Title
Essays on Health and Labor Markets

Author
Knepper, Matthew Mark

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on Health and Labor Markets

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Matthew Mark Knepper

Committee in charge:

Professor Julie Berry Cullen, Chair
Professor Eli Berman
Professor Edward Castillo
Professor Jeffrey Clemens
Professor David Guss

2015

The dissertation of Matthew Mark Knepper is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2015

DEDICATION

To my family and friends for always believing in me.

EPIGRAPH

There is scarcely any passion without struggle.

—Albert Camus

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

A special thanks to my dissertation committee members: Julie Berry Cullen, Eli Berman, Edward Castillo, Jeffrey Clemens, and David Guss. Also a large thanks to my co-author, Matthew Niedzwiecki. Additionally, I would like to thank the participants at the following seminars and conferences: UCSD, UCSF/UC-Berkeley, and SAEM 2014. Finally, I would like to thank the following people for their comments on this research: Gordon Dahl, Alex Imas, and Mitch Downey.

Chapter 1, in full, is currently being prepared for submission for publication of the material. Knepper, Matt; Niedzwiecki, Matthew. The dissertation author was a co-primary investigator and co-author of this material.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Knepper, Matt. The dissertation author was the primary investigator and author of this material.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Knepper, Matt; Castillo, Edward; Chan, Theodore; Guss, David. The dissertation author was the primary investigator and author of this material.

VITA

2009	B.A. in Mathematics, Economics *with honors*, University of North Carolina, Chapel Hill
2010	M.A. in Economics, University of California, San Diego
2015	Ph.D. in Economics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Essays on Health and Labor Markets

by

Matthew Mark Knepper

Doctor of Philosophy in Economics

University of California, San Diego, 2015

Professor Julie Berry Cullen, Chair

I focus on the interaction of the legal environment with health and labor market outcomes in a variety of settings. The first chapter examines the effect of state public health insurance expansion decisions on private insurance costs by comparing how private insurance plan costs have evolved over time according to whether a state has pursued a Medicaid coverage expansion as part of the Patient Protection & Affordable Care Act. The second chapter explores how one can test for judicial ingroup biases and applies the test to workplace sex discrimination cases. The third chapter evaluates the potential efficacy of federal incentive payments designed to boost physician access to electronic health records on patient outcomes in an Emergency Department.

Chapter 1

Public Expansions and Private Spillovers: What the 2014 Medicaid Expansions Mean for Private Insurance Markets

1.1 Introduction

Unlike most products, the cost of providing health insurance depends upon the consumers who ultimately purchase it. Recent innovations in the literature provide evidence that public health insurance expansions remove high-cost individuals from the private insurance market, which reduces the average cost of providing insurance for the remaining private enrollees. These studies detect such changes in the risk pool by estimating how private coverage rates (Clemens (2014)) and premiums (Tian (2013)) evolve following public insurance expansions. However, health plans are a complex composite product. In particular, the mapping between total medical costs and the amount that insurers collect from consumers is governed exclusively by the cost-sharing features that define a plan. In the presence of higher expected costs among health plan enrollees, it follows that profit-maximizing insurers will tighten this mapping by increasing cost-sharing and reducing overall plan generosity. In this paper, we study how private insurance firms alter plan features in response to changes in private insurance risk pools induced by the Affordable Care Act (ACA) 2014 State Medicaid Expansion decisions. We offer a comprehensive view of insurer responses by considering insurers' choice of cost-sharing features, premiums, and plan network types.

Seminal theoretical work by Rothschild and Stiglitz (1976) (hereafter RS) demonstrates that when insurers have perfect information on the risk characteristics of individuals, they offer full insurance policies at actuarially fair prices. Any restrictions that limit insurers' access to such information eliminates this outcome from the realm of what is possible. In its place, RS propose a "separating equilibrium" in which low-risk consumers reveal their type by selecting plans that offer a lower quantity of insurance than is desired by high-risk consumers. Prior to the ACA, much of the individual market for health insurance was experience-rated, operating in the unregulated benchmark RS environment. Beginning in January of 2014, however, the ACA acutely limited insurers' ability to obtain information on individual risk characteristics, let alone use it to make insurance policy pricing decisions. Generalizing from RS, insurers could respond to these regulations by altering any plan characteristic that is valued relatively more by high-risk consumers to produce a separating equilibrium isomorphic to the one described above. In particular, we propose that insurers increase cost-sharing requirements to sort for low-risk consumers.[1] Healthier individuals should expect to utilize a low amount of care and should accordingly be hurt less by higher usage fees than will sick individuals. Because the 2014 Medicaid expansions remove high-cost individuals from the private market risk pool, the average plan offering should have relatively lower cost-sharing requirements in states that expand Medicaid.

Certainly cost-sharing requirements are not the only plan features that can serve as an effective self-selection mechanism in community-rated markets. Others have found that HMO plans, which offer a lower 'quantity' of insurance by limiting specialist referrals and requiring that enrollees use a primary care gatekeeper, have historically been used in this capacity following the enactment of regulatory restrictions (Feldman and Dowd (2000); Buchmueller and Dinardo (2002); Buchmueller and Liu (2005); Lo Sasso and Lurie (2009)). Additionally, Vaithianathan (2004) found that, when prevented from pricing based on age, insurers offered both expensive plans with comprehensive coverage and cheaper plans that failed to cover services typically utilized by the elderly (such as treatment for cataracts) in order to circumvent the age discrimination regulations. However, no one has yet considered how insurers might

[1] More recent models have extended the RS framework to allow for enrollee heterogeneity along two dimensions – (1) expected cost and (2) risk preferences (Veiga and Weyl (2014); Cohen and Einav (2007)). Related empirical work has shown that in many insurance markets, risk aversion and expected costs are negatively correlated (Finkelstein and McGarry (2006); Fang et al. (2008)). To the extent that low-risk individuals are also more risk averse, insurers should tend to adjust deductibles and cost-sharing rates rather than out-of-pocket maximums.

respond to changes in adverse selection pressures by way of the cost-sharing channel. Perhaps this is due to the relative difficulty of interpreting how a change in plan features maps to consumer costs. More likely, a historical lack of detailed information on private plan features is the culprit.

We find evidence that opting out of the 2014 Medicaid expansion increases private plan deductibles by \$296, or 10.5% on average. These deductible increases are accompanied by an overall decrease in plan generosity of 0.8%. The decrease in plan generosity suggests that insurers are not making compensatory changes to other features of the cost-sharing structure (i.e., annual out-of-pocket limits or coinsurance rates) that would otherwise counterbalance the decrease in overall plan generosity owing to higher deductibles. In a simulation exercise using MEPS data, we produce a lower-bound estimate of the private out-of-pocket savings resulting from a Medicaid expansion of \$232 per person annually (6.2% of medical spending) for those in the intermediate spending range and \$61 per person annually (3.6% of out-of-pocket spending) overall. Medicaid expansions had no effect on HMO plan offerings, though overall HMO plan offerings increased from 5 to 30 percent following the ACA private market regulatory reforms. Lastly, we check to see that insurers do not reduce premiums on plans offered in non-expansion states, which could be done to compensate for less favorable cost-sharing provisions. We use a methodology similar to Dube et al. (2010), restricting our sample to rating areas which lie directly across from one another along state borders, and find null effects on premiums for the most highly subsidized plans, which allow us to rule out large (negative and positive) premium changes.[2]

[2] Currently, we are missing valid premium estimates for pre-ACA health insurance markets as they were excluded from the archived healthcare.gov databases from which we obtained all other plan features data. As such, we rely on a local identification approach with cross-sectional 2014 exchange data. When 2015 exchange data become available, we hope to implement panel methods to identify the effect of Medicaid expansions on premiums in the same way that we have for other plan characteristics.

Our paper is the first, to our knowledge, to examine how insurers respond to changes in risk pool composition along the cost-sharing design margin. Given our results, we conclude that a failure to take into account changes in private plan generosity significantly understates the private savings resulting from public health insurance expansions.

The rest of the paper proceeds as follows: section 2 provides institutional background, section 3 presents the model, section 4 describes the data, section 5 describes the identification, section 6 reports the results, and section 7 concludes.

1.2 Institutional Background

The ACA dramatically altered the private non-group health insurance market, which will be the focus of this paper.[3] In particular, the law established insurer regulations to prevent discrimination against unhealthy individuals, imposed an individual mandate penalty for foregoing insurance, and reduced insurance costs for low-income individuals through the use of subsidies. Lastly, it appropriated federal funds to support the expansion of Medicaid programs to cover all individuals earning no more than 138% of the Federal Poverty Level (FPL) in all states electing to do so.[4] The first three of these policies interacted with state Medicaid expansion decisions in a way that has different implications for the resulting risk pools. In the remainder of this section, we expand upon these details in order to motivate our predictions for how private insurers will differentially respond to state Medicaid expansion decisions.

[3] Similar reforms impacted the small-group markets but many of the reforms have not yet been enacted and premium data are largely unavailable for these markets.

[4] States receive 100% federal financing to cover the newly eligible for the first three years. Thereafter, the federal finiancing rate is gradually reduced to 90% by 2020.

1.2.1 The New Health Insurance Landscape: Insurer and Individual Regulations

The ACA established new minimum coverage requirements for health plans that vastly improved plan generosity; only 43% of all pre-ACA plans in our data met such requirements. These new standards include limitations on individual out-of-pocket maximums to $6,350, free preventive care, the abolition of lifetime benefit caps, and coverage for ten essential health services, such as mental health treatment and prescription drug coverage.[5] All plans are further classified by metal tiers–bronze, silver, gold, and platinum–that correspond to the actuarial value (AV) of the plan. The actuarial value is the percentage of total medical expenses, for a fixed population, that would be covered by the insurer on average. Bronze plans have an AV of 60%; silver plans have an AV of 70%; gold plans have an AV of 80%; and platinum plans have an AV of 90%. The AV of a plan is also allowed to fluctuate within 2 percentage points of the target AV associated with its metal tier, e.g. silver plans can have an AV of 68-72%.

The ACA also required states to create exchanges, or online marketplaces, which allowed consumers to browse these plans on the basis of price, provider networks, and cost-sharing/benefits structures before making a purchase decision.[6] The insurers offering these plans were subjected to a high degree of regulatory scrutiny through modified community rating and guaranteed issue rules. Plan offerings can

[5]Full list of essential health services include: emergency services; hospitalization; maternity and newborn care; mental health and substance use disorder services, including behavioral health treatment; prescription drugs; rehabilitative and habilitative services and devices; laboratory services; preventive and wellness services and chronic disease management; and pediatric services, including oral and vision care. source: https://www.healthcare.gov/glossary/essential-health-benefits/

[6]States had the additional option of setting up a joint state-federal partnership exchange or defaulting to allow the federal government to run and operate the exchange. Sixteen states and DC elected to operate their own exchanges, twenty-seven states are part of the federal exchange, while the remaining seven states have chosen a joint state-federal partnership exchange. Figure 1.13 shows a map that displays the type of exchange run in each U.S. state.

only vary at the rating area level, which typically consists of a group of two or more contiguous counties.[7] These new regulations marked a significant departure from the previous highly unregulated individual market, which, in all but seven states,[8] allowed insurers to charge different prices and/or deny coverage to individuals according to health status. Modified community rating prohibits insurers from pricing policies based on any enrollee characteristics, including pre-existing medical conditions, except for age, which is restricted to a rigid three-to-one rating band (i.e., the oldest enrollees cannot be charged more than three times the premium of the youngest enrollees).[9] Guaranteed issue ensures that no one can be denied coverage outright. The combination of these two regulations leads to adverse selection in health insurance markets since the benefit–but not the price–of enrollment increases with expected medical costs. The abolition of medical underwriting practices in favor of community-rated insurance also implies that the profitability of newly offered plans is inextricably linked to the composition of the entire enrollee risk pool.

