may also not be affiliated with a union; the regressions omit the non-union category. We calculate the average home price index (HPI) in a school district by taking the average zip-code level home price index across employees' zip-code of residence.[5]

[4] "Licensed" refers to the possession of a teaching license.

[5] We use 5-Digit zip-code level home price indices from Bogin, Doerner and Larson (2019). The data and

Pct. Republican measures the percent of households in a school district that are registered as Republicans as of 2016.[6]

We find that plan menus are less generous on average for part-time employees, are substantially less generous for substitute teachers, and are more generous for employees at community colleges. Certain union affiliations are also predictive of more or less generous plan menus. Across school districts, predicted actuarial value is decreasing in both the logged home price index as well as the percent of registered Republicans.

Heterogeneity in Moral Hazard. In section 3.3.2, we present evidence of heterogeneity in moral hazard across quartiles of household risk score. Here, we explore the extent to which this heterogeneity can be explained by variation in the intensity of treatment. Assignment into a lower or higher coverage plan could affect total spending by exposing consumers to lower or higher out-of-pocket costs. However, if a consumer is so healthy that they would almost always be consuming healthcare at levels below the deductible of both plans, then there is in fact no variation in coverage level for that consumer. The same could be true of very sick households that, knowing they will always spend the out-of-pocket maximum, will consume healthcare in the same way in both plans.

Table C.9 compares the realized spending outcomes of households in different risk quartiles with the variation in plan cost-sharing features that gives rise to different marginal out-of-pocket prices. The top panel of Table C.9 shows the observed distributions of total spending among the four quartiles of risk for individual and family households. The bottom panel shows the (in-network) deductible and out-of-pocket maximum (OOP Max.) for each of the Moda plans in 2008. We find that the heterogeneity in our moral hazard estimates in Table 3.3 lines up well with households' potential exposure to varying marginal out-of-pocket costs. For example, individual households in the first quartile have the majority of the density of their spending distribution around or below the plan deductibles. Individual

paper are accessible at http://www.fhfa.gov/papers/wp1601.aspx.

[6]Data on percent of registered voters by party is available at the county level; we construct school district measured by taking the average over employees' county of residence. Voter registration data in Oregon can be downloaded at https://data.oregon.gov/api/views/6a4f-ecbi.

households in the third and fourth quartiles of individual households have the majority of their spending near or above the plan out-of-pocket maximums.

The patterns of heterogeneity in our estimates of moral hazard in Table 3.3 correspond well to the likely variation in marginal out-of-pocket prices facing each type of household. For example, we estimate the largest amount of moral hazard for the second quartile of individual households, whose spending distribution more closely spans the range over which there would in fact be marginal out-of-pocket price variation across plans. Likewise for family households, those in the fourth quartile nearly all have spending above the highest out-of-pocket maximum, and we do not estimate any moral hazard within this group. While this exercise is merely suggestive, it points to the fact that a key dimension of heterogeneity is the extent to which households are exposed to differential out-of-pocket spending across nonlinear insurance contracts. Our theoretical and empirical models are well-equipped to capture this issue.

C.2 Estimation Details

C.2.1 Fenton-Wilkinson Approximation

As there is no known closed form solution for the distribution of the sum of lognormal random variables, the Fenton-Wilkinson approximation is widely used in practice.[7] Under the Fenton-Wilkinson approximation, the distribution of the sum of draws from independent lognormal distributions can be represented by a lognormal distribution. The parameters of the approximating lognormal distribution are chosen such that its first and second moments match the moments of the true distribution of the sum of lognormals, which it is simple to calculate. In our application, the sum of lognormals is the household's health state distribution and the independent lognormals being summed are the household's individuals' health state distributions. Individuals are assumed to face lognormal distributions of health

[7] See Fenton (1960), and for a summary, Cobb, Rumí and Salmerón (2012).

states according to:

$$\log(\tilde{l}_i + \kappa_i) \sim N(\mu_i, \sigma_i^2).$$

All parameters may vary over time (since individual demographics vary over time), but t subscripts are omitted here for simplicity. The moment matching conditions for the distribution of a household level health state are:

$$E(\tilde{l}_k + \kappa_k) = \sum_{i \in \mathcal{I}_k} E(\tilde{l}_i + \kappa_i), \tag{C.2}$$

$$Var(\tilde{l}_k + \kappa_k) = \sum_{i \in \mathcal{I}_k} Var(\tilde{l}_i + \kappa_i), \tag{C.3}$$

$$E(\tilde{l}_k) = \sum_{i \in \mathcal{I}_k} E(\tilde{l}_i). \tag{C.4}$$

where $\mathcal{I}_k$ is the set of individuals in household k. Equation (C.2) sets the mean of the household's health state distribution equal to the sum of the means of each individual's health state distributions. Equation (C.3) matches the variance. Because we have a third parameter to estimate (the shift, κ_k), we use a third moment matching condition to match the first moment of the unshifted distribution, shown in equation (C.4).

Under the approximating assumption that $l_k + \kappa_k$ is distributed lognormally, and substituting the analytical expressions for the mean and variable of a lognormal distribution, these equations become:

$$\exp(\mu_k + \frac{\sigma_k^2}{2}) = \sum_{i \in \mathcal{I}_k} \exp(\mu_i + \frac{\sigma_i^2}{2})$$

$$(\exp(\sigma_k^2) - 1)\exp(2\mu_k + \sigma_k^2) = \sum_{i \in \mathcal{I}_k} (\exp(\sigma_i^2) - 1)\exp(2\mu_i + \sigma^2)$$

$$\exp(\mu_k + \frac{\sigma_k^2}{2}) - \kappa_k = \sum_{i \in \mathcal{I}_k} \exp(\mu_i + \frac{\sigma_i^2}{2}) - \kappa_i$$

This leaves three equations in three unknowns for the parameters of a household's distribu-

tion. The solutions for μ_k, σ_k^2, and κ_k are as follows:

$$\sigma_k^2 = log[1 + \left[\sum_{i \in \mathcal{I}_k} \exp(\mu_i + \frac{\sigma_i^2}{2})\right]^{-2} \sum_{i \in \mathcal{I}_k} (\exp(\sigma_i^2) - 1) \exp(2\mu_i + \sigma_i^2)]$$

$$\mu_k = -\frac{\sigma_k^2}{2} + log[\sum_{i \in \mathcal{I}_k} \exp(\mu_i + \frac{\sigma_i^2}{2})]$$

$$\kappa_k = \sum_{i \in \mathcal{I}_k} \kappa_i$$

Given these algebraic solutions for the parameters of a household's distribution, we need only to estimate the individual-level parameters.

C.2.2 Estimation Algorithm

In this appendix we describe the details of the algorithm used to estimate our model of health insurance and healthcare demand. We estimate the model using a simulated maximum likelihood approach similar to that described in Revelt and Train (1998) and Train (2009), with the appropriate extension to a discrete/continuous choice model in the style of Dubin and McFadden (1984). The maximum likelihood estimator selects the parameter values that maximize the conditional probability density of households' observed total healthcare spending, given their plan choices.

The model contains three dimensions of unobservable heterogeneity: risk aversion, household health, and the moral hazard parameter. Random variables $\beta_{kt} = \{\psi_k, \mu_{kt}, \omega_k\}$ are distributed as described by equation (3.14). We denote the full set of model parameters to estimate as θ, which among other things contains the parameters of the distribution of the random variables. Given a guess of θ, we simulate the distribution of β_{kt} using Gaussian quadrature with 27 support points, yielding simulated points $\beta_{kts}(\theta) = \{\psi_{ks}, \mu_{kts}, \omega_{ks}\}$, as well as weights W_s.[8,9] For each simulation draw s, we then calculate the conditional

[8]Note that some components of ψ_{ks}, μ_{kts}, and ω_{ks} do not depend on unobservables, and are fixed functions of θ and household demographics.

[9]We use the Matlab program *qnwnorm* to implement this method, with three points in each dimension of unobserved heterogeneity. The program can be obtained as part of Mario Miranda and Paul Fackler's CompEcon Toolbox; for more information see Miranda and Fackler (2002).

density at households' observed total healthcare spending and the probability of households' observed plan choices.

We first construct individual-level health state distribution parameters μ_{it}, σ_{it}, and κ_{it} from θ and individual demographics, as described in equations 3.12. We then construct household-level health state distribution parameters μ_{kts}, σ_{kt}, and κ_{kt} using the formulas given in equations 3.13 and the draws of $\beta_{kts}(\theta)$. The model predicts that upon realizing their health state l, households choose total healthcare spending m by trading off the benefit of healthcare utilization with its out-of-pocket cost. Specifically, accounting for the fact that zero spending arises from negative health states, the model predicts optimal healthcare spending $m^*_{jt}(l, \omega_{ks}) = \max(0\ ,\ \omega_{ks}(1 - c'_{jt}(m^*)) + l)$ if household k were enrolled in plan j in year t. Inverting the expression, the implied health state l_{kjts} that would have given rise to observed spending m_{kt} under moral hazard parameter ω_{ks} is given by

$$l_{kjts} : \begin{cases} l_{kjts} < 0 & m_{kt} = 0 \\ l_{kjts} = m_{kt} - \omega_{ks}(1 - c'_{jt}(m_{kt})) & m_{kt} > 0. \end{cases}$$

Note that $c'_{jt}(m^*) = 1$ when $m_{kt} = 0$.