The ACA also coupled these community rating and guaranteed issue regulations with other measures to increase rates of insurance coverage. One such measure is the individual mandate, or the requirement that all individuals obtain health insurance coverage. The mandate is enforced by income penalties levied against those electing not to purchase insurance. It unambiguously improves the risk pool by encouraging exchange participation among healthy individuals with a low value of insurance who might otherwise opt out. The penalties are the greater of 1% of income and $95 per person in 2014.[10] Importantly, individuals earning less than 138% FPL without

[7]The rating areas of Idaho, Nebraska, Massachusetts, and the county of Los Angeles, California are determined by collection of zip9 codes.

[8]These states include Maine, Massachusetts, New York, New Jersey, Oregon, Vermont, and Washington.

[9]Smokers may also be charged up to 50% more for any given policy.

[10]These penalties rise to the greater of 2% of income and $325 per person in 2015, and 2.5% of income and $695 in 2016. In subsequent years, these penalties will increase annually to keep pace with cost-of-living adjustments.

access to affordable Employer Sponsored Insurance (ESI)[11] or Medicaid are exempt from the mandate penalty.[12]

Individuals earning 138-400% FPL in states that expand Medicaid and individuals earning 100-400% FPL in states that do not expand Medicaid are also eligible for exchange subsidies that reduce the premium price paid. Thus, exchange participation should begin at 138% FPL in states that expand Medicaid and at 100% FPL in states electing not to expand Medicaid. This feature lies at the heart of our identification strategy, which exploits differences in exchange participation incentives for individuals earning between 100% and 138% FPL across states and rating areas on the basis of whether the individual resides in an area that expanded Medicaid coverage.

These premium subsidies are tied to the second-lowest price silver plan in a given rating area. The subsidy limits what individuals pay in premiums from 2% of income for individuals earning between 100% and 133% of the FPL to 9.5% of income for individuals earning between 300% and 400% of the FPL on a sliding scale. For example, a 27 year-old non-smoker earning between 100% and 138% of the FPL should expect to pay between $19 and $43 per month for a silver plan valued at $211 per month,[13] which amounts to a price reduction of 80-91%. In addition to premium subsidies, cost-sharing credits reduce cost-sharing amounts and annual out of pocket limits for individuals earning between 100% and 250% of the FPL who purchase silver plans.[14] Thus, we expect that silver plans will induce the most severe adverse selection

[11] Affordable ESI defined as an insurance plan that has an actuarial value of at least 60% and does not exceed 9.5% of the individual's income

[12] Those earning less than the federal income tax filing limit ($10,000 annually for individuals and $20,000 annually for couples) are also exempt from the mandate.

[13] This is the national average monthly premium for a silver plan.

[14] These credits reduce cost-sharing such that The average actuarial value (AV) of a silver plan is 70%. These cost-sharing credits increase the AV on a sliding scale according to income so that individuals earning 100-150% have their plan AV increased to 94%, those earning 150-200% have their plan AV increased to 87%, and individuals earning between 200% and 250% receive a plan AV of 73%. In other words, these groups of individuals can expect to pay 6%, 13%, and 27% of total incurred health costs on average, respectively.

pressures among individuals earning less than 138% FPL who are without access to Medicaid coverage.

1.2.2 State Medicaid Expansion Decisions

When the ACA was passed in March 2010, all states were expected to expand Medicaid income eligibility limits from pre-reform levels up to 138% FPL since the federal government retained the right to withhold all matching funds to states that did not comply.[15][16] In the June 2012 *National Federation of Independent Business (NFIB) v. Sebelius* decision, the U.S. Supreme Court upheld the individual mandate but struck down the provision that allowed the federal government to withhold state Medicaid funds on the grounds that it was "unconstitutionally coercive."[17] This decision effectively left the Medicaid expansion decisions to the states; 24 of them opted out. Figure 1.5 shows a map of U.S. states classified by state Medicaid expansion decisions. Shor (2013) shows that these decisions were largely based on legislator ideology and not any kind of state-specific insurance market conditions, though we are careful to control for any such differences in our estimation framework. Figure 1.6 shows the change in Medicaid income eligibility limits across all states from 2013 to 2014 by expansion status. It is worth noting that childless adults will be the most highly impacted by Medicaid expansions; eligibility limits increased more than five-fold from 29% FPL to 147% FPL in expansion states whereas all childless adults in non-expansion states are ineligible for Medicaid coverage across both periods. Overall, Medicaid eligibility among those earning < 138% FPL increased from 43% to 100% in expansion states but remained at 18% in non-expansion states across the reform pe-

[15] The federal government spent nearly $265 billion on state Medicaid programs in 2013, or $5.3 billion per state on average.

[16] See Congressional Budget Office, The Budget and Economic Outlook. http://www.cbo.gov/publication/45229

[17] source: http://kaiserfamilyfoundation.files.wordpress.com/2013/01/8347.pdf

riod.[18] Due in no small part to these expansions, Medicaid is now the fastest-growing entitlement program with spending projected to double by 2020.[19] Hence, the unexpected opt-out provision creates a large degree of quasi-random variation that we exploit for the purposes of identifying the effect of Medicaid expansions on private plan characteristics.

1.2.3 Exchange Rules and their Influence on Private Market Risk Pools

Table 1.1 displays subsidy availability and mandate penalty exemption status partitioned by income categories across Medicaid expansion and non-expansion states. The expected difference in private risk pools across expansion and non-expansion states is governed by the exchange participation incentives of the group earning between 100% and 138% FPL, which are highlighted in green. In Medicaid expansion states, these individuals are not eligible for exchange subsidies since they have access to a public insurance option. Given the relatively high price of unsubsidized private insurance (an average bronze plan costs $2101 annually, for a 27 year old non-smoker, which accounts for approximately 20% of income for a single individual living at the poverty line), these individuals will likely either opt for Medicaid or remain uninsured. Conversely, individuals earning between 100% and 138% of the FPL in states that do not expand Medicaid face strong incentives to enter the private health insurance exchanges since they qualify for heavily subsidized private insurance. This group is also exempt from the mandate penalty, which allows healthy individuals to costlessly opt out, further exacerbating adverse selection pressures. Since income verification standards are lax and incomes for the group earning less than the FPL are volatile, we also expect that unhealthy individuals earning less than 100% may overstate their

[18] These estimates are derived using the American Community Survey (ACS).
[19] source: http://www.healthreformgps.org/wp-content/uploads/CBOMeansFeb2013.pdf

incomes in order to take advantage of the generous exchange subsidies in states that do not expand Medicaid (Eibner (2014)). These features combined predict that the most adversely selected individuals earning less than 138% of the FPL will enter the private market risk pool in non-expansion states whereas individuals in this same income group that reside in Medicaid expansion states will not participate in the exchanges.

How these differences in the income distributions of market participants in the expansion and non-expansion states affect the cost of providing insurance depends upon the relationship between health care spending and income. There is a wide body of evidence in the epidemiology literature (Marmot (2002); Braveman et al. (2011); Elo (2009)) and economics literature (Clemens (2014); Cutler et al. (2008)) which suggests that poorer individuals are more likely to be of worse health status and more costly to insure on average. When we account for the fact that even among the group earning less than 138% FPL in non-expansion states, the particularly healthy individuals can opt out of insurance, we predict that the non-expansion state risk pools will be relatively more costly to insure.

We compared expected health spending from the 2007-2009 Medical Expenditure Panel Survey (MEPS) for uninsured adults, aged 18-64 for two groups: (1) those earning less than 138% FPL and (2) those earning more than 138% FPL. Since the group earning less than 138% FPL is uninsured, we expect that it includes some poor, healthy individuals who were eligible for public insurance but neglected to take it up. Nonetheless, the expected spending for uninsured adults earning less than 138% FPL exceeds that of those earning more than 138% FPL by nearly $280, or 20 percentage points, annually ($1,645.67 vs. $1,366.64). Figure 1.7 shows the distribution of expected healthcare spending for the two income groups. The spending distribution of the poorer group has less mass at the low spending portion of the graph, more mass in the middle parts of the distribution around $\$4,000 - \$5,000$, and has a much longer, thicker right tail than the higher income group. It is this group of particularly costly individuals in the right tail of the spending distribution that could plausibly push up premium prices and deductibles on insurance exchanges in states for which these individuals lack access to Medicaid coverage. Using data from the Consumer Population Survey (CPS), we also confirm that the low-income group has a self-reported health status that is 21% worse than that of the full population (2.69 vs. 2.22, on a scale of 1 to 5 where 1=Excellent and 5=Poor).

1.2.4 Prior work on Public Insurance Expansions and Private Insurance Markets

The literature on the interaction between public insurance coverage expansions and private health plan costs is sparse, which is unsurprising given the historical lack of transparency of private plan characteristics in the individual market. One of the goals of the ACA was to make healthcare costs and pricing more transparent, which is a feature that we take advantage of in this paper.

Clemens (2014) studies the effect of public insurance expansions on private coverage rates in community rated markets. Consistent with adverse selection pressures, he finds that private coverage rates decline when community rating rules are enacted since costly individuals enter the private market when insurers are prohibited from charging rates based on individual health measures. The analogous prediction in our setting is that private insurance costs increase in non-expansion states after the reforms occur, i.e. community rating and guaranteed issue rules, are enforced in private insurance markets. Moreover, Clemens finds that private insurance coverage rates largely recover following subsequent public insurance expansions targeted at the medically needy. In our setting, this policy environment is similar to that of the expansion states, where community rating rules and public insurance expansions are enacted simultaneously. The prediction in our paper is that private health plan costs will fall relative to non-expansion states. This finding would be consistent with Clemens' conclusion that private coverage rates recovered following Medicaid expansions owing to a reduction of adverse selection pressures on the private markets.

Tian (2013) also studies the relationship between public insurance expansions and private insurance markets by estimating how the coverage expansions of the State Children's Health Insurance Program (SCHIP) in the late 1990's affected the premiums of ESI plans. She finds that increasing the SCHIP threshold by 100 percent of the FPL decreases ESI premiums by 2-3%. The proposed mechanism by which these premium prices are affected is through adults that drop ESI when their children gain coverage for SCHIP, which is a more indirect channel than is the case in our study setting. Thus, we should expect to find a larger impact.

Eibner (2014) uses data from the 2008 Survey of Income and Program Participation (SIPP), the MEPS, and the Kaiser/HRET Survey of Employer Benefits to simulate the effect of Medicaid expansion decisions on private health premiums. Her simulations predict an increase in exchange premiums for states that do not expand Medicaid due to the participation of and adverse selection among individuals earning less than 138% FPL. We investigate a larger number of insurer responses, including changes to plan networks and cost-sharing design.

Evidence on the new 2014 health insurance exchanges is limited given that data have only recently become available. Pauly et al. (2014) provide the first comprehensive study of the changes in premiums and expected out of pocket payments due to the private market reforms of the ACA. Using survey data from the CPS on individuals with private insurance from 2010-2012 in addition to plan data from states on the Federally-facilitated marketplace (FFM) and California, they construct a state-level panel that they then use to calculate the expected change in total cost of plans purchased on the individual market. They find that the total average expected costs of individual health insurance increased by between 14 and 28 percent due to the ACA reforms.

Kowalski (2014) offers a full state-by-state analysis of the change in total welfare attributable to the establishment of the new health insurance exchanges. She uses quarterly pre-ACA and post-ACA data on coverage, average costs, and premiums collected by the National Association of Insurance Commisioners (NAIC) in order to estimate the change in welfare resulting from the ACA's impact on average enrollee cost and markups in the individual health insurance market. In particular, she finds that Medicaid expansion has no statistically significant effect on total welfare. However, one core assumption underlying her analysis is that plan generosity remains unchanged after the establishment of the new health insurance exchanges. On the contrary, we argue that average plan generosity should have increased mechanically due to the imposition of minimum essential benefit and actuarial value requirements, which implies that she understates the change in welfare associated with Medicaid expansions. Moreover, we predict and test for a heterogenous behavioral response to anticipated changes in local and state risk pools along the cost-sharing design margin.

Lastly, Gruber et al. (2014) use cross-sectional data from the 34 states participating in the FFM of 2013-2014 to study the effect of competition on pricing. They use United Healthcare's decision not to participate in the exchanges as an instrument for measuring how expected changes to the Herfindahl-Hirschman Index (HHI) affect premiums. They find that United's nonparticipation decision decreased private market premiums on the individual market by 5.4% but do not find any association between premium prices and Medicaid expansion status. We hope to complement such evidence by using an alternate approach that allows us to measure changes in health plan costs before and after the ACA reforms.

In our study, we focus on estimating private cost spillovers (cost-sharing, premiums, and plan network choices) induced by the participation of poor individuals who are ineligible for Medicaid coverage due to the fact that their state did not expand public eligibility for the program. We are the first to use sub-state level variation in the size of the Medicaid ineligible population to answer this question. Our study also adds to the literature on the interaction of private and public health insurance markets and represents a methodological improvement over prior studies that have not taken into account changes in plan generosity engendered by the ACA.

1.3 Model

In this section, we begin by defining the relationship between medical spending, plan features, consumer costs, and insurer costs. We then add a contingent of high-cost individuals to the private insurance market, which simulates the effect of a Medicaid expansion opt-out on the risk pool. Finally, we analyze the set of behavioral responses available to insurers.

Figure 1.1 shows the relationship between total medical spending and consumer out-of-pocket spending for a stylized plan with a deductible equal to \$2,000, a cost-sharing rate of 40%, and an annual out-of-pocket limit of \$3,000. Because deductibles specify the amount that a consumer must cover in full before an insurer begins to pay out the first dollar of benefits, the slope equals 1 until the deductible is met. The next segment of the graph has a slope equal to that of the coinsurance rate, which corresponds to how much the consumer must cover for each additional dollar of medical spending. Finally, the slope flattens out to 0 once the consumer reaches his or her out-of-pocket limit, since insurers must cover 100% of any remaining medical bills.

In Figure 1.2, we overlay the insurer cost schedule with the consumer cost schedule. These functions are complementary in the sense that their sum should be a 45 degree line through the origin since all medical spending must be covered by either the consumer or the insurer.[20] We focus on this latter function for the purposes of characterizing insurer behavior.

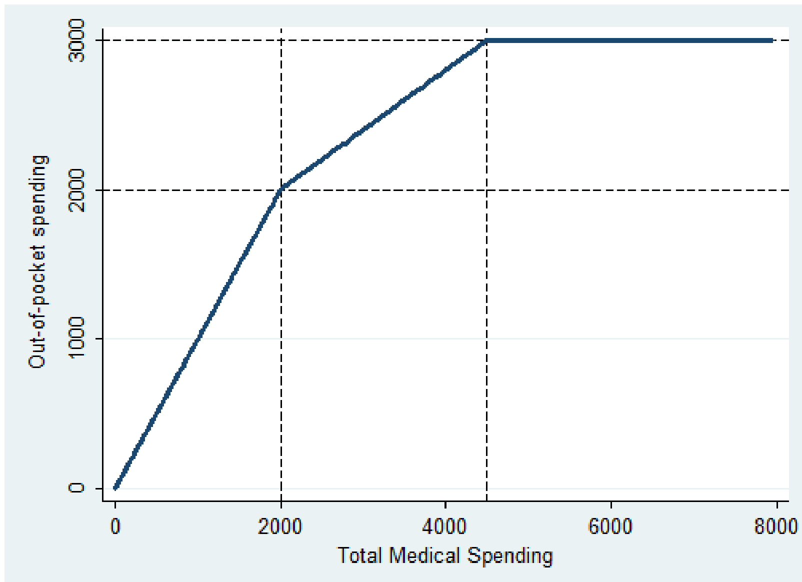


Figure 1.1: Stylized Plan with Mapping to Out-of-pocket costs

We now focus our attention on the cost-sharing design of silver plans, as they have attracted the largest enrollee pool by a wide margin. While insurers have some flexibility in setting deductibles, coinsurance, and out-of-pocket limits, ACA regulations require that a plan's overall actuarial value (AV) must not differ from the target AV of 70% by more than 2%. Importantly, the Center for Medicare and Medicaid Services (CMS) calculates AV based on a fixed population.

We start under the condition that all states expand Medicaid income eligibility limits to 138% FPL, as was the initial expectation prior to the *National Federation of Independent Business (NFIB) v. Sebelius* Supreme Court decision. Let s = total

[20]However, the sum of the slopes of these lines may be less than 1 if the government provides cost-sharing subsidies to certain individuals.

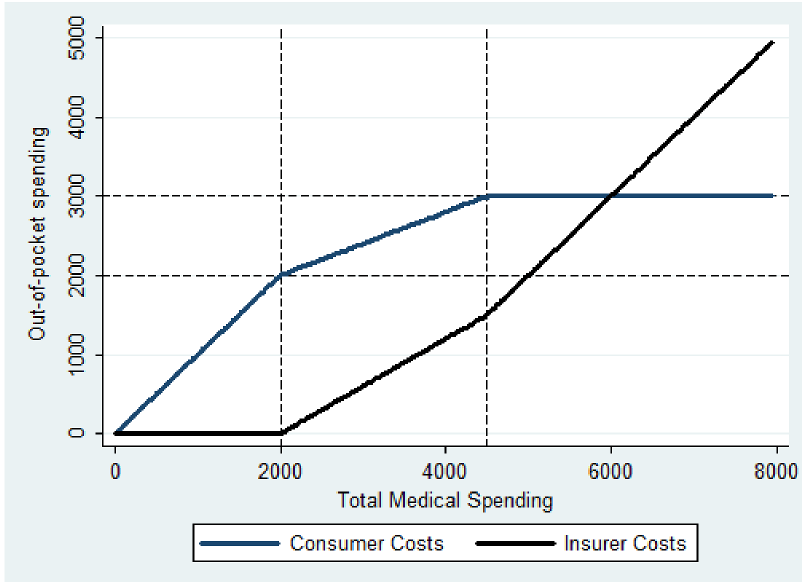


Figure 1.2: Consumer Costs and Insurer Costs by Total Medical Spending

medical spending, d = deductible, r = coinsurance rate, and m = out-of-pocket maximum. Fix the plan features offered.

Then we define the insurer cost function, $f(s) =$

$$\begin{cases} 0 & \text{if } s < d \\ (1-r) \times (s-d) & \text{if } d < s < s(m) \\ s - m & \text{if } s \geq s(m) \end{cases}$$

For the sake of simplicity, we assume no moral hazard.[21] Further suppose that $g(s)$ is a continuous function that characterizes the medical spending distribution across the risk pool, where $s \in [a, b]$.

Then let $w(s) \equiv f(s)g(s)$, which is the insurer cost function weighted by the medical spending distribution.

Claim: There exists some $c \in [a, b]$ where $w(c) =$ the average cost of insuring

[21]Though moral hazard will tend to exacerbate the extent to which raising cost-sharing requirements are effective in controlling costs.

all individuals in the risk pool.[22]

Community rating rules lead to an outcome where insurers set the premium equal to $w(c)$, which implicitly defines the relationship between the premium price and cost-sharing features. We assume that all individuals with $s_i < c$ participate due to sufficiently high risk-aversion or because the mandate penalty induces participation. Then the insurer cost schedule, or mapping from total medical spending to the insurer's portion of the bill, is represented in Figure 1.3a.

Now suppose that a hypothetical expansion state opts out of the Medicaid expansion. Then the group of high-risk individuals that would have been covered by public insurance enters the private risk pool. This shifts the medical spending distribution to the right as in Figure 1.3b. If insurers do not alter cost-sharing requirements, premiums will rise to $w(c')$ (where $c' > c$), which is the weighted average cost of insuring the new risk pool under the original insurer cost-sharing function, $f(s)$. However, raising the premium will worsen the risk pool as some individuals with $s_i < c'$ will exit provided that they were marginal beforehand due to risk aversion or the mandate penalty.

Insurers could alternatively respond by raising the plan deductible, which should not affect the consumer costs of the low-risk individuals whose expected medical spending is less than that of the plan deductible. Increasing the deductible thus raises consumer costs for those individuals initially in the coinsurance range, which shifts the insurer cost function down over the corresponding domain. This is represented in Figure 1.4 by a rightward extension of the flat portion of the insurer cost function. Intuitively, this is equivalent to reducing the plan membership fee in favor

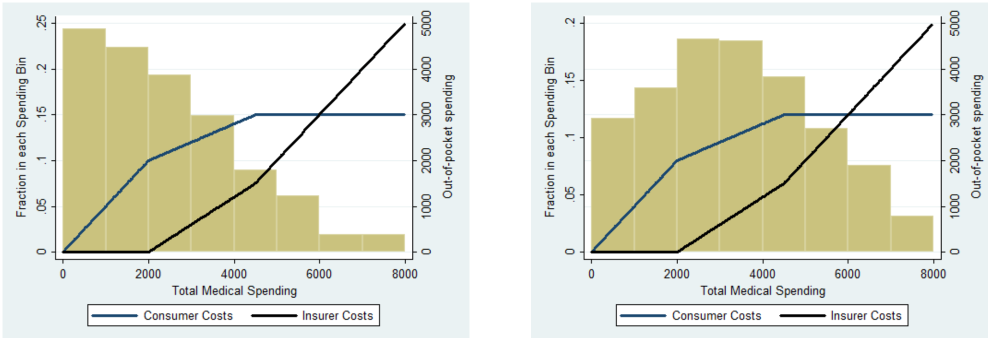
[22]Proof: Since f and g are continuous functions and w is the product of two continuous functions, $w(s)$ is also continuous on the interval [a,b]. Set $W(s) = \int_a^s w(t)\mathrm{d}t$. By the Mean Value Theorem for Integrals, $\exists\, c \in (a, b)$; $\frac{1}{b-a}\int_a^b w(t)\mathrm{d}t = w(c)$.

of raising the usage fee, which shifts the burden of covering the new entrants from all individuals to those in the intermediate range of the medical spending distribution. Insurers could instead achieve a similar result by raising coinsurance or the out-of-pocket limits. However, to the extent that low-risk individuals are more risk-averse, they will find increases in out-of-pocket maximums relatively unattractive.

Because AV is calculated using a fixed base population, increasing cost-sharing requirements is more efficient in generating dollar-for-dollar decreases in insurer costs the more adversely selected is the risk pool. Thus, the returns to raising cost-sharing requirements are necessarily greater in non-expansion states than in states that expand Medicaid.

Another margin on which insurers may respond to an increase in adverse selection is through plan network choice. The use of HMO plans–which limit referrals to specialists and require that individuals pay the full costs of out-of-network coverage–can shift the entire medical distribution to the left by selecting for good risks. Recent empirical work by Gruber and McKnight (2014) shows that narrow network plans can reduce spending on the intensive margin as well by encouraging primary care utilization in place of specialist or emergency care.[23] Thus, we predict that insurers with more restrictive plan networks (i.e., HMOs) will see greater expansion, as proxied for by a larger number of plans offered, in Medicaid opt-out states.

[23] Although not all HMO plans use narrow networks, they are commonly associated.