Household monetary health states are distributed lognormally according to:

$$l = \phi_f \tilde{l}$$
$$\log(\tilde{l} + \kappa_{kt}) \sim N(\mu_{kts}, \sigma^2_{kt})$$

There are two possibilities to consider. If m_{kt} is equal to zero, the implied health state l_{kjts} is negative. Given the monetary health state l_{kjts}, the implied "quantity" health state is equal to $\tilde{l}_{kjts} = \phi_f^{-1} l_{kjts}$, where f is the insurer offering plan j. Since $\phi_f > 0$, the probability of observing negative l_{kjts} is the probability of observing $\tilde{l}_{kjts} \leq \kappa_{kt}$ if $\tilde{l}_{kjts}$ is lognormally distributed with mean and variance parameters μ_{kts} and σ^2_{kt}. If m_{kt} is greater than zero, it is useful to define $\lambda_{kjts} = \phi_f^{-1} l_{kjts} + \kappa_{kt}$. The density of m_{kt} in this case is given by the density of λ_{kjts} conditional on $m_{kt} > 0$. Taken together, the probability density of total spending m conditional on plan, parameters, and household observables $\mathbf{X}_k$ is given by

$f_m(m_{kt}|c_{jt}, \beta_{kts}, \theta, \mathbf{X}_{kt}) = P(m = m_{kt}|c_{jt}, \beta_{kts}, \theta, \mathbf{X}_{kt})$, where

$$f_m(m_{kt}|c_{jt}, \beta_{ks}, \theta, \mathbf{X}_{kt}) = \begin{cases} \Phi\left(\frac{\log(\kappa_{kt}) - \mu_{kt}}{\sigma_{kt}}\right) & m_{kt} = 0, \\ \phi_f^{-1}\Phi'\left(\frac{\log(\lambda_{kjts}) - \mu_{kt}}{\sigma_{kt}}\right) & m_{kt} > 0, \end{cases}$$

and $\Phi(\cdot)$ is the standard normal cumulative distribution function. For a given guess of parameters, there are certain values of m_{kt} for which the probability density is zero. In order to rationalize the data at all possible parameter guesses, we use a convolution of $f_m(m_{kt}|c_{jt}, \beta_{ks}, \theta, \mathbf{X}_{kt})$ and a uniform distribution over the range [-1e-75, 1e75].[10]

Next, we calculate the probability of a household's observed plan choice. Given θ and β_{kts}, we simulate the distribution of monetary health states l_{kjtsd} using $D = 30$ support points:

$$l_{kjtsd} = \phi_f\left(e^{\mu_{kts} + \sigma_{kt} Z_d} - \kappa_{kt}\right),$$

where Z_d is a vector of points that approximates a standard normal distribution using Gaussian quadrature, with associated weights W_d. We then calculate the optimal healthcare spending choice m_{kjtsd} associated with each potential health state, according to $m^*_{kjtsd} = \max(0\ ,\ \omega_{ks}(1 - c'_{jt}(m^*_{kjtsd})) + l_{kjtsd})$. Because marginal out-of-pocket costs depend on where the out-of-pocket cost function is evaluated, there is not a closed-form solution for m^*_{kjtsd}. Instead, we derive cutoff values on the health state that determine which out-of-pocket cost "region" a household will find optimal.

Plans in our empirical setting are characterized by a deductible, a coinsurance rate, and an out-of-pocket maximum. Because the plans are piece-wise linear (in three pieces), one must only try out three candidate values of $c'(m)$, and then compare optimized utility in each case in order to find the global optimal spending choice. Specifically, $c'(m) = 1$ if spending m is in the deductible region, $c'(m) = c$ in the coinsurance region, and $c'(m) = 0$ in the out-of-pocket maximum region. By performing a generic version of this calculation, we can construct the relevant cutoff values for the health state. Define a

[10] We have experimented with varying these bounds and found that it does not affect parameter estimates as long as the uniform density is sufficiently small.

plan to consist of a deductible D, a coinsurance rate C, and an out-of-pocket maximum O. Define $A = C^{-1}(O - D(1-C))$ to be the level of total spending above which the consumer would reach their out-of-pocket maximum. Under moral hazard parameter ω, the relevant cutoff values are

$$Z_1 = D - \omega(1-C)/2$$
$$Z_2 = O - \omega/2$$
$$Z_3 = A - \omega(1 - C/2),$$

where $Z_1 \leq Z_2 \leq Z_3$ so long as $O \geq D$ and $C \in [0,1]$. There are two types of plans to consider. If D and A are sufficiently far apart (there is a sufficiently large coinsurance region), then only the cutoffs Z_1 and Z_3 matter, and it may be optimal to be in any of the three regions, depending on where the health state is relative to those two cutoff values. If D and A are close together, it will never be optimal to be in the coinsurance region (better to burn right though it and into the free healthcare of the out-of-pocket maximum region), and the cutoff Z_2 will determine whether the deductible or out-of-pocket maximum region is optimal. If the realized health state is negative, optimal spending will equal zero. In sum, optimal spending m^* conditional on health state realization l, moral hazard parameter ω and plan characteristics $\{D, C, O\}$ is given by

If $A - D > \omega/2$:

$$m^* = \begin{cases} \max(0, l) & l \leq Z_1, \\ l + \omega(1-C) & Z_1 < l \leq Z_3, \\ l + \omega & Z_3 < l; \end{cases}$$

If $A - D \leq \omega/2$:

$$m^* = \begin{cases} \max(0, l) & l \leq Z_2, \\ l + \omega & Z_2 < l. \end{cases}$$

Derivations are available upon request. A graphical example (of the case in which the coinsurance region is sufficiently large) is shown in Figure C.2b. All plans in our empirical setting have $A - D > \omega/2$ at reasonable values of ω.

With distributions of m^*_{kjtsd} in hand for each household, plan, year, and draw of β_{ks}, we can calculate households' expected utility from enrolling in each potential plan in

their choice set. We construct the numerical approximation to equation (3.10) using the quadrature weights W_d:

$$U_{kjts} = -\sum_{d=1}^{D} \left[W_d \ \exp\left(-\psi_k x_{kjts}(l_{kjtsd})\right) \right],$$

where the monetary payoff x is calculated as in equation (3.11). To avoid numerical issues arising from double-exponentiation, we estimate the model in terms of certainty equivalent units of U_{kjts}:

$$U^{CE}_{kjts} = \bar{x}_{kjts} - \frac{1}{\psi_k} \log \left(\sum_{d=1}^{D} \left[W_d \ \exp\left(-\psi_k (x_{kjts}(l_{kjtsd}) - \bar{x}_{kjts})\right) \right] \right),$$

where $\bar{x}_{kjts} = \mathbb{E}_d[x_{kjts}(l_{kjtsd})]$.

Choice probabilities, conditional on β_{kts}, are given by the standard logit formula:

$$L_{kjts} = \frac{\exp(U^{CE}_{kjts}/\sigma_\epsilon)}{\sum_{i \in \mathcal{J}_{kt}} \exp(U^{CE}_{kits}/\sigma_\epsilon)}.$$

The numerical approximation to the likelihood of the sequence of choices and healthcare spending amounts for a given household is given by

$$LL_k = \sum_{j=1}^{J} d_{kjt} \sum_{s=1}^{S} W_s \prod_{t=1}^{T} f_m(m_{kt}|\theta, \beta_{kts}, c_{jt}, \mathbf{X}_{kt}) L_{kjts},$$

where $d_{kjt} = 1$ if household k chose plan j in year t and zero otherwise. The simulated log-likelihood function for parameters θ is

$$SLL(\theta) = \sum_{k=1}^{K} \log\left(LL_k\right).$$

C.2.3 Recovering household-specific types

We assume that household types $\beta_{kt}(\theta) = \{\psi_k, \mu_{kt}, \omega_k\}$ are distributed according to equation (3.14). After estimating the model and obtaining $\hat{\theta}$, we want to use each household's observed choices to back out which type they themselves are likely to be. Let $g(\beta|\hat{\theta})$ denote the population distribution of types. Let $h(\beta|\hat{\theta}, y)$ denote the density of β conditional on parameters $\hat{\theta}$ and a sequence of observed healthcare spending amounts and plan chocies y.

Using what Revelt and Train (2001) term the "conditioning of individual tastes" method, we recover households' posterior distribution of β using Bayes' rule:

$$h(\beta|\hat{\theta}, y) = \frac{p(y|\beta)g(\beta|\hat{\theta})}{p(y|\hat{\theta})}.$$

Taking the numerical approximations, $p(y|\hat{\theta})$ is simply the household-specific likelihood function LL_k for an observed sequence of spending amounts and choices, $g(\beta|\hat{\theta})$ is the quadrature weights W_s on each simulated point, and $p(y|\beta)$ is the *conditional* household likelihood:

$$LL_{ks} = \sum_{j=1}^{J} d_{kjt} \prod_{t=1}^{T} f_m(m_{kt}|\theta, \beta_{ks}, c_{jt}, \mathbf{X}_{kt}) L_{kjts}.$$

Taken together, the numerical approximation to each household's posterior distribution of unobserved heterogeneity is given by

$$h_{ks}(\beta|\hat{\theta}, y_k) = \frac{LL_{ks}\, W_s}{LL_k},$$

where $\sum_s h_{ks}(\beta|\hat{\theta}, y_k) = 1$.

We use these household specific distributions over types to calculated expected quantities of interest for each household. In particular, we calculate WTP_{kjt} and SS_{kjt} as

$$WTP_{kjt} = \sum_s h_{ks}(\beta|\hat{\theta}, y_k) WTP_{kjts},$$

$$SS_{kjt} = \sum_s h_{ks}(\beta|\hat{\theta}, y_k) SS_{kjts}.$$

C.2.4 Joint distribution of household types

The joint distribution of household types is of central importance to this paper. Here, we investigate the distribution implied by our primary estimates in column 3 of Table 3.4. For each household, we first calculate the expectation of their type with respect to their posterior

distribution of unobservable heterogeneity:

$$\psi_k = \sum_s h_{ks}(\beta|\hat{\theta}, y_k)\psi_{ks},$$
$$\omega_k = \sum_s h_{ks}(\beta|\hat{\theta}, y_k)\omega_{ks}.$$

In place of μ_{kt}, a more relevant measure of household health is the expected health state, or in other words, expected total unavoidable spending. Using the expectation of a shifted lognormal variable and price parameter $\phi = 1$, the expected health state $\bar{l}_{kt}$ is given by

$$\bar{l}_{kt} = \sum_s h_{ks}(\beta|\hat{\theta}, y_k)(\exp(\mu_{kts} + \frac{\sigma_{kt}^2}{2}) - \kappa_{kt}).$$

To limit our focus to one type for each household, we look at $\bar{l}_{kt}$ for the first year each household appears in the data. Figure C.3 presents the joint distribution of household types along the dimensions of risk aversion (ψ), moral hazard (ω), and expected health state ($\log(\mathbb{E}[\text{Health state}])$). We measure the health state on a log scale for readability.

Table C.1: *Sample Construction*

Criteria	2009	2010	2011	2012	2013
Individuals in membership file	161,502	162,363	156,113	156,042	157,799
Not eligible for coverage	7,370	8,265	8,422	8,719	8,388
Retiree, COBRA, or oldest member over 65	13,180	12,567	12,057	11,603	11,840
Partial year coverage	17,115	18,649	19,283	21,281	23,074
Covered by multiple plans	1,447	1,947	2,038	2,239	2,336
Opted out	3,241	4,205	4,321	4,576	4,529
Not in intact family	8,389	9,188	9,181	8,925	10,265
No prior year of data	6,175	3,947	2,455	3,104	3,702
Missing premium or contribution data	25,653	28,466	22,755	23,284	30,401
Final total	78,932	75,129	75,601	72,311	63,264

Notes: This table shows counts of individuals dropped due to each sample selection criterion. Drops are made in the order in which criteria appear.