(a) Expansion States

(b) Non-Expansion States

Figure 1.3: Medical Spending Distributions and Mapping to Costs

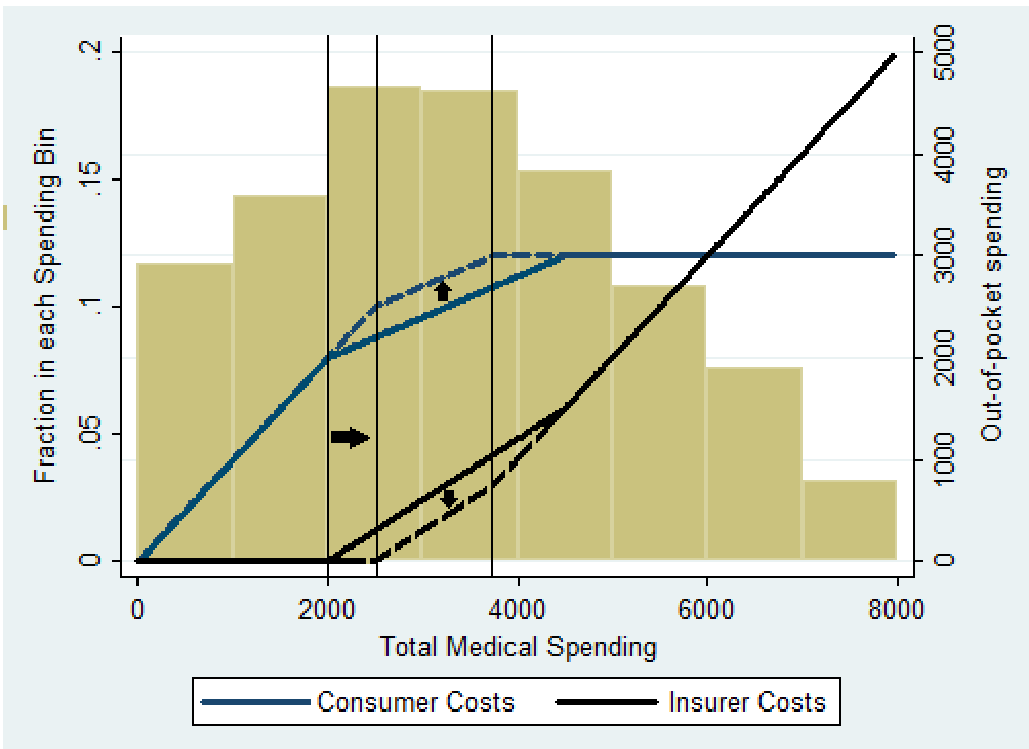


Figure 1.4: Insurer Response to an Increase in Adverse Selection
This illustrates the effect of increasing the deductible on both the consumer and insurer cost schedules. The flat region of the insurer cost schedule is extended while consumer costs increase for all individuals initially in the plan coinsurance region.

1.4 Data

To carry out our analysis, we construct a novel dataset that combines cross-sectional data scraped from federal and state exchange websites in 2014 with archived 2011 and 2013 databases of private health plans listed on Healthcare.gov. The data is at the health insurance plan level. In order to make plans comparable over time, we first drop all pre-ACA plans that would have failed to meet the minimum "essential health benefits" set forth by the law in 2014.[24] We find that just 43% of all pre-ACA plans offered on healthcare.gov would have met these new standards. This highlights the fact that a failure to account for changes in plan generosity invalidates any pre/post plan comparison; new plans are required to offer more comprehensive coverage while limiting right-tail spending for consumers.

Next, we calculate the Actuarial Value (AV) of all surviving pre-ACA plans using the methodology developed by the Center for Medicare and Medicaid Services (CMS) to assign 2014 exchange plans to metal tiers.[25] The three main inputs to the AV calculation are the general coinsurance rate, the deductible, and the annual out-of-pocket spending limit. We use these variables in combination with continuance tables of expected spending, which are based on a standard commercial database of individuals enrolled in plans during the 2010 calendar year, to estimate the expected fraction of overall health spending that will be covered by the insurer for all plans in our database. We then assign plans with an AV of 60% to the bronze tier, those with an AV of 70% to the silver tier, those with an AV of 80% to the gold tier, and those with an AV of 90% to the platinum tier, while allowing for a 2% fluctuation in either direction. These steps allow us to directly compare plans with similar benefit structures before and after the reforms, which represents a critical methodological

[24] These standards are detailed in Section 2.1.

[25] The methodology used is decribed here: http://www.cms.gov/CCIIO/Resources/Files /Downloads/av-calculator-methodology.pdf

improvement over prior studies. We focus exclusively on the 35,258 silver plans in the sample since these are the plans that low-income folks will find most attractive in the post-reform era. Also, insurers must offer at least one silver plan in any rating area in which they choose to participate. Preliminary enrollment reports indicate that silver plans accounted for 76% of plans purchased by individuals receiving financial assistance and 65% overall.[26]

To isolate the effect of Medicaid expansions on private markets independently from that of other market regulations imposed by the ACA, we drop all states that had pre-ACA community rating or guaranteed issue rules. These states include Maine, Massachusetts, New Jersey, New York, Oregon, Vermont, and Washington – all states that expanded Medicaid. States that had already enacted community rating regulations prior to the reforms would expect a much smaller level increase in expected enrollee costs. Thus, including them in our analysis would bias our results towards finding that the Medicaid expansion decreases plan generosity. We also exclude “active purchaser” states. Active purchaser states – California, Rhode Island, Massachusetts, New York, Oregon, and Vermont – had the power to reject health insurance plans for not offering competitive rates, providing wide enough provider networks, or a host of other reasons, even if the plans otherwise met all exchange requirements. Since these six states all expanded Medicaid, including them in our sample would similarly bias our results towards finding a spuriously larger negative effect of Medicaid expansions on private costs. We also excluded Wisconsin because their Medicaid program, Badgercare, had already covered childless adults earning up to 100% FPL prior to the ACA reforms but the state did not expand Medicaid and so classification of its expansion status is somewhat ambiguous. In Virginia, 2014 plan premiums averaged over $1,000 per month for all age groups. We follow the convention of Gruber et

[26] source: $http://aspe.hhs.gov/health/reports/2014/MarketPlaceEnrollment/Apr2014$

al. (2014) in dropping this extreme outlier. Finally, we excluded Hawaii due to the absence of pre-ACA plan data. Table 1.9 summarizes all of the states included and excluded from the sample. In total, we are left with a repeated cross-section of plan characteristics, such as deductibles, coinsurance rates for various services, rejection and up-charge rates, and out-of-pocket spending limits, corresponding to all plans offered in 39 states (16 expansion and 23 non-expansion) across 415 rating areas.

After applying the above sample selection techniques to our data, we observe that only 36.5% of pre-reform plans offered in non-expansion states met minimum essential benefits as compared to 51.8% in expansion states. These differences are depicted graphically in Figure 1.8, which shows the distribution of qualifying pre-ACA plans by Medicaid expansion status across states. This again highlights the importance of accounting for changes in plan generosity for the purposes of evaluating any policy enacted simultaneously with the ACA reforms. Despite these differences in pre-ACA plan quality across Medicaid expansion status, Table 1.5 shows that the distribution of qualifying pre-ACA plans are similar across treatment and control groups.

Last, we augment our plan characteristics variables with small area measures of population demographics, income distributions, medical spending patterns, and health status at the state and rating area level (when available) using 2008-2012 samples from the American Community Survey (ACS) and Current Population Survey (CPS). We report these variables for both the full sample and for those earning $<$ 138% FPL in order to highlight how we expect Medicaid expansion decisions to impact the private insurance risk pools.

1.5 Regression Framework and Identification

To estimate the effect of state Medicaid expansions on private health insurance plan characteristics in the individual market, we use a difference-in-differences identification strategy that measures how plan characteristics evolve separately for expansion and non-expansion states following the ACA reforms. We then estimate a triple-difference model that exploits rating area level variation in the fraction of the population that would have gained Medicaid eligibility under a hypothetical expansion. The latter framework measures the level of exposure to this low-income group faced by private insurance markets.

We also compare our estimates to those produced using data from the post-reform cross-section only. If the ACA Medicaid expansion is truly a natural experiment, the cross-sectional results should match those generated by using across-time variation as well. Thus, we compare our cross-sectional results side-by-side with the across-time results as a way of assessing experimental validity.

1.5.1 Difference-in-Differences

We begin with our difference-in-differences (DD) estimating equation to explore the effect of state Medicaid expansions on private market deductibles at the rating area level, where subscripts i denote the rating area, s the state, and t the year (2011, 2013, or 2014), as follows:

$$\begin{split} deduct_{ist} = \beta_0 + \beta_1(post_t \times noexpansion_s) + \beta_2 state_s + \beta_3 year_t \\ + \beta_4 \phi_{ist} + \beta_5(year_t \times \delta_{is}) + \epsilon_{ist} \end{split} \tag{1.1}$$

Here, *deduct* is the average silver plan deductible offered in a rating area. *post* and *noexpansion* are indicator variables that denote that the plans are being offered on the 2014 exchanges and in a non-expansion state, respectively. The coefficient on the interaction of these two terms, *post*noexpansion*, is the coefficient of interest. The interpretation on β_1 is the change in deductible amounts offered on the private insurance exchange due to not expanding the Medicaid income eligibility threshold in a given state. We also estimate similar regressions with different dependent variables–out-of-pocket maximums, plan AV, and coinsurance rate–to detect (1) whether insurers are also reducing overall plan generosity in response to a Medicaid non-expansion and (2) if insurers are making compensatory changes to other plan characteristics in order to maintain the same AV. Additionally, we run this specification with an indicator for HMO plans as the dependent variable to see whether insurers offering plans in HMO networks are better able to expand in non-expansion markets.

β_1 will be an unbiased estimate of the effect of the Medicaid opt-out on plan characteristics provided that two main identifying assumptions are satisfied: (ID.1) Medicaid expansion decisions are based on the political ideology of state legislatures; (ID.2) Other than through the Medicaid expansion channel, changes to the composition of private insurance market risk pools would have been the same across expansion and non-expansion states (i.e., parallel trends).

Shor (2013) constructs political ideology scores and shows that these predict Medicaid expansion decisions with 94% accuracy. Moreover, a joint study by the Robert Wood Johnson Foundation (RWJF) and Urban Institute finds that, because the newest Medicaid expansions are largely financed by federal dollars, the 24 states opting out of the Medicaid expansion are foregoing $423.6 billion in savings between 2013 and 2022.[27] From a state budgetary perspective, these opt-out decisions are not individually rational, which provides additional evidence that political ideology

[27]source: $http://www.rwjf.org/content/dam/farm/reports/issue_briefs/2014$ $/rwjf414946$

is playing a pivotal role in determining treatment.

For the purposes of identification, it would also be problematic if Medicaid expansion decisions were correlated with ex-ante expected changes to the non-group market risk pools. However, once states with pre-ACA regulatory reforms are removed (i.e., New England states and states touching the Pacific Ocean), expansion and non-expansion states appear to be quite similar by measures of medical spending, pre-ACA composition of private insurance risk pools, and anticipated changes to the risk pool induced by all ACA measures outside of the Medicaid expansion. In Table 1.3, we compare individual characteristics across expansion and non-expansion states for both the full sample and the group earning less than 138% FPL. We see that medical out-of-pocket spending ($3,996 in E states vs. $3,827 in NE states) and non-group market participation rates (11% in E states vs. 10% in NE states) are similar across treatment and control. Measures of self-reported health status from the CPS are also similar (2.18 in E states vs. 2.25 in NE states). Additionally, we check to see whether non-expansion and expansion states will be differentially affected by the employer mandate. The employer mandate requires that all employers with 50 or more employees provide access to affordable insurance or else pay a "shared responsibility payment" of $2,000 per employee. Although the employer mandate was delayed until 2015, it is possible that firms will preemptively respond by dropping Employer Sponsored Insurance (ESI), which could affect private-market risk pools. However, Table 1.3 also shows that individuals in expansion and non-expansion states are equally likely to be employed by firms with 50 or more employees. These features combined lend credibility to the notion that in our study sample, Medicaid expansion decisions are not correlated with other factors that would have independently predicted differences in post-reform private risk pools.

In addition to controlling for the standard main effects in the interaction term (state and year fixed effects), we also control for ϕ, a vector that includes both the percentage of individuals rated up from the base premium rate and the percentage of individuals denied coverage for each plan. These medical underwriting intensity variables proxy for the anticipated changes in the risk pool due to community rating and guaranteed issue rules, respectively. ϕ also includes the fraction of pre-ACA plans meeting post-ACA requirements, which proxies for the segment of the market being served by ACA-eligible plans prior to the reforms. We also control for demographic variables interacted with year effects in $year \times \delta$ to account for the possibility that differences in population characteristics are correlated with expected health care spending patterns. These demographic variables include education, race, metro status, percent unemployed, and various other measures listed in Table 1.3. If inherent residual differences in 'red' states versus 'blue' states are driving the estimated effects then the coefficient of interest should be highly sensitive to the inclusion of this interaction term.

1.5.2 Difference-in-Difference-in-Differences

We next add our third source of variation, the fraction of the population in a given rating area that is low income ($<$ 138% FPL), ineligible for Medicaid, and uninsured or insured through a non-group insurance plan, i.e. the group that the Medicaid expansion would remove from the private market. From this point forward, we refer to this variable as the *Expansion Ineligibles*. We use pre-reform Medicaid income eligibility thresholds when constructing this variable. Figure 1.9 shows variation in our independent variable, which ranges from 0.6 percent to 11.5 percent across all rating areas in the United States.

We estimate the Difference-in-Difference-in-Differences (DDD) specification as follows:

$$
\begin{aligned}
deduct_{ist} = \alpha_0 &+ \alpha_1(post_t \times noexpansion_s \times Expansion\,Ineligibles_i) \\
&+ \alpha_2(post_t \times noexpansion_s) + \alpha_3(state_s \times Expansion\,Ineligibles_i) \\
&+ \alpha_4(year_t \times Expansion\,Ineligibles_i) + \alpha_5 year_t + \alpha_6 state_s \\
&+ \alpha_7 Expansion\,Ineligibles_i + \alpha_8 \phi_{ist} + \alpha_9(year_t \times \delta_{is}) + \epsilon_{ist}
\end{aligned}
\tag{1.2}
$$

The coefficient of interest is now α_1, which is multiplied by the triple interaction term, $(post_t \times noexpansion_s \times Expansion\,Ineligibles_i)$. α_1 tells us how a 1% increase in the size of the *Expansion Ineligibles* affects average plan deductibles offered in a non-expansion state. In this framework, we exploit both across and within-state variation in levels of exposure to the *Expansion Ineligibles.*

We similarly include the vectors, ϕ and $year_t \times \delta_{is}$ to control for medical underwriting intensity in the pre-ACA period and the potentially time-varying effects of demographic characteristics, respectively. To estimate the implied average effect of the Medicaid opt-out on deductibles, we multiply the triple interaction term by 5.9, which is the average size of *Expansion Ineligibles* in non-expansion states, and then add this product to α_2, the coefficient on $post_t \times noexpansion_s$. We obtain standard errors on this linear combination of coefficients by using the delta method. Finally, we estimate regressions for the four other dependent variables–annual out-of-pocket spending limits, general coinsurance rates, plan AV, and HMO plan offerings–in the DDD specification as well.

1.5.3 Validity check: Post-reform cross-section

If Medicaid expansion status were randomly assigned, one could simply compare plan characteristics in the post-ACA cross-section to determine causal effects.

Inclusion of the pre-ACA cross-sections in the analysis would only serve to improve precision. While Medicaid expansion decisions were not determined by dice rolls, we explore the possibility that they were quasi-randomly assigned with respect to the health plan offerings on state private insurance markets. This conjecture requires a stronger assumption in place of the parallel trends requirement: plan characteristics must not have only been trending the same across expansion and non-expansion states; they must have been the same in levels as well. We view this exercise as an overly strong check on the validity of the estimates of the difference-in-differences and triple difference framework. We follow the convention of comparing our main results side-by-side with their analogous estimates in the cross-section.

1.5.4 Comparison of Premiums in Rating Areas Along a State Border

[*We use this approach in the interim while we attempt to acquire pre-ACA premium data.*]

Following the methodology of Dube et al. (2010), we compare 2014 plan premiums in contiguous counties straddling a state border. This is a local identification approach that relies on the assumption that the rating areas on either side of a state border are approximately the same in terms of the health costs of the populations and the structure of the health insurance markets, except in relation to whether the states pursued a Medicaid expansion.

After restricting the sample to contiguous rating area pairs, we estimate the following OLS equation:

$$ln(premium)_{is} = \gamma_0 + \gamma_1(noexpansion_{is}) + \gamma_2(\delta_{is}) + \theta_{is} + \epsilon_{is} \tag{1.3}$$

δ is defined as before and ϕ is excluded since law dictates that both the upcharge and denial rates must be 0 following the insurance market reforms. We also add a contiguous border fixed effect, θ, to ensure that our comparison is restricted only to those rating areas sharing a state boundary line. Our identification assumption is: $Cov(noexpansion_{is}, \epsilon_{is}|\delta_{is}, \theta_{is}) = 0$, which requires that differences in Medicaid expansion decisions within the contiguous rating area pairs are conditionally uncorrelated with residual differences in premium prices in the rating areas. This essentially boils down to a matching technique for which the main matching criterion is geography.

One potential threat to this identification strategy is endogenous sorting of individuals across state borders. If unhealthy individuals near the poverty line in non-expansion states were to move across the state border in response to a Medicaid expansion in the contiguous state, this would invalidate our research design. However, a recent study by Schwartz and Sommers (2014) finds that Medicaid eligibility expansions did not induce interstate migration by low-income individuals. Another possible threat might arise if health care markets are very different across states. However, the Dartmouth Atlas constructs Hospital Service Areas (HSAs), or "local health care markets for hospital care," based on a collection of zip codes corresponding to residents who receive hospitalizations in the same area. These HSAs do not respect state boundaries, which provides a basis for our assumption that local health care markets are not state-specific.

1.6 Results

1.6.1 Visual Representation of Plan Characteristics

Figure 1.10 shows how the five key plan characteristics evolve separately across expansion and non-expansion states. One feature apparent in the data is that not only were nearly all plan characteristics trending similarly across expansion status prior to the reforms; they were also the same in levels. This provides evidence in support of the hypothesis that the Medicaid expansion decisions were quasi-random with respect to ex-ante plan features.

Subfigures 1.10a and 1.10b highlight the key findings in the paper; states opting out of the Medicaid expansions experienced a sharp increase in plan deductibles and accompanying decrease in plan generosity following the ACA reforms. As exhibited in subfigure 1.10d, coinsurance rates in non-expansion states also appear to have caught up to those in expansion states following the reforms. Subfigures 1.10c and 1.10e show that out-of-pocket limits and HMO plan offerings both experienced indistinguishably large increases in treatment and control states; HMO plan offerings, for example, increased from 5% to 30% across the entire sample. One explanation for why HMOs expanded similarly across the board is that HMOs in unregulated markets were far less likely to deny coverage or vary premiums on the basis of enrollee risk (Gabel (1997); Hall (2000)). Thus, they should have been less impeded by new community rating regulations than other types of networks since HMOs' primary method for controlling costs (i.e., restrictions on referrals and limited provider networks) were not affected by the ACA reforms.

To explore the mechanism underlying how Medicaid opt-out decisions affected plan deductibles and decreased plan generosity, we turn to the triple-difference results in Figure 1.11. Since *Expansion Ineligibles* is a continuous measure of exposure to the pivotal group, we distinguish rating areas on the basis of whether they were above or below the median within the sample. The distribution of *Expansion Ineligibles* is roughly symmetric about the mean and so we view the median as a natural partition for distinguishing between rating areas with high and low exposure. In subfigures 1.11a and 1.11b, it becomes clear that the relative increase in deductibles and decrease in generosity in non-expansion states is being driven entirely by rating areas with greater exposure to the *Expansion Ineligibles*. This finding provides evidence that the most recent round of Medicaid expansions removed high-cost individuals from the markets. It also suggests that insurers responded by raising cost-sharing requirements in markets that were particularly exposed to this adversely selected group. That plan characteristics in expansion states evolved identically across rating areas with differing levels of pre-ACA exposure to the *Expansion Ineligibles* provides additional evidence in support of this mechanism.

1.6.2 Regression Analysis of Plan Characteristics

In this section, we provide estimates of equations (1) and (2) for the full sample, which includes 2011, Quarter 1 of 2013, and 2014. We also estimate equations (1) and (2) using only the 2014 cross-section in order to assess the experimental validity of the Medicaid expansion treatment. We also provide placebo tests and robustness checks on all specifications. Additionally, we estimate equation (3) in both the full cross-section and set of contiguous areas straddling state border lines in order to provide a preliminary estimate of how Medicaid expansions impact private market premiums. Lastly, we combine our regression estimates with data from MEPS to simulate the average private out-of-pocket savings realized due to the Medicaid expansions.

In Table 1.6, we present our main difference-in-differences results for the effect of a Medicaid non-expansion on plan characteristics. Column (1) shows that non-expansion states experienced an increase in deductibles by \$300 (or 10.5%) and a decrease in plan generosity by 0.8%. While the point estimates are imprecise due to the short duration of the panel, the estimates are highly insensitive to the addition of demographic controls and pre-ACA underwriting intensity measures, which lends credibility to the initial point estimates. Moreover, when we estimate the analogous equation in the post-reform cross-section in Table 1.7, we similarly find that Medicaid opt-outs increase plan deductibles by \$266 and decrease plan generosity by 0.8-0.9%.

Next, we report our estimates of the more precise triple-difference specification, where treatment intensity varies both across and within states according to the level of private market exposure to the $< 138\%$ FPL group. Column (1) of Table 1.8 indicates that a 1% increase in the *Expansion Ineligibles* in NE states increases deductibles by a precisely estimated \$210. To calculate the implied average effect of a Medicaid opt-out on average deductibles, we multiply this point estimate by 5.9–the average size of the *Expansion Ineligibles*–and add it to the coefficient on the $post_t \times noexpansion_s$ term. Using this procedure, we find that a Medicaid non-expansion increases plan deductibles by a statistically significant \$410, or 14.4%. This estimate differs from that obtained in the difference-in-difference framework by only a few percentage points even though the two identification strategies are not mechanically linked to one another.

We also estimate the analogous triple-difference specification in the post-reform cross-section. Columns (1) and (3) of Table 1.9 reveal that a 1% increase in the *Expansion Ineligibles* increases deductibles by \$225 and decreases plan generosity by 0.9-1.0%. These estimates are again remarkably similar to those that we obtain when using across-time variation. The closeness of these estimates verifies that private plan characteristics were not only trending the same across expansion status but that they were also the same in levels. It also suggests that the June 2012 *National Federation of*

Independent Business (NFIB) v. Sebelius decision generated quasi-random variation in Medicaid opt-out decisions among the states whose individual markets were not tightly regulated prior to the ACA reforms.

1.6.3 Placebo Tests and Robustness Checks

We perform a series of robustness checks on our difference-in-differences and triple difference framework. First, we replace the post-ACA plan characteristics with plan data from Quarter 1 of 2013–a time period before the new ACA market regulations and health insurance exchanges went into effect–and re-run the two prior specifications as a placebo test for pre-trends. Tables 1.10 and 1.11 show that the estimated effect on plan characteristics is now close to 0 in the double-difference and triple-difference frameworks, respectively. This finding rules out the possibility that insurers were altering plan characteristics differently by Medicaid expansion status in anticipation of the ACA reforms.

We also check to see that the estimates are not sensitive to different sample selection criteria. One might be concerned that individuals in the Deep South have higher health costs on average. If this were true, then the insurers in this region could respond by raising cost-sharing measures to limit their financial exposure to this high-cost population. To address the possibility that these states are driving the results, we re-estimate equations (1) and (2) after removing Alabama, Georgia, Louisiana, Mississippi, and South Carolina from the sample. Results reported in Table 1.12 alleviate this concern as the estimated effect of non-expansion on deductibles and plan generosity are stable at 10% and -1.0%, respectively. In Table 1.13, we report the results for the triple-difference specification. We too find that the implied average effect of a Medicaid non-expansion on deductibles is an 11.7% increase, which is also indistinct from our full-sample estimate.

1.6.4 Preliminary Evidence on Plan Premiums

In Table 1.14, we present our results for the comparison of 2014 premiums in contiguous rating areas straddling a state border. We also include results for the conditional association between expansion status and premium prices for in the full sample for comparison. Medicaid expansions appear not to have any effect on silver plans using either cross-sectional sample. This rules out the possibility that non-expansion states are reducing premiums in order to compensate for higher cost-sharing requirements, which would have had ambiguous implications for consumer costs.