Table C.2: *Plan Characteristics*

2008					
Plan	AV	Insurer Premium ($)	Deductible ($)	OOP Max. ($)	Market Share
Kaiser - 1	0.97	10,567	0	1,200	0.07
Kaiser - 2	0.96	10,098	0	2,000	0.10
Moda - 1	0.92	11,955	300	500	0.28
Moda - 2	0.89	11,481	300	1,000	0.06
Moda - 3	0.88	10,841	600	1,000	0.11
Moda - 4	0.86	10,382	900	1,500	0.07
Moda - 5	0.82	9,615	1,500	2,000	0.12
Moda - 6	0.78	8,689	3,000	3,000	0.03
Moda - 7	0.68	6,643	3,000	10,000	0.00
Providence - 1	0.96	11,564	900	1,200	0.14
Providence - 2	0.95	11,475	900	2,000	0.02

2010					
Plan	AV	Insurer Premium ($)	Deductible ($)	OOP Max. ($)	Market Share
Kaiser - 1	0.96	12,537	0	2,400	0.17
Kaiser - 2	0.95	12,150	0	3,000	0.03
Moda - 1	0.89	17,042	600	1,200	0.10
Moda - 2	0.86	15,817	600	1,500	0.01
Moda - 3	0.85	14,344	600	1,800	0.17
Moda - 4	0.84	12,877	900	2,000	0.12
Moda - 5	0.82	11,781	1,500	2,000	0.21
Moda - 6	0.78	10,596	3,000	3,000	0.09
Moda - 7	0.75	8,083	3,000	10,000	0.02
Providence - 1	0.91	18,121	1,200	1,200	0.04
Providence - 2	0.89	17,647	1,800	1,800	0.01

Notes: Actuarial value (AV) is calculated as the ratio of average insured spending to average total spending among all households, using counterfactual calculations of insured spending for households that did not choose a certain plan. Insurer premium reflects the premium negotiated between OEBB and the insurer. The deductible and out-of-pocket maximum shown are for in-network services for a family household.

Table C.2: *Plan Characteristics, cont.*

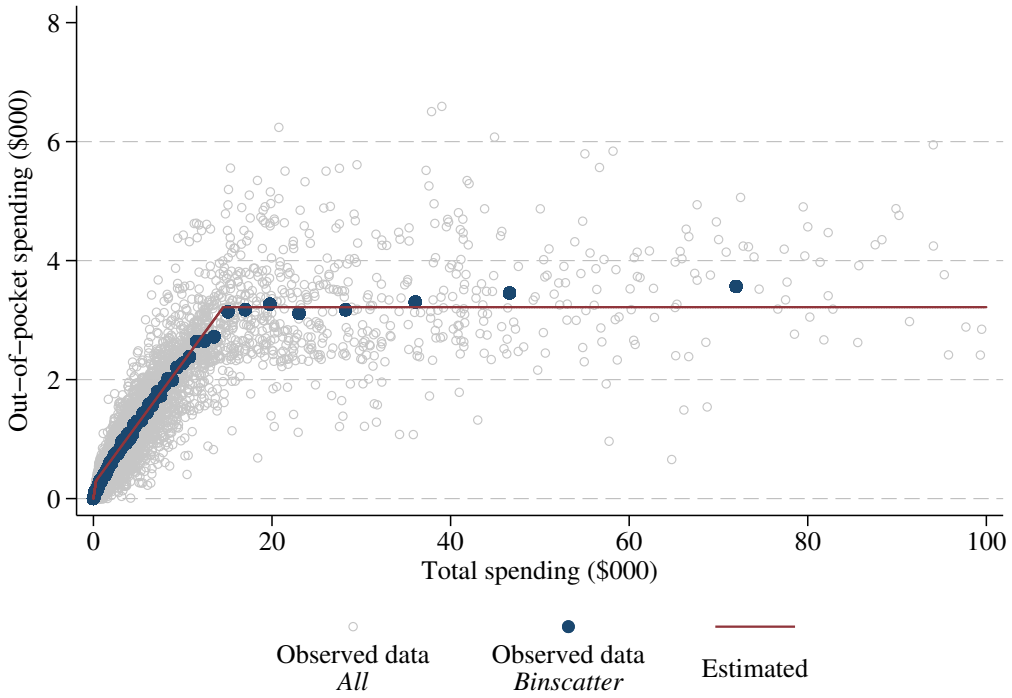
2011 Plan	AV	Insurer Premium ($)	Deductible ($)	OOP Max. ($)	Market Share
Kaiser - 1	0.95	11,958	0	2,400	0.16
Kaiser - 2	0.92	10,954	300	4,000	0.04
Moda - 1	0.86	16,900	600	4,500	0.06
Moda - 2	0.84	13,405	900	6,000	0.00
Moda - 3	0.84	13,726	900	6,000	0.15
Moda - 4	0.83	12,261	1,200	6,300	0.09
Moda - 5	0.82	11,021	1,500	6,600	0.24
Moda - 6	0.78	9,481	3,000	6,600	0.15
Moda - 7	0.75	8,445	3,000	10,000	0.05
Providence - 1	0.87	16,168	300	3,600	0.02
Providence - 2	0.84	15,090	900	6,000	0.00

2012 Plan	AV	Insurer Premium ($)	Deductible ($)	OOP Max. ($)	Market Share
Kaiser - 1	0.95	14,508	0	2,400	0.18
Kaiser - 2	0.93	13,283	450	4,000	0.04
Moda - 1	0.87	20,029	600	4,500	0.06
Moda - 2	0.85	15,469	900	6,000	0.01
Moda - 3	0.85	16,616	900	6,000	0.12
Moda - 4	0.84	15,039	1,200	6,300	0.06
Moda - 5	0.83	13,707	1,500	6,600	0.22
Moda - 6	0.79	12,051	3,000	6,600	0.17
Moda - 7	0.76	9,082	3,000	10,000	0.11

2013 Plan	AV	Insurer Premium ($)	Deductible ($)	OOP Max. ($)	Market Share
Kaiser - 1	0.95	15,369	0	3,000	0.20
Kaiser - 2	0.94	13,950	600	4,400	0.03
Moda - 1	0.87	21,285	600	6,000	0.03
Moda - 2	0.85	17,055	1,050	7,200	0.08
Moda - 3	0.84	14,234	1,500	7,800	0.22
Moda - 4	0.82	13,211	2,250	8,400	0.06
Moda - 5	0.80	12,362	3,000	9,000	0.11
Moda - 6	0.78	11,337	3,750	12,000	0.05
Moda - 7	0.77	9,276	3,000	10,000	0.13
Moda - 8	0.76	10,250	4,500	15,000	0.05

Notes: Actuarial value (AV) is calculated as the ratio of average insured spending to average total spending among all households, using counterfactual calculations of insured spending for households that did not choose a certain plan. Insurer premium reflects the premium negotiated between OEBB and the insurer. The deductible and out-of-pocket maximum shown are for in-network services for a family household.

Figure C.1: *Example of Plan Cost Sharing Features Estimation*



Notes: The figure shows the data used to estimate the cost sharing features of Moda Plan 3 for individual households in 2012. Each gray dot represents a household. The blue dots are a binscatter plot of the gray data, using 100 data points. The basic cost sharing features of the plan are a deductible of $300, non-specialist coinsurance rate of 20 percent, and in-network out-of-pocket maximum of $2,000. The red line shows the "estimated" cost sharing schedule of the plan that minimizes the sum of squared errors between predicted and observed out-of-pocket spending. The estimated coinsurance rate is 20.5 percent and the estimated out-of-pocket maximum is $3,218.

Table C.3: *Estimated Plan Characteristics*

2009	Individuals			Families		
Plan	Ded.	Coins.	OOP Max.	Ded.	Coins.	OOP Max.
Kaiser - 1	0	0.03	564	0	0.03	645
Kaiser - 2	0	0.03	684	0	0.04	760
Kaiser - 3	0	0.03	734	0	0.04	791
Moda - 1	100	0.10	1,613	300	0.10	2,009
Moda - 2	100	0.18	1,922	300	0.15	2,662
Moda - 3	200	0.20	2,081	600	0.15	3,062
Moda - 4	300	0.19	2,796	900	0.15	3,835
Moda - 5	500	0.22	3,164	1,500	0.16	4,296
Moda - 6	1,000	0.22	3,713	3,000	0.12	5,422
Moda - 7	1,500	0.42	4,693	3,000	0.30	8,086
Providence - 1	300	0.02	790	900	0.00	900
Providence - 2	300	0.03	867	900	0.00	986
Providence - 3	300	0.04	1,116	900	0.01	1,296

2010	Individuals			Families		
Plan	Ded.	Coins.	OOP Max.	Ded.	Coins.	OOP Max.
Kaiser - 1	0	0.03	697	0	0.04	805
Kaiser - 2	0	0.04	820	0	0.05	885
Moda - 1	200	0.14	2,526	600	0.12	3,430
Moda - 2	200	0.21	2,846	600	0.18	3,967
Moda - 3	200	0.21	3,189	600	0.18	4,299
Moda - 4	300	0.22	3,109	900	0.18	4,079
Moda - 5	500	0.22	3,321	1,500	0.16	4,572
Moda - 6	1,000	0.22	3,844	3,000	0.12	5,684
Moda - 7	1,500	0.19	4,913	3,000	0.15	7,579
Providence - 1	400	0.05	1,523	1,200	0.02	1,851
Providence - 2	600	0.06	1,998	1,800	0.02	2,473

2011	Individuals			Families		
Plan	Ded.	Coins.	OOP Max.	Ded.	Coins.	OOP Max.
Kaiser - 1	0	0.04	883	0	0.06	974
Kaiser - 2	100	0.06	1,340	300	0.06	1,831
Moda - 1	200	0.22	2,608	600	0.18	4,316
Moda - 2	300	0.22	3,201	900	0.17	5,094
Moda - 3	300	0.22	3,246	900	0.17	5,202
Moda - 4	400	0.22	3,324	1,200	0.17	5,367
Moda - 5	500	0.22	3,529	1,500	0.16	5,727
Moda - 6	1,000	0.22	4,061	3,000	0.13	6,728
Moda - 7	1,500	0.21	4,914	3,000	0.15	7,663
Providence - 1	100	0.18	2,164	300	0.16	3,496
Providence - 2	300	0.15	2,911	900	0.13	4,378

Notes: Table shows plan deductibles (Ded.), estimated coinsurance rates (Coins.), and estimated out-of-pocket maximums (OOP Max.).

Table C.3: *Estimated Plan Characteristics, cont.*

2012	Individuals			Families		
Plan	Ded.	Coins.	OOP Max.	Ded.	Coins.	OOP Max.
Kaiser - 1	0	0.04	911	0	0.06	995
Kaiser - 2	150	0.07	1,709	450	0.05	2,160
Moda - 1	200	0.21	2,571	600	0.17	4,154
Moda - 2	300	0.21	3,187	900	0.17	4,981
Moda - 3	300	0.20	3,218	900	0.17	5,025
Moda - 4	400	0.21	3,291	1,200	0.16	5,104
Moda - 5	500	0.21	3,493	1,500	0.16	5,498
Moda - 6	1,000	0.21	4,000	3,000	0.12	6,608
Moda - 7	1,500	0.21	4,927	3,000	0.15	7,662

2013	Individuals			Families		
Plan	Ded.	Coins.	OOP Max.	Ded.	Coins.	OOP Max.
Kaiser - 1	0	0.04	911	0	0.06	1,040
Kaiser - 2	200	0.03	867	600	0.01	951
Moda - 1	200	0.20	3,237	600	0.17	4,893
Moda - 2	350	0.20	3,842	1,050	0.16	5,647
Moda - 3	500	0.20	4,175	1,500	0.15	6,160
Moda - 4	750	0.20	4,704	2,250	0.14	6,989
Moda - 5	1,000	0.19	5,186	3,000	0.12	7,714
Moda - 6	1,250	0.19	6,414	3,750	0.12	9,187
Moda - 7	1,500	0.21	4,865	3,000	0.15	7,650
Moda - 8	1,500	0.19	7,620	4,500	0.11	10,614

Notes: The table shows plan deductibles (Ded.), estimated coinsurance rates (Coins.), and estimated out-of-pocket maximums (OOP Max.).

Table C.4: *Household Summary Statistics (2008)*

Sample demographics	2008
Number of households	45,012
Number of enrollees	116,267
Employee age, mean (med.)	45.5
	(47.0)
Enrollee age, mean (med.)	38.2
	(35.8)
Enrollee percent female	0.53
Premiums	
Employee premium ($), mean (med.)	596
	(0)
Full premium ($), mean (med.)	10,107
	(10,605)
Household health spending	
Total spending ($), mean (med.)	9,956
	(4,485)
OOP spending ($), mean (med.)	957
	(620)
Household structure (percent)	
Individual	0.25
Family	0.75
Region (percent)	
Portland-Salem	0.64
Eugene-Medford	0.26
Bend-Spokane-Boise	0.10

Notes: Summary statistics are shown for households in the 2008 analysis sample used in our descriptive analyses. Enrollees are employees plus their family members. Statistics about premiums are for households' chosen plans, as opposed to for all possible plans. Sample medians are shown in parentheses.

Table C.5: *Plan Choice Logit Model (equation (3.7))*

	2008	2009	2010	2011	2012	2013
Employee premium ($000)	-0.789	-0.674	-0.505	-0.372	-0.515	-0.490
	(0.017)	(0.014)	(0.008)	(0.010)	(0.008)	(0.008)
HRA/HSA contrib. ($000)	0.112		0.358	0.134	0.269	0.534
	(0.759)		(0.044)	(0.024)	(0.019)	(0.015)
Vision/dental contrib. ($000)	0.654	0.408	0.480	0.794	0.553	0.710
	(0.021)	(0.022)	(0.019)	(0.017)	(0.017)	(0.017)
Kaiser - 1	-0.771	-0.728				
	(0.026)	(0.030)				
Kaiser - 2	-1.287	-1.112	-0.846	-0.469	-0.375	-0.074
	(0.031)	(0.032)	(0.034)	(0.035)	(0.034)	(0.044)
Kaiser - 3		-1.563	-1.042	-0.985	-1.629	-1.820
		(0.384)	(0.056)	(0.051)	(0.048)	(0.058)
Moda - 1	0.000†	0.000†	0.000†	0.000†	0.000†	0.000†
Moda - 2	-1.113	-1.184	-0.911	-2.088	-2.578	-0.593
	(0.026)	(0.032)	(0.058)	(0.163)	(0.072)	(0.045)
Moda - 3	-1.226	-1.110	-0.518	-0.373	-0.389	-0.957
	(0.022)	(0.025)	(0.029)	(0.034)	(0.033)	(0.046)
Moda - 4	-1.751	-1.540	-1.356	-1.192	-1.554	-2.261
	(0.028)	(0.030)	(0.034)	(0.037)	(0.039)	(0.055)
Moda - 5	-1.951	-1.881	-1.341	-0.878	-0.999	-2.391
	(0.034)	(0.037)	(0.040)	(0.039)	(0.037)	(0.055)
Moda - 6	-2.785	-2.871	-2.205	-1.406	-1.917	-3.182
	(0.048)	(0.051)	(0.050)	(0.043)	(0.046)	(0.065)
Moda - 7	-4.391	-4.260	-3.388	-1.959	-3.007	-3.492
	(0.098)	(0.098)	(0.074)	(0.050)	(0.060)	(0.073)
Moda - 8						-3.679
						(0.068)
Providence - 1	0.001	0.048	0.135	-0.778		
	(0.019)	(0.028)	(0.038)	(0.053)		
Providence - 2	-0.600	-0.314				
	(0.043)	(0.049)				
Providence - 3		-0.048	-0.159	-0.939		
		(0.078)	(0.083)	(0.436)		
Number of observations	163,431	121,744	116,541	114,527	163,278	163,683

Notes: This table presents the parameter estimates from the conditional logit model described by equation (3.7), presented separately for each year. The unit of observation is the household-plan. Moda plan 1 (the highest coverage Moda plan) is the omitted plan.
†By normalization.

Table C.6: *Plan Menu Generosity and Household Health*

	2008	2009	2010	2011	2012	2013
Household Risk Score	-0.006	0.017	0.020	0.002	0.006	0.000
	(0.039)	(0.016)	(0.011)*	(0.009)	(0.010)	(0.012)
Family Type						
Employee Alone	0.000†	0.000†	0.000†	0.000†	0.000†	0.000†
Employee + Spouse	-1.389	-1.369	-1.498	-1.040	-1.626	-1.612
	(0.077)***	(0.040)***	(0.029)***	(0.025)***	(0.026)***	(0.031)***
Employee + Child	-0.542	-0.634	-0.907	-0.616	-1.092	-0.937
	(0.084)***	(0.053)***	(0.039)***	(0.031)***	(0.031)***	(0.037)***
Employee + Family	-1.792	-1.882	-1.804	-1.306	-2.147	-2.102
	(0.064)***	(0.037)***	(0.028)***	(0.023)***	(0.025)***	(0.029)***
Dependent variable mean	88.7	88.5	84.6	82.7	83.3	82.6
R^2	0.020	0.084	0.154	0.115	0.242	0.220
Number of observations	37,666	31,074	29,538	29,279	27,897	24,283

Notes: The dependent variable is plan menu generosity as measured by predicted actuarial value conditional on choosing Moda, $\widehat{AV}_{d,Moda}$, as estimated by the logit model in equation (3.7) and calculated according to equation (3.8). $\widehat{AV}_{d,Moda}$ is multiplied by 100 to increase parameter magnitudes. The level of observation is the household. Household risk score is the mean risk score among all individuals in a household, and has been z-scored such that the variable has a mean of zero and a standard deviation of one within each year. * p<0.10, ** p<0.05, *** p<.01.
†By normalization

Table C.7: *Explaining Plan Menu Generosity: 2008*

	(1)	(2)	(3)	(4)
Household Risk Score	-0.006	0.016	0.011	0.025
	(0.039)	(0.039)	(0.038)	(0.040)
Family Type				
Employee Alone	0.000†	0.000†	0.000†	0.000†
Employee + Spouse	-1.389	-1.374	-1.251	-1.085
	(0.077)***	(0.083)***	(0.083)***	(0.085)***
Employee + Child	-0.542	-0.535	-0.478	-0.462
	(0.084)***	(0.085)***	(0.084)***	(0.082)***
Employee + Family	-1.792	-1.819	-1.688	-1.437
	(0.064)***	(0.071)***	(0.071)***	(0.074)***
Part-time		-0.428	-0.448	-0.867
		(0.133)***	(0.133)***	(0.139)***
Occupation Type				
Admin.		-1.745	-1.883	-2.685
		(0.455)***	(0.459)***	(0.501)***
Classified		-0.598	-0.469	-0.155
		(0.283)**	(0.414)	(0.457)
Comm. Coll. Fac.		0.553	1.138	1.044
		(0.287)*	(0.430)***	(0.470)**
Comm. Coll. Non-Fac.		0.671	0.457	0.077
		(0.288)**	(0.288)	(0.302)
Confidential		-2.759	-2.883	-3.133
		(0.855)***	(0.856)***	(0.915)***
Licensed		0.001	1.645	1.628
		(0.278)	(0.459)***	(0.505)***
Substitute		-11.051	-9.312	-9.354
		(0.283)***	(0.457)***	(0.496)***
Union Affiliation				
AFT			0.251	-0.398
			(0.374)	(0.432)
IAFE			0.758	1.222
			(0.404)*	(0.458)***
OACE			2.671	1.617
			(0.389)***	(0.449)***
OEA			-1.799	-1.765
			(0.434)***	(0.491)***
OSEA			-0.086	-0.426
			(0.395)	(0.449)
District characteristics				
ln(HPI)				-0.876
				(0.085)***
Pct. Republican				-14.077
				(0.467)***
Dependent variable mean	88.7	89.0	89.1	98.3
R^2	0.020	0.031	0.046	0.073
Number of observations	37,666	37,666	37,666	35,698

Notes: The dependent variable is plan menu generosity as measured by predicted actuarial value conditional on choosing Moda, $\widehat{AV}_{d,Moda}$, as estimated by the logit model in equation (3.7) and calculated according to equation (3.8). $\widehat{AV}_{d,Moda}$ is multiplied by 100 to increase parameter magnitudes. The level of observation is the household. Household risk score is the mean risk score among all individuals in a household, and has been z-scored such that the variable has a mean of zero and a standard deviation of one within each year. * p<0.10, ** p<0.05, *** p<.01. †By normalization