1.6.5 Simulated Reduction in Out-of-Pocket Spending Resulting from a Medicaid Expansion

Following the methodology of Pauly et al. (2014), we use MEPS data to simulate the change in expected average out-of-pocket costs resulting from a Medicaid expansion-induced increase in plan generosity. We do so by feeding individual MEPS spending data through the average silver plan characteristics in non-expansion states following ACA implementation[28] to determine the expected out-of-pocket costs for the given spending distribution. We subtract our regression-estimated average change in the deductible, coinsurance rate, and out-of-pocket maximum from the corresponding average non-expansion state plan characteristics to determine the counterfactual plan characteristics that would prevail under a hypothetical Medicaid expansion. We then re-feed the MEPs data through these counterfactual plan characteristics to estimate the associated expected out-of-pocket expenses. We refer to the resulting difference in expected out-of-pocket costs as the out-of-pocket savings attributable to a public health insurance expansion.[29]

[28]The average deductible, coinsurance rate, and out-of-pocket spending limit are \$3,115.87, 18.50%, and \$5,775.90, respectively

[29]We provide a detailed description of this simulation exercise in the appendix.

Using this methodology, we calculate the expected out-of-pocket spending in non-expansion states to be \$1,703.71 and the expected out-of-pocket spending in non-expansion states under a hypothetical Medicaid expansion to be \$1,642.44. This implies that expanding Medicaid reduces expected out-of-pocket spending by \$61.27, or 3.6 percentage points, through the channel of increased plan generosity. Although this reduction appears to be small, it is worth noting that 964 out of 1,309 MEPS observations face the same out-of-pocket costs whether a state expands Medicaid or not. 41 of these observations correspond to individuals whose total medical care spending exceeds \$19,000. The remaining 923 observations correspond to individuals spending less than \$2,815.87 (the implied average deductible when a non-expansion state expands Medicaid). Moreover, a quarter of the sample has reported medical spending that is less than \$125. Thus, the smallness of the main estimates are largely a function of the small fraction of individuals that fall in the intermediate range of spending (345 of 1309 observations, or 26.4% of the sample). On the contrary, the difference in expected out-of-pocket spending for this affected group is much larger at \$232.48, which represents a 6.2 percentage point reduction in expected out-of-pocket spending.

In Figure 1.12, we show the relationship between expected out-of-pocket spending and total medical spending under an average silver plan in non-expansion states and the analogous relationship for an average silver plan under a counterfactual Medicaid expansion. We then overlay these functions onto the empirical distribution of total medical spending in the MEPS database. Integrating the area between the two lines over this empitical distribution of total medical spending yields the expected reduction in out-of-pocket costs that would result from a Medicaid expansion. This graph illustrates that the expected difference in out-of-pocket spending owing to a Medicaid expansion is highly sensitive to the density of individuals falling in the intermediate range of spending (those whose medical expenditures are more than the average plan deductible but less than the amount of spending required to reach the out-of-pocket maximum).

These estimates likely understate the true out-of-pocket savings engendered by a Medicaid expansion for two reasons. First, our MEPS observations correspond to individuals participating in the non-group market from 2005-2011. To the extent that community rating reforms increased adverse selection on the non-group market, the medical spending distribution will have shifted to the right in post-reform era. Second, our simple back-of-the-envelope calculation assumes no moral hazard. Because overall plan generosity increased following the reforms, medical spending induced by lower prices will have increased spending within each bin, even among the same individuals. We accordingly view our simulated savings resulting from a Medicaid expansion as a lower bound on the true effect.

1.7 Conclusion

In this study, we evaluate the effect of a state's decision not to expand Medicaid income eligibility limits up to 138% FPL on deductibles and actuarial value for plans offered on the 2014 individual private health insurance exchanges. We assess these effects by constructing a data set of private health plan data spanning the period before and after the implementation of the 2014 private market reforms induced by the ACA. The incentive structure imposed by the ACA regulatory setting implies that the decision not to expand Medicaid increases the expected medical costs of those individuals who enter the private aggregate risk pool. We find that when faced with a more adversely selected risk pool, insurers raise the cost of private insurance by increasing plan deductibles and decreasing the actuarial value of plans.

Specifically, the decision not to expand Medicaid increases private plan deductibles by an average of 10.5% for silver plans, which equates to an increase of approximately $300 in annual deductibles paid on the private markets. These findings suggest that Medicaid expansion opt-outs impose negative externalities on individuals purchasing private insurance plans through exchanges. The size of this effect is proportional to the size of the poor population that enters the private markets due to the non-expansion; deductibles increase by roughly 8 percent (or $220) for each one percentage point increase in the fraction of individuals uninsured or privately insured, earning less than 138% FPL, and ineligible for Medicaid. We also find that insurers decrease plan generosity by roughly 1% for the types of plans most likely to be selected by this pivotal population and again find consistent evidence at the rating area level using variation in the fraction of individuals earning less than 138% FPL. Our MEPS simulations further indicate that the increased plan generosity resulting from a Medicaid expansion reduces private annual out-of-pocket costs by $232 per person (6.2% of medical spending) for those in the intermediate range of the medical spending distribution and by $61 per person (3.6% of out-of-pocket spending) over the entire risk pool. These findings also suggest that the assumption of constant plan generosity over time will understate the change in welfare attributable to the ACA, particularly in states that opted into the Medicaid expansion.

In a study of the early state-by-state impact of the ACA, Kowalski (2014) finds that both private coverage rates and average enrollee costs experienced relatively greater increases in Medicaid non-expansion states. This finding is consistent with the mechanism that we propose in this paper; namely, that *Expansion Ineligibles* enter the private markets.[30]

While non-expansion insurance markets are no more likely to have an increase in HMO plan offerings relative to expansion insurance markets, we do find evidence

[30] However, as Kowalski notes, this finding is observationally equivalent to an increase in crowd-out in Medicaid expansion states. Either mechanism is supported by our findings.

that overall HMO plan offerings increase from 5% to 30% after the ACA regulatory reforms. These results are consistent with those of Buchmueller and Liu (2005), Buchmueller and Dinardo (2002), and Lo Sasso and Lurie (2009), who find that these networks become increasingly prevalent as a tool for attracting better risks following the enactment of community rating regulations. Lastly, we assess the effect of Medicaid expansions on cross-sectional premium prices by comparing contiguous rating areas that straddle a state border for which either state has a different expansion status. Here we find no effect for silver plans, which is consistent with the findings of Gruber et al. (2014).

We have shown that Medicaid expansions generate externalities that also benefit consumers in private health insurance markets. Our study highlights the importance of considering many non-premium mechanisms by which insurers might respond to minimize losses on adversely selected insurance enrollees; it is the first to show that private insurers respond to increases in adverse selection pressures by increasing plan cost-sharing requirements. We hope this will add to the ongoing policy discussion about expanding Medicaid eligibility. Our current research design also provides a useful framework for studying the impact of future changes to state Medicaid income eligibility limits on private market premiums and plan characteristics on the 2015 private exchanges and beyond. As data become more readily available, researchers will be able to answer these and related questions with increasing precision.

1.8 Figures

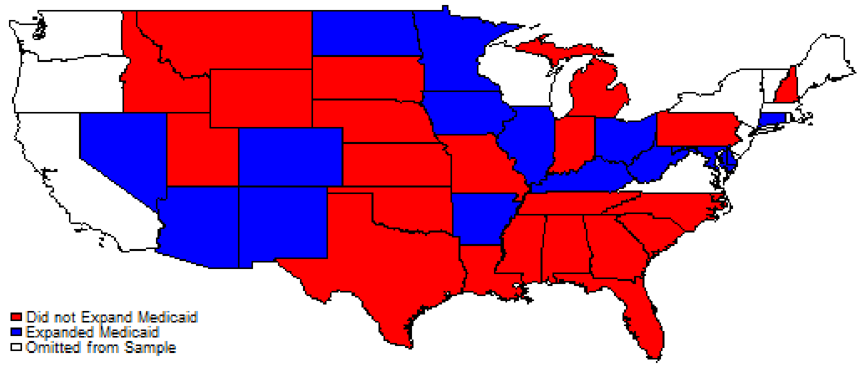


Figure 1.5: Medicaid Expansion Map

States that expand Medicaid are colored blue, while states that did not expand Medicaid by the end of the spring 2014 open enrollment period (3/31/14) are colored red. Omitted states include those with pre-ACA community rating and/or guaranteed issue rules, states with "active purchaser exchanges," and three miscellaneous states (Hawaii, Virginia, and Wisconsin)

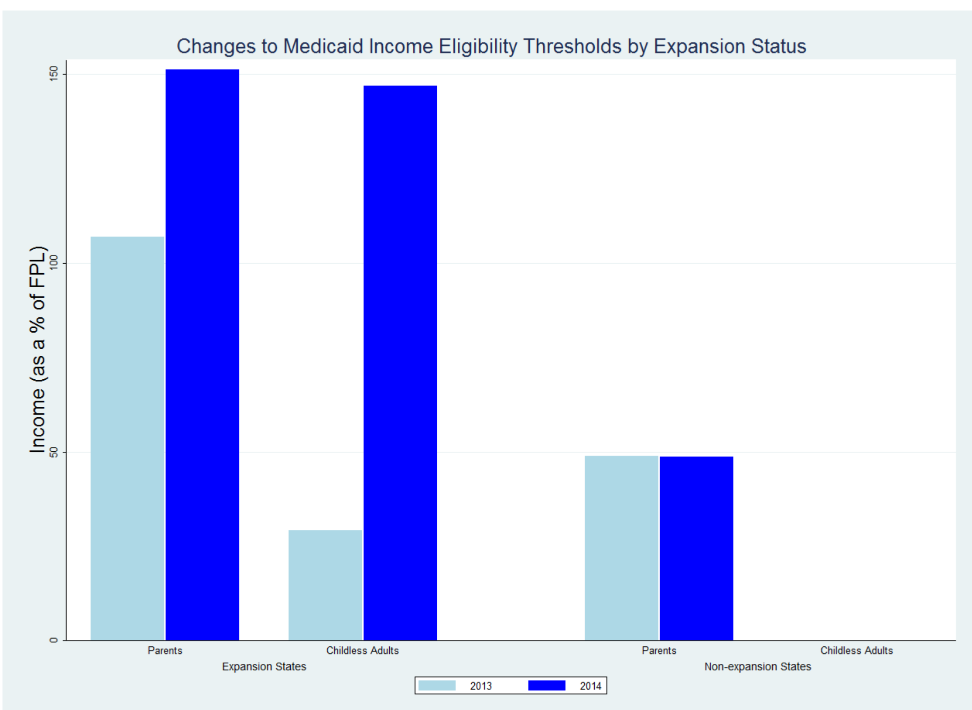


Figure 1.6: New Eligibility Thresholds

Omitted states include those with pre-ACA community rating and/or guaranteed issue rules, states with "active purchaser exchanges," and three miscellaneous states (Hawaii, Virginia, and Wisconsin)

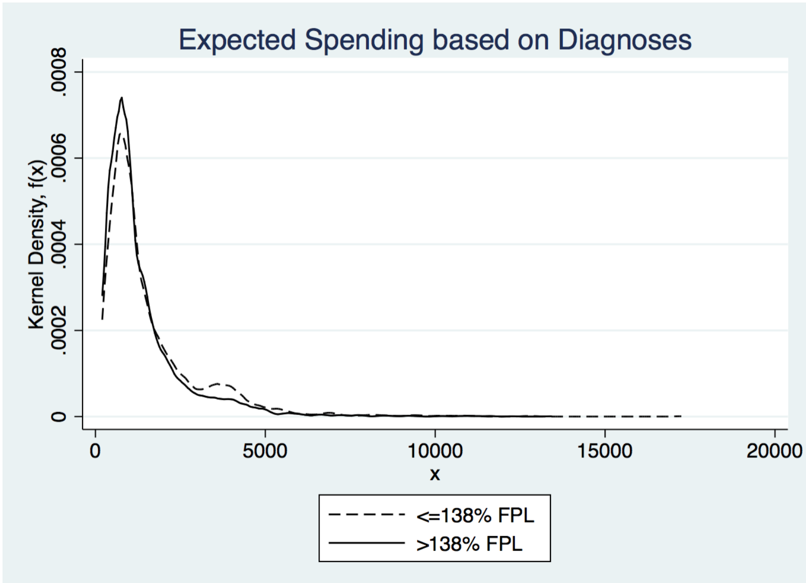


Figure 1.7: Distribution of Costs by Income Group
Using 2007-2009 MEPS data, we construct the distributions of expected healthcare spending for individuals based on whether they would be eligible for Medicaid coverage under a hypothetical ACA public insurance expansion. One would be eligible for coverage provided that he or she earns no more than 138% of the Federal Poverty Level.