Table C.8: *Conditional Logit Model of Household Plan Choice in 2008*

	Ind. Q_1	Fam. Q_1	Ind. Q_2	Fam. Q_2	Ind. Q_3	Fam. Q_3	Ind. Q_4	Fam. Q_4
Employee premium ($000)	-1.602***	-1.014***	-1.345***	-1.019***	-1.401***	-0.949***	-1.302***	-0.870***
	(0.128)	(0.047)	(0.114)	(0.049)	(0.113)	(0.053)	(0.108)	(0.056)
Vision/dental contrib. ($000)	1.301***	0.943***	1.254***	0.884***	1.089***	0.621***	1.042***	0.495***
	(0.092)	(0.061)	(0.094)	(0.065)	(0.094)	(0.071)	(0.099)	(0.076)
HSA/HRA contrib. ($000)				-6.871		2.774***		-6.703
				(318.561)		(1.068)		(526.706)
Kaiser - 1	-0.074	1.351**	-1.452**	-0.856	1.069	0.863	2.149***	0.525
	(0.420)	(0.531)	(0.671)	(0.747)	(0.799)	(0.918)	(0.782)	(0.801)
Kaiser - 2	0.575	1.765***	-0.960	-0.278	1.483*	1.376	2.468***	1.135
	(0.410)	(0.517)	(0.657)	(0.731)	(0.791)	(0.899)	(0.774)	(0.789)
Moda - 1	0.000†	0.000†	0.000†	0.000†	0.000†	0.000†	0.000†	0.000†
Moda - 2	-1.175***	-0.425***	-1.077***	-1.011***	-0.498*	-0.571**	-0.644**	-0.930***
	(0.185)	(0.161)	(0.242)	(0.215)	(0.260)	(0.254)	(0.270)	(0.214)
Moda - 3	-0.865***	-0.298	-0.880***	-1.162***	-0.290	-0.395	-0.108	-0.810**
	(0.202)	(0.240)	(0.332)	(0.334)	(0.372)	(0.399)	(0.383)	(0.333)
Moda - 4	-1.265***	-0.331	-1.535***	-1.719***	-0.370	-0.535	-0.100	-1.194**
	(0.280)	(0.349)	(0.477)	(0.488)	(0.534)	(0.584)	(0.553)	(0.486)
Moda - 5	-1.083***	-0.065	-1.419**	-1.896**	0.386	-0.119	0.623	-1.029
	(0.407)	(0.527)	(0.713)	(0.740)	(0.805)	(0.885)	(0.832)	(0.737)
Moda - 6	-1.053*	-0.086	-1.903*	-2.678**	0.515	-0.517	1.390	-1.634
	(0.592)	(0.770)	(1.048)	(1.084)	(1.171)	(1.295)	(1.210)	(1.082)
Moda - 7	-2.060**	0.093	-3.330*	-5.027***	0.880	-0.940	1.879	-1.986
	(0.997)	(1.304)	(1.757)	(1.854)	(1.968)	(2.225)	(2.058)	(1.842)
Providence - 1	-0.251	1.141*	-1.448*	-0.696	0.474	2.210**	0.840	-0.613
	(0.566)	(0.659)	(0.863)	(0.850)	(0.920)	(0.938)	(0.922)	(0.747)
Providence - 2	0.300	1.533**	-1.022	-0.194	1.017	2.809***	1.215	-0.121
	(0.536)	(0.639)	(0.836)	(0.830)	(0.894)	(0.915)	(0.915)	(0.728)
Number of osbervations	8,487	25,054	8,367	25,416	8,285	25,393	8,077	25,326

Notes: The table presents the results of estimating equation (3.7) separately by quartile of household risk score within individual and family households in 2008. The columns indicate which sample is being used: Individuals (Ind.) versus families (Fam.) and the household risk quartile Q_n, where Q_4 is the sickest households. The omitted plan fixed effect is for Moda plan 1 (the most generous Moda plan). The coefficient on employee premium (measured in thousands of dollars) is normalized to -1.

†By normalization.

Table C.9: *Spending Distributions and Moda Plan Characteristics, 2008*

Panel A: Total Spending Distributions by Risk Quartile

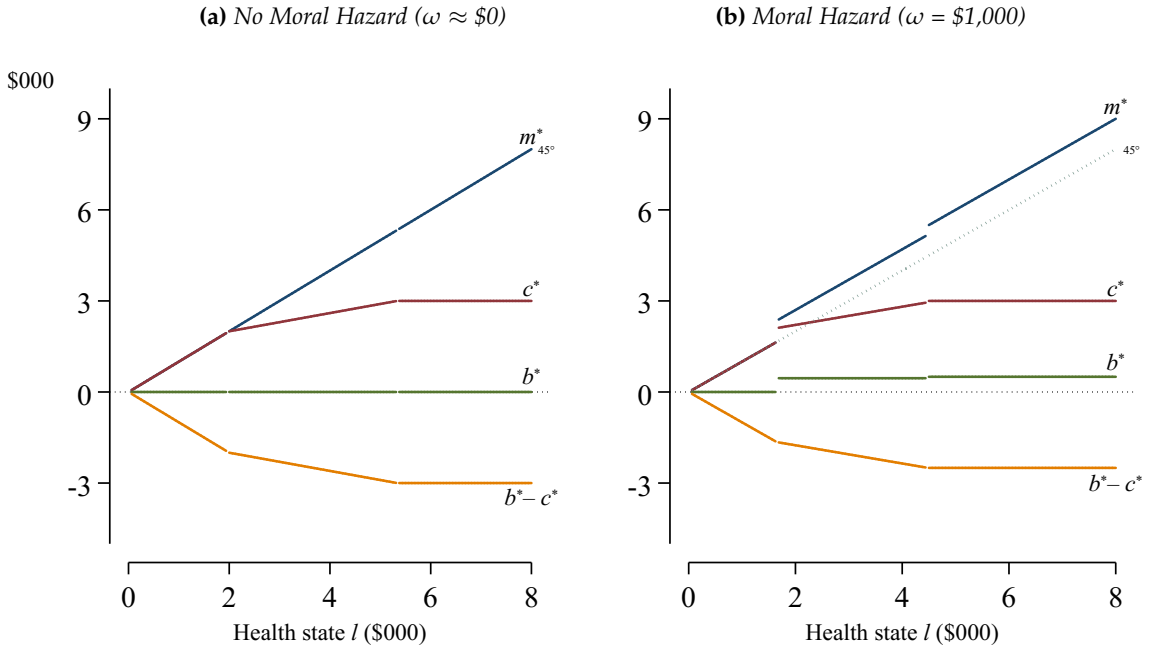
	Percentile of total spending				
Risk quartile	10th	25th	50th	75th	90th
Individuals					
*Q*1	0	30	381	851	1,454
*Q*2	293	721	1,286	1,984	3,025
*Q*3	782	1,688	2,861	4,266	5,987
*Q*4	1,869	4,134	7,155	12,765	21,240
Families					
*Q*1	418	985	1,959	3,508	6,718
*Q*2	1,489	2,567	4,212	6,584	10,984
*Q*3	3,373	5,261	7,811	11,745	17,301
*Q*4	5,096	9,820	15,401	22,637	29,615

Panel B: Plan Characteristics

	Moda plan						
	Plan 1	Plan 2	Plan 3	Plan 4	Plan 5	Plan 6	Plan 7
Individuals							
Deductible	100	100	200	300	500	1,000	1,500
OOP Max.	500	1,000	1,000	1,500	2,000	2,000	5,000
Families							
Deductible	300	300	600	900	1,500	3,000	3,000
OOP Max.	500	1,000	1,000	1,500	2,000	3,000	10,000

Notes: This table shows the distributions of household realized total healthcare spending and the plan characteristics of Moda plans in 2008. Panel A shows the spending distributions, by quartile of household risk score within Individual and Family households. Panel B shows the in-network deductible and out-of-pocket maximum (OOP Max.) for each of the Moda plans.