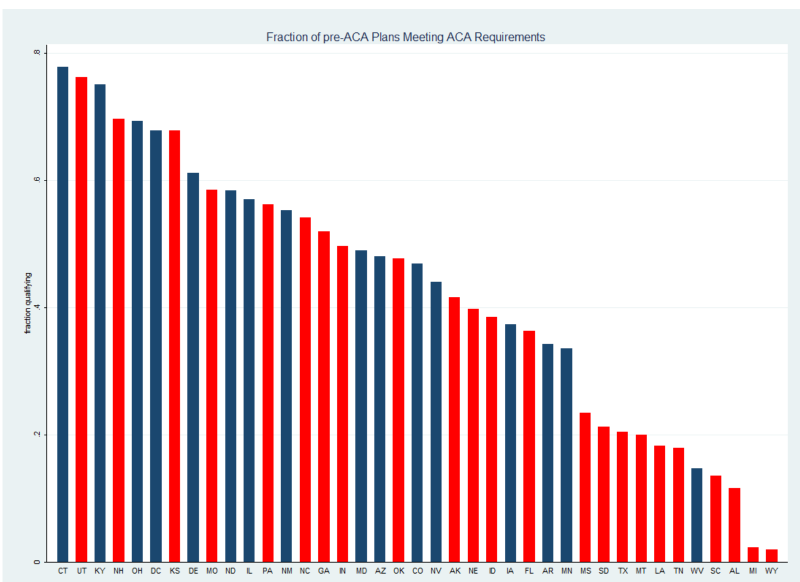


Figure 1.8: Fraction of Pre-ACA Plans Qualifying by Expansion Status
States that expand Medicaid are colored blue, while states that did not expand Medicaid by the end of the spring 2014 open enrollment period (3/31/14) are colored red. Overall, 51.8% of expansion state plans in the pre-ACA period would have met minimum benchmark requirements, as compared to only 36.5% of non-expansion state plans.

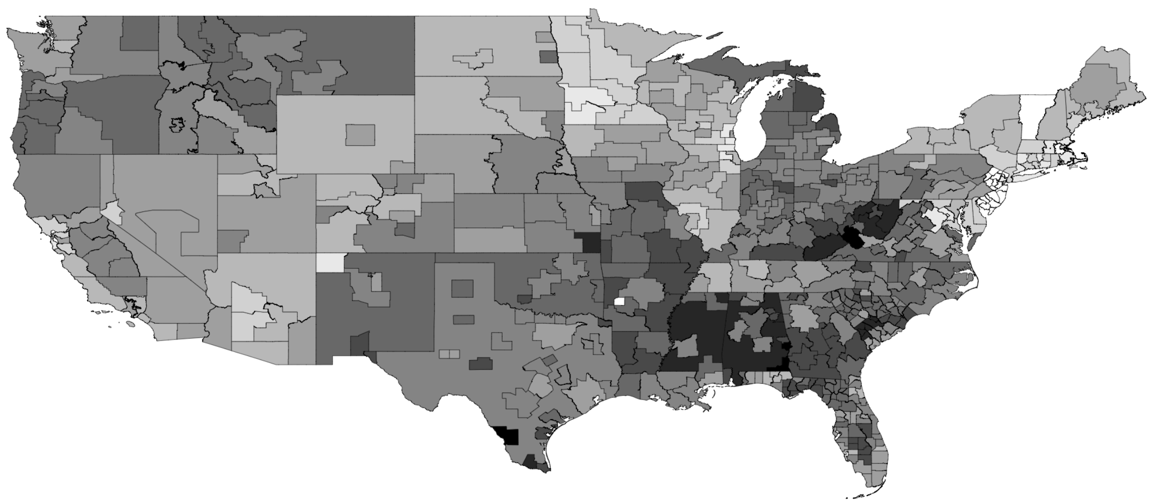
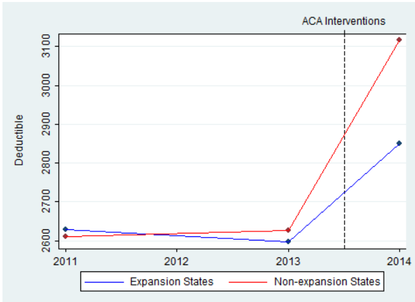
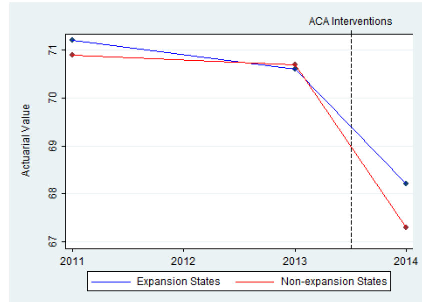


Figure 1.9: Fraction Uninsured, < 138% FPL by Rating Area

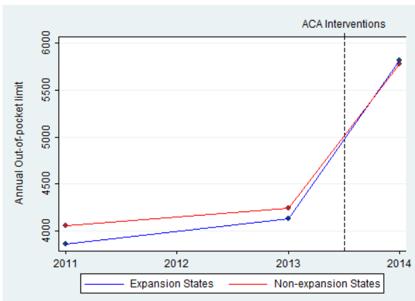
This map shows the fraction of residents (19-64 years old) in a rating area who earn < 138% FPL, ineligible for Medicaid using pre-ACA income thresholds, and uninsured or privately insured through the non-group market, by rating area. The color tone increases in proportion to this fraction, ranging from 0.5-11.5%.



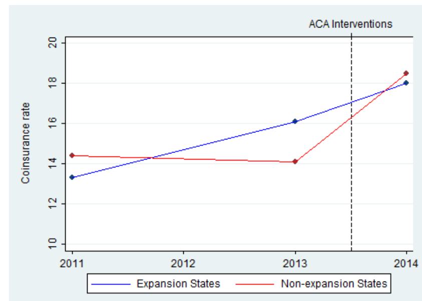
(a) Deductibles on Silver Plans



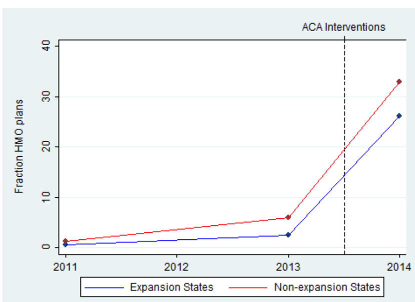
(b) Actuarial Value on Silver Plans



(c) Out-of-pocket Limits on Silver Plans

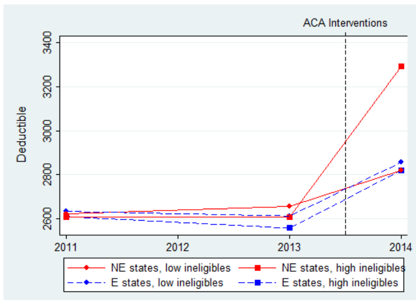


(d) Coinsurance Rates on Silver Plans

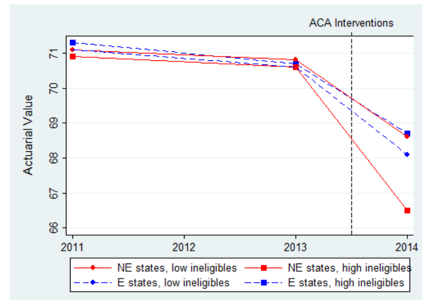


(e) Fraction HMO Plans (Silver)

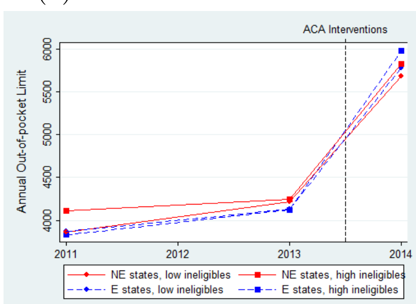
Figure 1.10: Plots of Silver Plan Characteristics, Double Difference



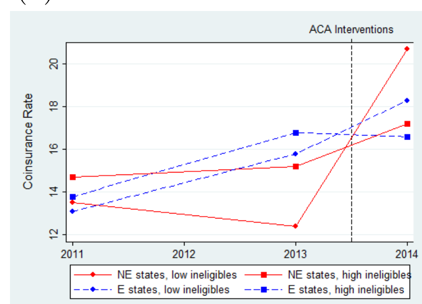
(a) Deductibles on Silver Plans



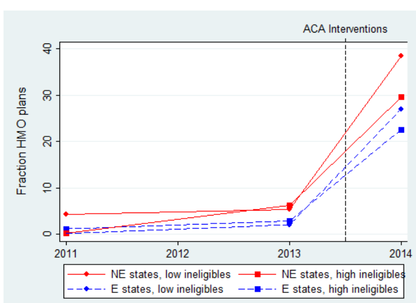
(b) Actuarial Value on Silver Plans



(c) Out-of-pocket Limits on Silver Plans



(d) Coinsurance Rates on Silver Plans



(e) Fraction HMO Plans (Silver)

Figure 1.11: Plots of Silver Plan Characteristics, Triple Difference

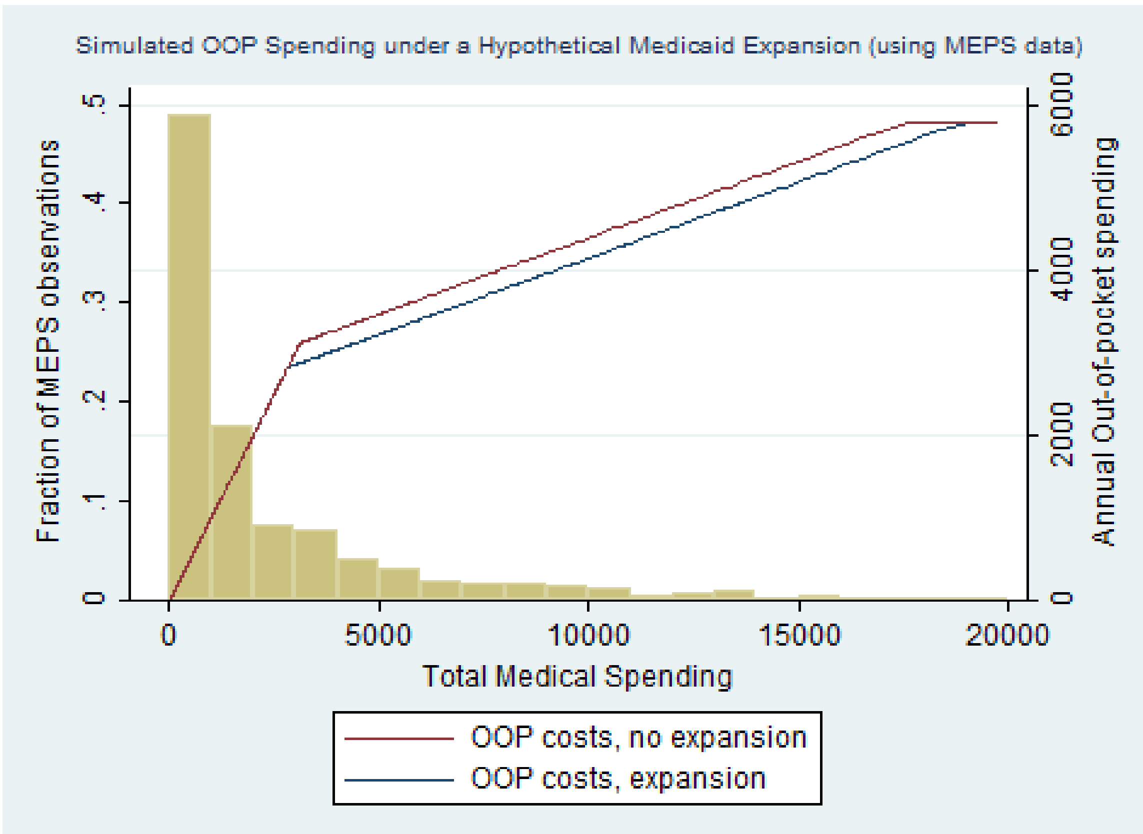


Figure 1.12: Simulated Change in Out-of-pocket Costs under a Hypothetical Medicaid Expansion

Using 2005-2011 MEPS data, we estimate the relationship between expected out-of-pocket costs and total medical spending using average silver plan characteristics in a non-expansion state. We then estimate the analogous relationship for the average silver plan that would prevail under a hypothetical Medicaid expansion. Integrating the area between the two lines over the empitical distribution of total medical spending (given by the histogram) yields the expected reduction in out-of-pocket costs that would result from a Medicaid expansion.