Figure C.2: *Healthcare Spending Choice Example*



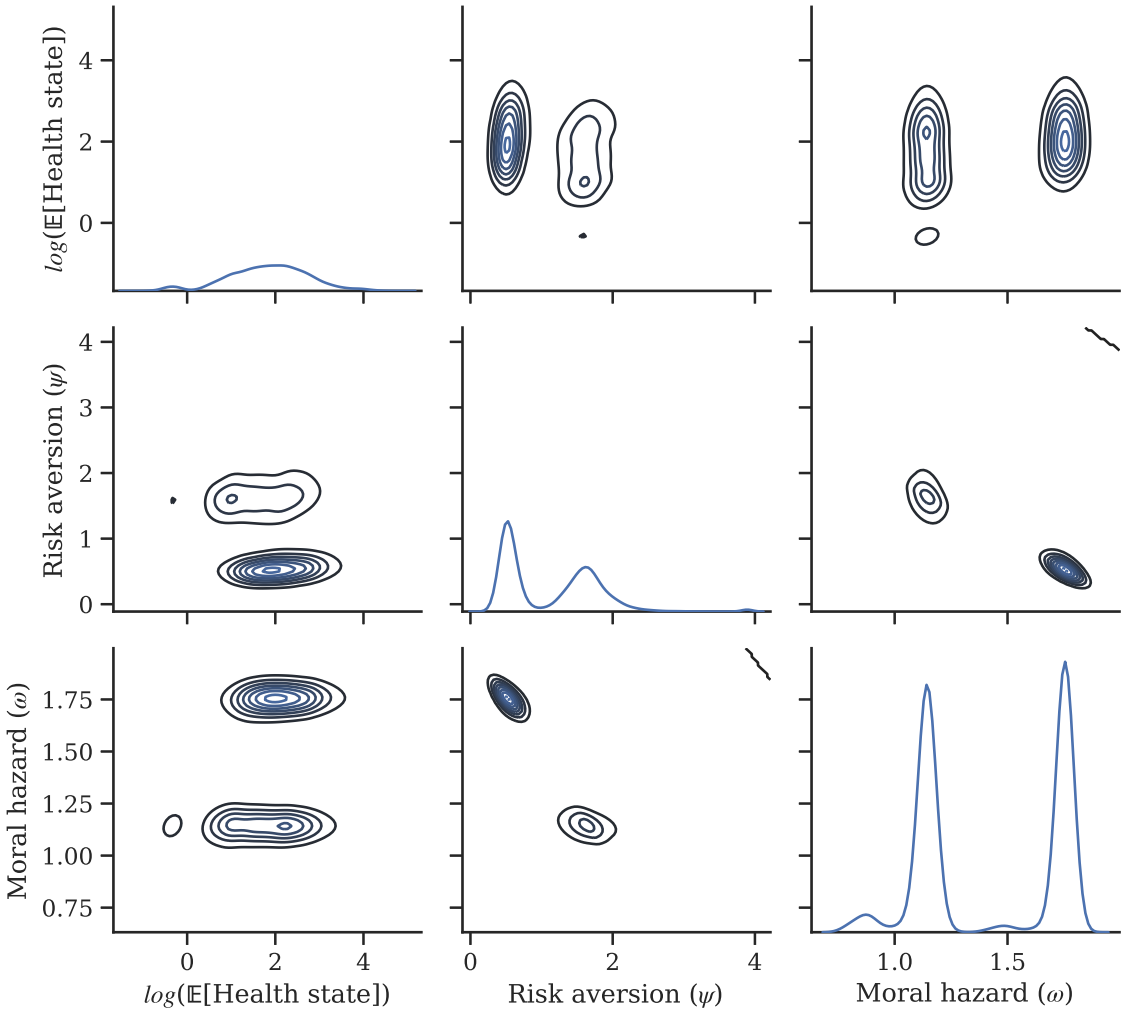
Notes: The figure shows optimal healthcare spending behavior predicted by our specification of household preferences over healthcare utilization (given in equation (3.9)). Optimal behavior is calculated assuming enrollment in an insurance contract with a deductible of $2,000, a coinsurance rate of 30%, and an out-of-pocket maximum of $3,000. Predicted behavior is shown under no moral hazard in panel (a) and under some moral hazard (ω = $1,000) in panel (b). Possible health state realizations are plotted on the horizontal axis. Optimal total healthcare spending m^* is shown for each health state; when there is no moral hazard, it is optimal to set total spending equal to the health state. Optimal healthcare spending m^* implies some correspondingly optimal out-of-pocket costs c^*, utility from healthcare utilization b^*, and *net* utility from healthcare utilization $b^* - c^*$. Conditional on plan choice, households face a lottery over net utility $b^* - c^*$, where the uncertainty is with respect to their distribution of health states.

Table C.10: *Additional Demand Model Parameter Estimates*

	(1)		(2)		(3)	
Variable	Parameter	Std. Err.	Parameter	Std. Err.	Parameter	Std. Err.
Insurer fixed effects						
Kaiser * (Age−40) ($000s)	−0.073	0.005	−0.078	0.005	−0.071	0.005
Providence * (Age−40) ($000s)	−0.073	0.008	−0.122	0.009	−0.074	0.008
Kaiser * **1**[Children] ($000s)	−1.608	0.119	−1.509	0.120	−0.546	0.124
Providence * **1**[Children] ($000s)	−1.373	0.174	−2.116	0.199	−0.480	0.177
Kaiser * Region 1 ($000s)	−1.692	0.093	−1.477	0.091	−1.976	0.095
Kaiser * Region 2 ($000s)	−5.112	0.254	−4.949	0.254	−5.343	0.252
Providence * Region 1 ($000s)	−4.420	0.156	−3.899	0.158	−4.530	0.159
Providence * Region 2 ($000s)	−5.727	0.211	−5.301	0.213	−5.701	0.213
Providence * Region 3 ($000s)	−5.153	0.233	−4.716	0.235	−5.633	0.234
Health state distributions						
κ	0.167	0.002				
κ * Risk QT 1			0.123	0.004	0.184	0.000
κ * Risk QT 2			0.174	0.004	0.201	0.000
κ * Risk QT 3			0.162	0.004	0.302	0.000
κ * Risk QT 4			0.095	0.037	0.182	0.022
κ * Risk QT <4 * Risk score			0.156	0.023	0.270	0.017
μ	0.618	0.006				
μ * Female 18–30			0.142	0.014	0.059	0.016
μ * Age $<$ 18			0.020	0.014	−0.015	0.016
μ * Risk QT 1			−0.267	0.025	−0.421	0.021
μ * Risk QT 2			0.555	0.012	0.212	0.010
μ * Risk QT 3			0.709	0.008	0.420	0.007
μ * Risk QT 4			1.355	0.015	1.279	0.013
μ * Risk QT <4 * Risk score			1.025	0.016	1.184	0.018
μ * Risk QT 4 * Risk score			0.311	0.005	0.326	0.004
σ	1.117	0.002				
σ * Risk QT 1			1.408	0.010	1.450	0.008
σ * Risk QT 2			1.129	0.005	1.392	0.004
σ * Risk QT 3			1.067	0.003	1.244	0.003
σ * Risk QT 4			0.992	0.005	1.047	0.005

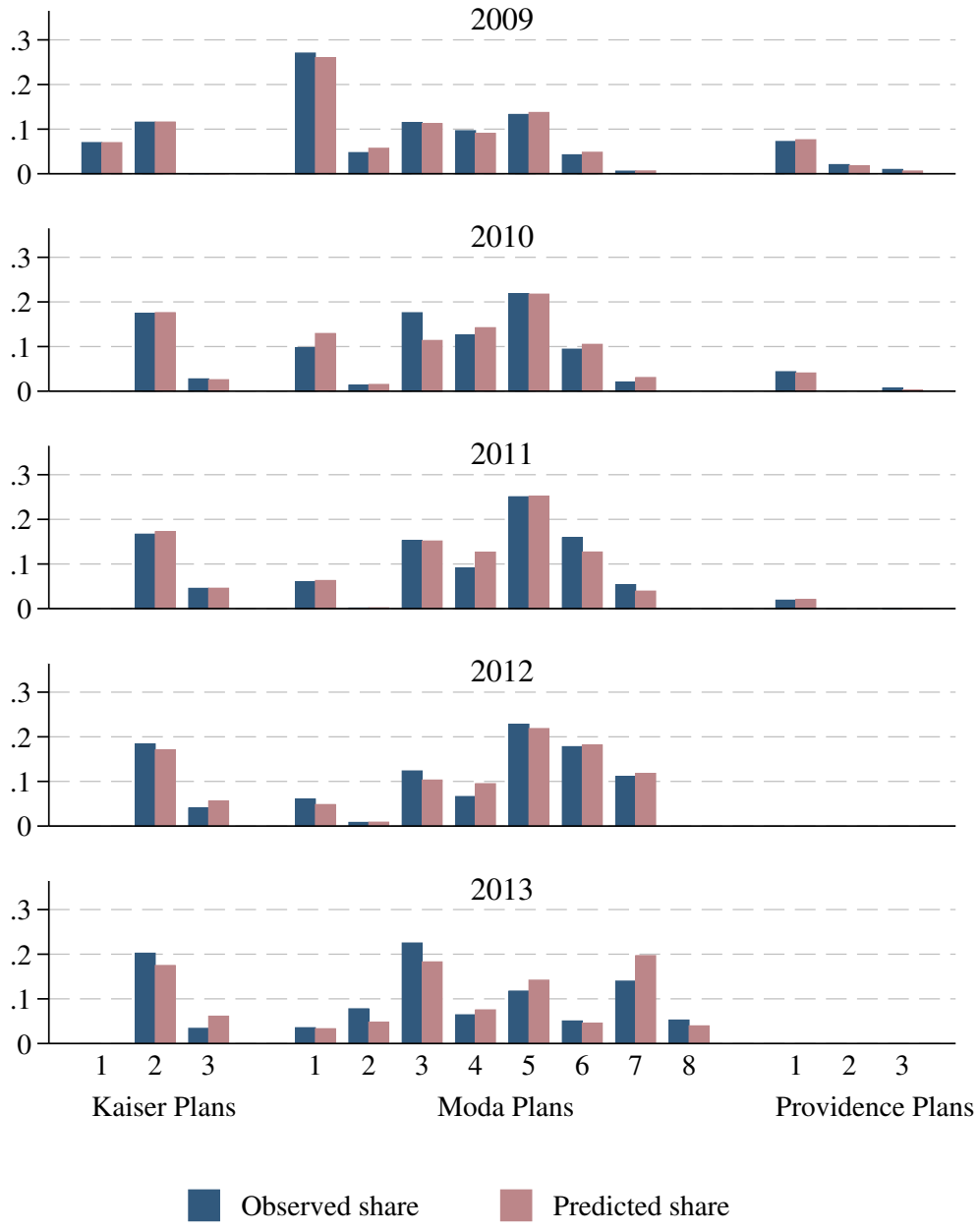
Notes: This table presents the parameter estimates that were not presented in the main table (Table 3.4), including insurer fixed effects and the health state distribution parameters. Column 1 estimates a model without individual observable heterogeneity. "Risk QT" refers to an indicator for an individual's risk quartile, where "Risk QT 4" is the sickest individuals. To make non-interacted coefficients more readily interpretable, Age is adjusted to be (Age−40). Higher risk scores correspond to worse predicted health.

Figure C.3: *Joint Distribution of Household Types*



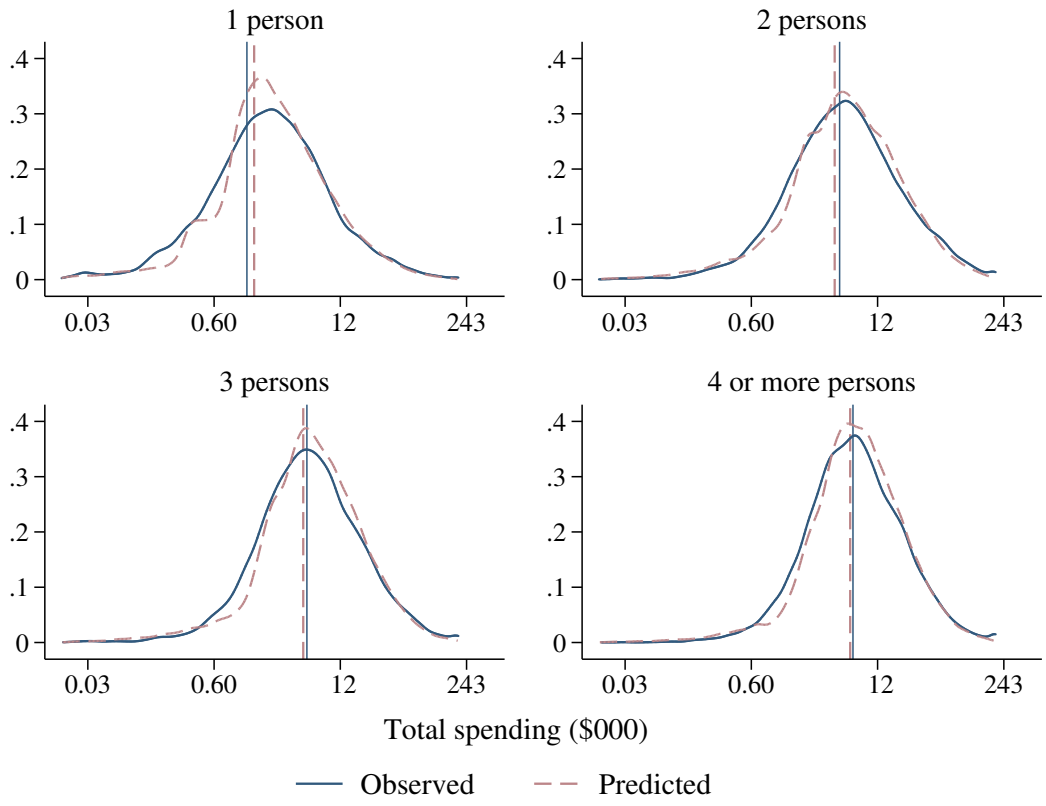
Notes: The figure shows the joint distribution of household types implied by the estimates in column 3 of Table 3.4. Households are assigned to a particular type according to the procedure described in Section C.2.4. Because expected health shock can vary over years within a household, for the purposes of this figure we use the first year a household appears in the data. Expected health state ($\mathbb{E}$[Health state]) is equal to a household's expected total unavoidable healthcare spending.

Figure C.4: *Model Fit: Plan Choices Year by Year*



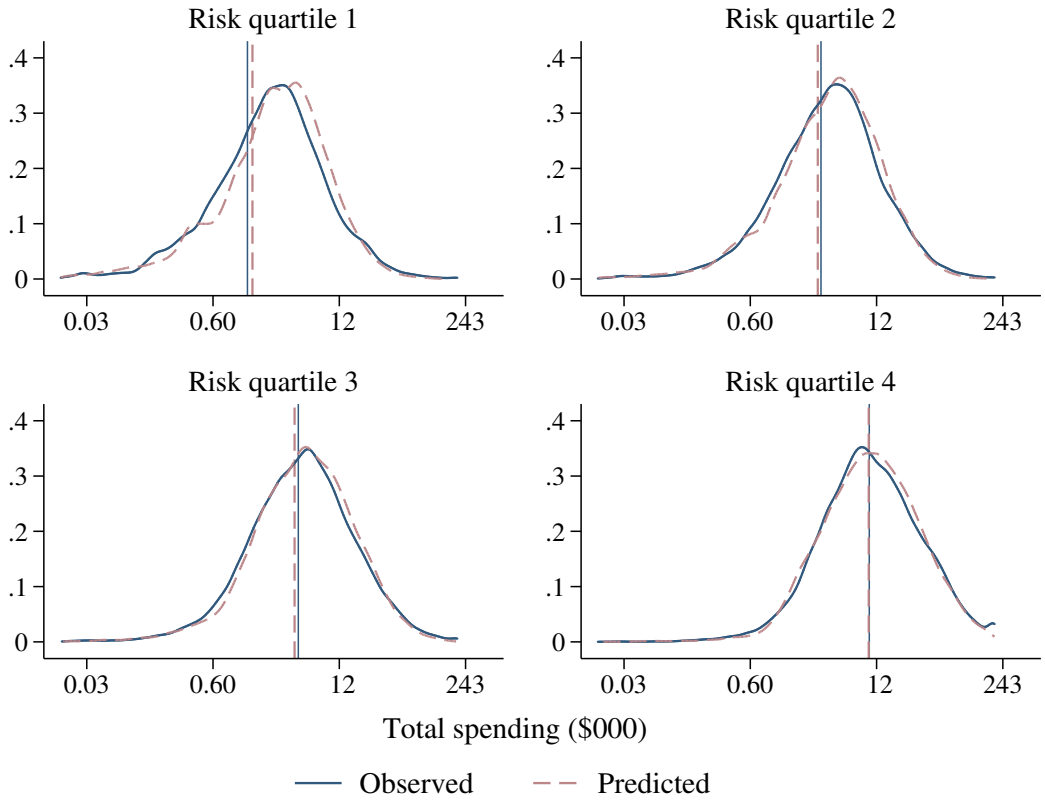
Notes: The figures shows predicted and observed market shares at the plan level. In each year, the level of observation is the household. Predicted shares are estimated using the parameters in column 3 of Table 3.4.

Figure C.5: *Model Fit: Healthcare Spending by Number of Family Members*



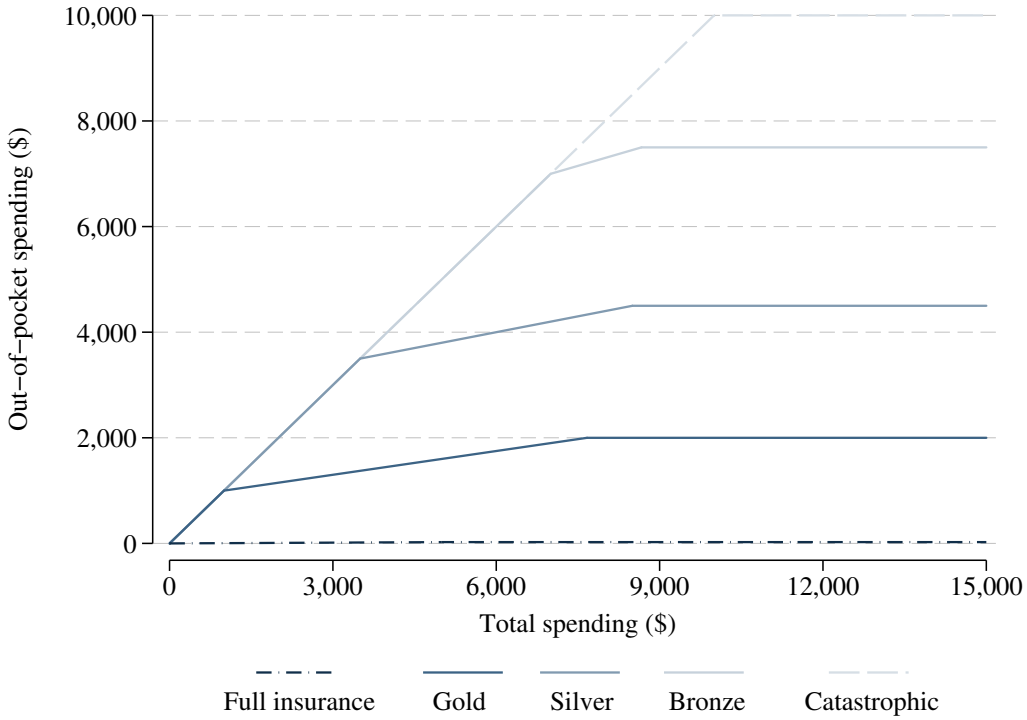
Notes: The figure shows kernel density plots of the predicted and observed distribution of total healthcare spending on a log scale, separately among households with different numbers of family members. All years are pooled together, so the observation is the household-year. The vertical lines represent the mean of the respective distribution. Predicted distributions are estimated using the parameters in column 3 of Table 3.4.

Figure C.6: *Model Fit: Healthcare Spending by Household Health Risk*



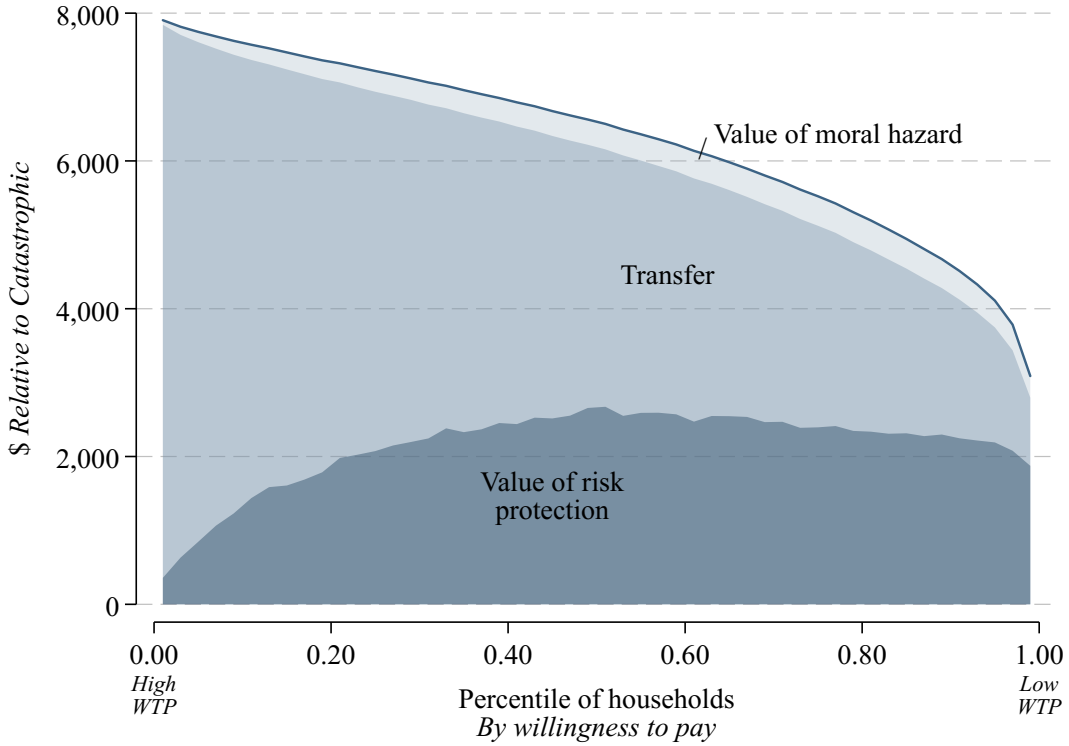
Notes: The figure shows kernel density plots of the predicted and observed distribution of total healthcare spending on a log scale, separately among households in each quartile of household health risk. Household health risk is measured as the mean risk score across individuals in the household. Quartile 4 is the sickest households. All years are pooled together, so the observation is the household-year. The vertical lines represent the mean of the respective distribution. Predicted distributions are estimated using the parameters in column 3 of Table 3.4.

Figure C.7: *Counterfactual Potential Plans: Out-of-pocket Cost Functions*



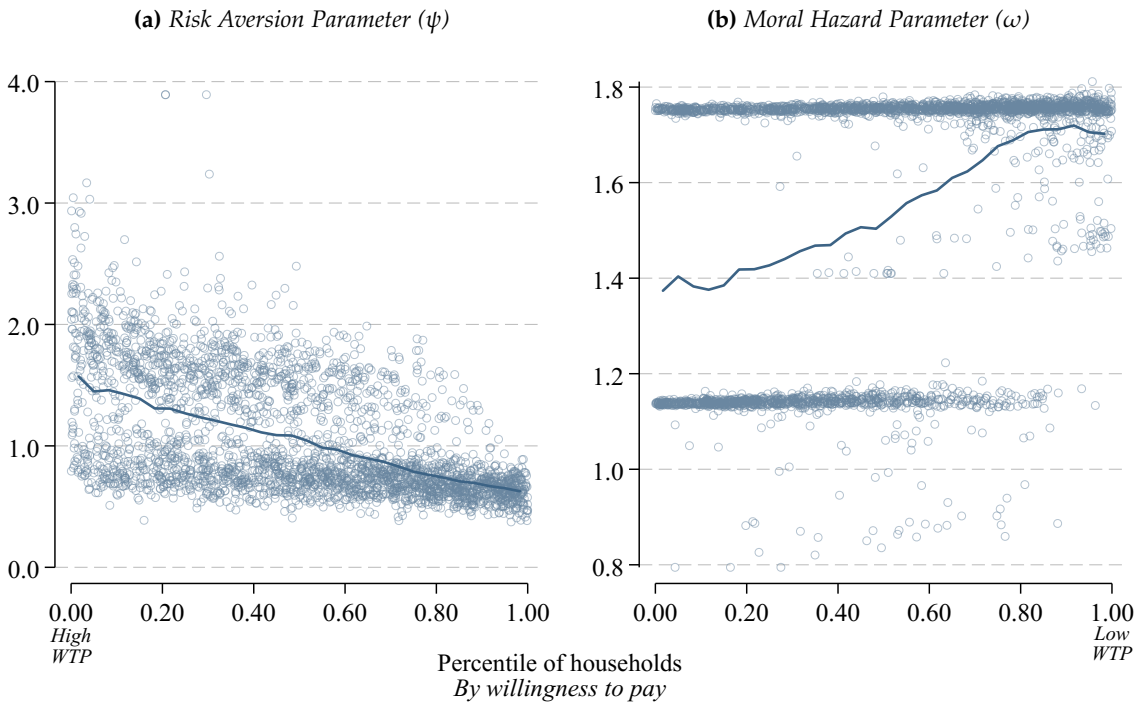
Notes: The figure shows the cost sharing schedules for the five potential plans we consider in our counterfactuals. These plans are chosen to align with the plan designs and coverage levels of typical plans on the Affordable Care Act exchanges. The exact deductible, coinsurance rate, and out-of-pocket maximum of the plans are $1,000, 15%, $2,000 for Gold; $3,500, 20%, $4,500 for Silver; $7,000, 30%, $7,500 for Bronze; and $10,000, 30%, $10,000 for Catastrophic.

Figure C.8: *Breakdown of Willingness to Pay for Gold Plan*



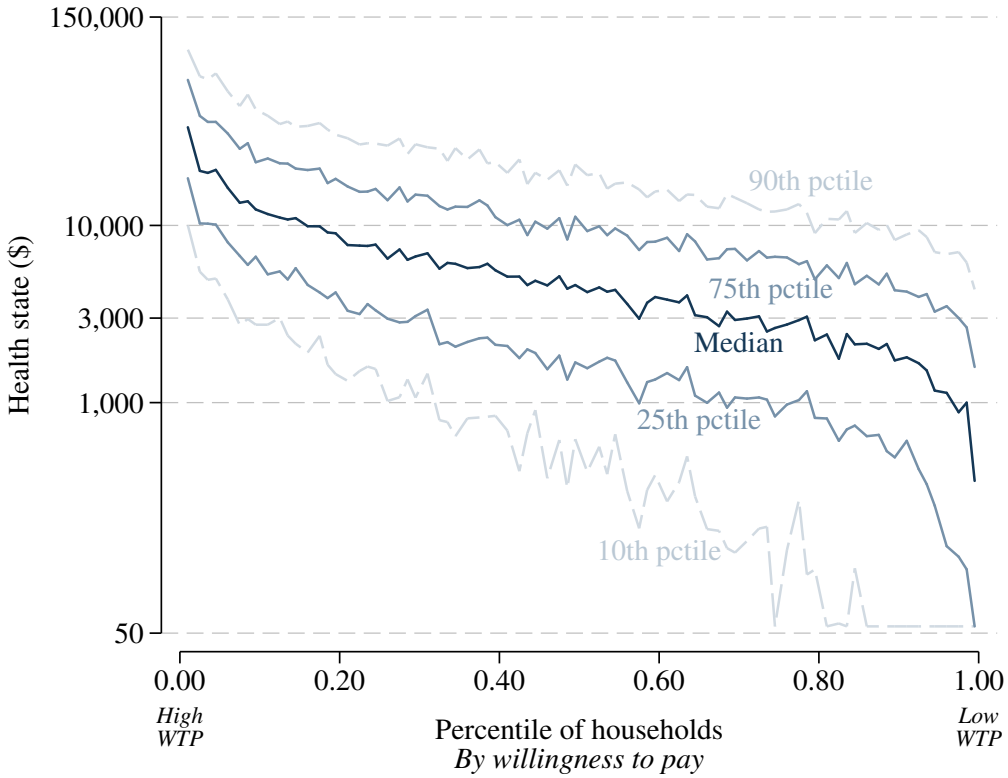
Notes: The figure shows the breakdown of willingness to pay for the Gold plan relative to the Catastrophic plan into its three component parts: mean reduced out-of-pocket costs from unavoidable medical spending, the value of risk protection, and mean benefit from moral hazard spending. Households are arranged on the horizontal axis according to their willingness to pay. The height of the shaded areas represent the average of each component of willingness to pay for households at that percentile of willingness to pay.

Figure C.9: *Risk Aversion and Moral Hazard Parameters by Willingness to Pay*



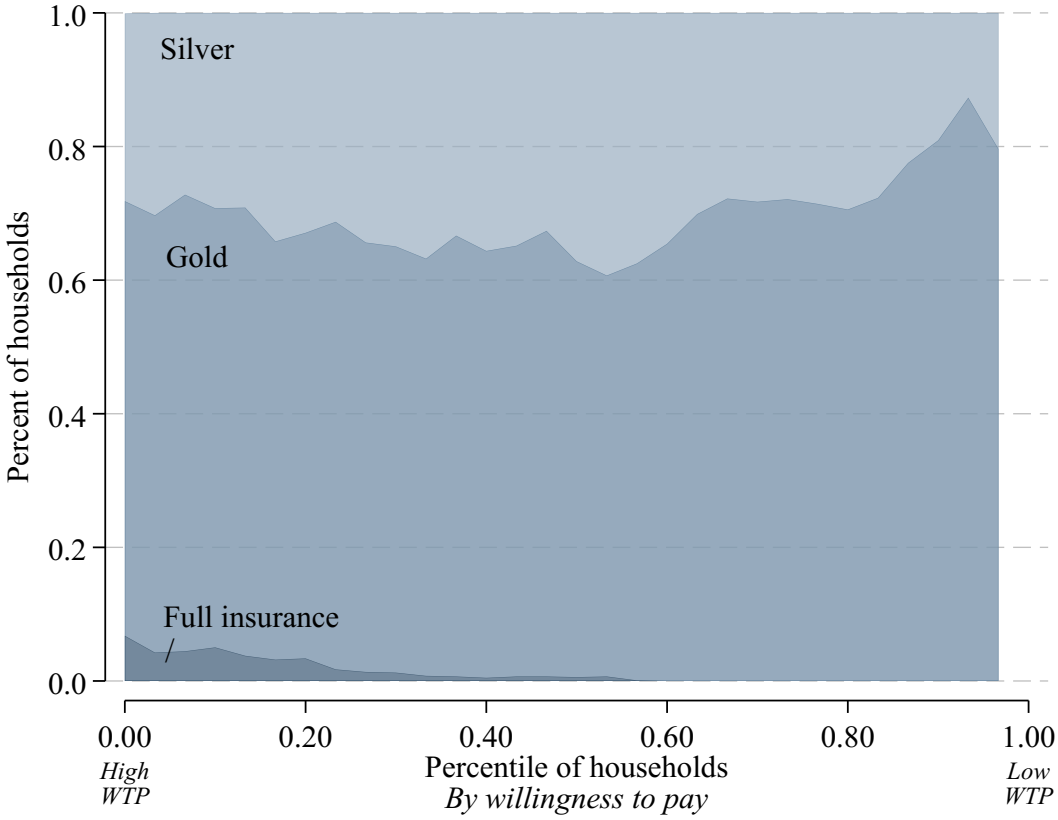
Notes: The figure shows the distribution of households' risk aversion parameter and moral hazard parameter across the distribution of willingness to pay. Each dot represents a household, for a 10 percent random sample of households. The dark line is a binscatter plot over all households, representing the mean value of the vertical axis variable at each percentile of willingness to pay. The clumping at certain parameter values is driven by the intercepts (children versus no children) coupled with the normality assumption on unobserved heterogeneity.

Figure C.10: *Household Health State Distributions by Willingness to Pay*



Notes: The figure shows the distribution of health states faced by the set of households at each percentile of willingness to pay. Health state distributions are represented by their 10th, 25th, 50th, 75th, and 90th percentiles. A health state realization is equal to unavoidable total healthcare spending. The vertical axis is on a log scale in order to show more clearly the relationship between health state distributions and relevant values of the out-of-pocket cost schedule of the plans we consider in Section 3.5.2.

Figure C.11: *Efficient Coverage Level by Willingness to Pay*



Notes: The figure shows the fraction of households at each percentile of willingness to pay for which each level of coverage is optimal. Households are ordered on the horizontal axis according to their willingness to pay. For example, among the top one percent of households by willingness to pay, full insurance is the efficient level of coverage for 6 percent of households, Gold is efficient for 66 percent of households, and Silver is efficient for the remaining 28 percent.