1.9 Tables

Table 1.1: Exchange Rules by Income Category and Expansion Status

Income Category (as a % of FPL)	Expansion Status	Medicaid Eligibility	Maximum Premium Cost	Mandate Exemption	Cost-Sharing Reductions*
<100	E	Yes	Not subsidy-eligible	Some**	No
<100	NE	Some***	Not subsidy-eligible	Yes	No
100-138	E	Yes	Not subsidy-eligible	No	No
100-138	NE	No	2-3.3% of income	Yes	Yes
138-250	E + NE	No	3.3-8.05% of income	No	Yes
250-400	E + NE	No	8.05-9.5% of income	No	No
>400	E + NE	No	Not subsidy-eligible	No	No

*Cost-sharing reductions apply to silver plans only
**Those earning under the income tax filing limits are exempt ($10,000 for individuals and $20,000 for couples)
***Some parents are eligible for Medicaid in non-expansion states according to income thresholds; all childless adults are ineligible

Table 1.2: Medicaid Expansion and Exchange Status

	In study sample	Omitted from sample
Expanded Medicaid	AZ AR CO CT DE DC IL IA KY MD ND OH MN NV NM WV	CA^{+} HI** MA*$^{+}$ NJ* NY*$^{+}$ OR*$^{+}$ RI^{+} VT*$^{+}$ WA*
Did Not Expand Medicaid	AL AK FL GA ID IN KS LA MI MS MO MT NE NH NC OK PA SC SD TN TX UT WY	ME* VA$^{\phi}$ WI***

* pre-ACA community rating

$^{+}$dropped due to exchange's "active purchaser" status

**Missing pre-ACA data

***Wisconsin is the only non-expansion state that covers childless adults (provided that they earn no more than the Federal Poverty Level) and so we omit it from the sample

ϕ 2014 silver plan premiums in Virginia averaged over $1,000 per month for all age groups. We follow the convention of Gruber et al. (2014) in dropping this extreme outlier.

Table 1.3: Means Table: Individual-Level Variables, Full Sample

	Expansion	Non-Expansion
Full Sample		
Age	41.15	41.07
Black	0.10	0.16
Hispanic	0.22	0.16
High School or Equivalent	0.41	0.46
College or More	0.56	0.51
Metro Area	0.80	0.73
Mean Income as % Federal Poverty Line (FPL)	337.51	313.32
Employed	0.92	0.90
Employed at Firm with 50+ Workers	0.63	0.63
Earned Income Tax Credit	156.79	191.70
Any Health Insurance Coverage	0.85	0.81
Medicaid Eligible Pre-ACA	0.09	0.04
Non-Group Coverage	0.11	0.10
Medical Out-of-Pocket Spending	3996.69	3827.47
Total Family Spending on Health Insurance	2035.35	1928.76
Self-Reported Health Status (1=Excellent to 5=Poor)	2.18	2.25
States	16	23
Rating Areas	117	302

Summary stats for those states that appear in the main sample (which excludes those states with pre-ACA community rating and those states that are active purchasers on the exchange). All variables from the American Community Survey (ACS) 2008-2012, except for "Employed at Firm with 50+ Workers" and "Earned Income Tax Credit," which are taken from the CPS 2008-2013, and "Medical Out-of-Pocket Spending," and "Total Family Spending on Health Insurance," which are taken from the CPS 2011-2013. CPS Variables use the cutoff $< 125\% FPL$ in the lower section due to the coarseness of the CPS data. All data is restricted to individuals 18-64 years old and drops non-citizens and Native Americans for whom many of the provisions of the law do not apply.

Table 1.4: Means Table: Individual-Level Variables, Low-Income Groups

	Expansion	Non-Expansion
$< 138\%\ FPL$		
Age	37.19	37.19
Black	0.17	0.28
Hispanic	0.31	0.19
High School or Equivalent	0.54	0.57
College or More	0.40	0.36
Metro Area	0.75	0.68
Mean Income as % Federal Poverty Line (FPL)	71.32	72.34
Employed at Firm with 50+ Workers	0.51	0.53
Earned Income Tax Credit	566.23	652.03
Any Health Insurance Coverage	0.68	0.60
Medicaid Eligible Pre-ACA	0.43	0.18
Non-Group Coverage	0.10	0.09
Medical Out-of-Pocket Spending	1486.24	1604.23
Total Family Spending on Health Insurance	568.34	538.02
Self-Reported Health Status (1=Excellent to 5=Poor)	2.63	2.73
States	16	23
Rating Areas	117	302

Summary stats for those states that appear in the main sample (which excludes those states with pre-ACA community rating and those states that are active purchasers on the exchange). All variables from the American Community Survey (ACS) 2008-2012, except for "Employed at Firm with 50+ Workers" and "Earned Income Tax Credit," which are taken from the CPS 2008-2013, and "Medical Out-of-Pocket Spending," and "Total Family Spending on Health Insurance," which are taken from the CPS 2011-2013. CPS Variables use the cutoff $< 125\% FPL$ in the lower section due to the coarseness of the CPS data. All data is restricted to individuals 18-64 years old and drops non-citizens and Native Americans for whom many of the provisions of the law do not apply.

Table 1.5: Means Table: Plan-Level Variables

	Expansion	Non-Expansion
Metal Tier		
bronze plan ($0.58 < AV < 0.62$)	0.22	0.21
silver plan ($0.68 < AV < 0.72$)	0.39	0.41
gold plan ($0.78 < AV < 0.82$)	0.29	0.28
platinum plan ($0.88 < AV < 0.92$)	0.10	0.10
Insurers		
number of insurers	5.29	4.24
number of plans	51.30	58.28
applicants charged more (pre-ACA only)	25.81	16.17
applicants denied (pre-ACA only)	22.51	18.37
N	25,182	73,425
States	16	23
Rating Areas	117	302

Summary stats for those states that appear in the main sample (which excludes those states with pre-ACA community rating and those states that are active purchasers on the exchange).

Table 1.6: Silver Plan Characteristics (Full Panel), Difference-in-Differences

	(1) Deductible	(2) Out-of-pocket limit	(3) AV	(4) Coinsurance	(5) HMO plan
Non-Expansion State× Post-ACA	300.5	-116.4	-0.00898	0.0105	0.0440
	(422.5)	(147.4)	(0.0174)	(0.0422)	(0.0921)
R^2	0.093	0.238	0.144	0.050	0.261
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	NO	NO	NO	NO	NO
Fraction Meeting ACA Requirements	NO	NO	NO	NO	NO
Demographics×year	NO	NO	NO	NO	NO
Non-Expansion State× Post-ACA	296.9	-141.7	-0.00808	0.0172	0.0781
	(258.1)	(103.2)	(0.0151)	(0.0209)	(0.0888)
R^2	0.145	0.247	0.167	0.084	0.306
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	YES	YES	YES	YES	YES
Fraction Meeting ACA Requirements	YES	YES	YES	YES	YES
Demographics×year	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 35,258 silver plans across 38 states and 414 rating areas from 2011, 2013, and 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.7: Silver Plan Characteristics (Post-reform Cross-section) Analogous Single Difference

	(1) Deductible	(2) Out-of-pocket limit	(3) AV	(4) Coinsurance	(5) HMO plan
Non-Expansion State	266.4	-45.11	-0.00929	0.00449	0.0672
	(423.3)	(110.6)	(0.0170)	(0.0371)	(0.0953)
R^2	0.054	0.657	0.159	0.181	0.363
Demographics	NO	NO	NO	NO	NO
Non-Expansion State	263.4	-112.0	-0.00883	-0.00175	0.133
	(275.9)	(93.53)	(0.0146)	(0.0220)	(0.0833)
R^2	0.085	0.030	0.043	0.115	0.106
Demographics	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include percent upcharged, percent denied, state fixed effects, percent black, percent hispanic, percent graduated high school, percent graduated college, percent unemployed, metropolitan status, percent aged 0-18 , percent aged 35-49, and percent aged 50-64. Standard errors, clustered at the state level, are in parentheses.. Sample includes 5,503 silver plans across 39 states and 414 rating areas in 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.8: Silver Plan Characteristics (Full Panel), Triple Difference

	(1) Deductible	(2) OOP limit	(3) AV	(4) Coinsurance	(5) HMO plan
Expansion Ineligibles × NE State × Post-ACA	215.2**	-51.73	-0.00808	-0.0279**	-0.00722
	(70.45)	(36.59)	(0.00504)	(0.00891)	(0.0294)
Non-Expansion State × Post-ACA	-934.8*	65.00	0.0353	0.171**	0.0653
	(409.3)	(235.1)	(0.0280)	(0.0570)	(0.163)
R^2	0.105	0.239	0.150	0.057	0.265
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	NO	NO	NO	NO	NO
Fraction Meeting ACA Requirements	NO	NO	NO	NO	NO
Demographics×year	NO	NO	NO	NO	NO
Expansion Ineligibles × NE State × Post-ACA	206.3**	-67.86	-0.00820+	-0.0249**	0.0176
	(66.25)	(46.75)	(0.00427)	(0.00776)	(0.0274)
Non-Expansion State × Post-ACA	-806.3+	133.0	0.0339	0.158**	-0.0595
	(458.1)	(247.5)	(0.0269)	(0.0492)	(0.157)
R^2	0.152	0.247	0.175	0.088	0.310
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	YES	YES	YES	YES	YES
Fraction Meeting ACA Requirements	YES	YES	YES	YES	YES
Demographics×year	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 35,258 silver plans across 414 rating areas from 2011, 2013, and 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.9: Silver Plan Characteristics (Post-reform Cross-section), Analogous Double Difference

	(1) Deductible	(2) OOP limit	(3) AV	(4) Coinsurance	(5) HMO plan
Expansion Ineligibles × NE State	225.3**	43.59	-0.00899+	-0.0192**	-0.0172
	(71.98)	(42.94)	(0.00483)	(0.00631)	(0.0329)
Non-Expansion State	-991.4*	-318.2	0.0402	0.115**	0.152
	(366.7)	(208.9)	(0.0261)	(0.0317)	(0.203)
R^2	0.029	0.006	0.019	0.025	0.005
Demographics	NO	NO	NO	NO	NO
Expansion Ineligibles × NE State	224.3**	34.66	-0.00974*	-0.0160**	-0.00635
	(68.39)	(37.71)	(0.00429)	(0.00575)	(0.0276)
Non-Expansion State	-919.7+	-338.7	0.0443	0.0901*	0.127
	(456.0)	(200.2)	(0.0263)	(0.0386)	(0.174)
R^2	0.100	0.033	0.056	0.128	0.110
Demographics	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 35,258 silver plans across 414 rating areas from 2011, 2013, and 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.10: Placebo Test for Pre-trends, Double Difference

	(1) Deductible	(2) Out-of-pocket limit	(3) AV	(4) Coinsurance	(5) HMO plan
Non-Expansion State× Post-ACA	36.59	-9.320	0.000867	-0.0211+	0.0373
	(27.49)	(86.70)	(0.00169)	(0.0111)	(0.0241)
R^2	0.035	0.070	0.036	0.065	0.222
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	YES	YES	YES	YES	YES
Fraction Meeting ACA Requirements	YES	YES	YES	YES	YES
Demographics×year	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 29,755 silver plans across 39 states and 414 rating areas in 2011 and 2013. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.11: Placebo Test for Pre-trends, Triple Difference

	(1) Deductible	(2) Out-of-pocket limit	(3) AV	(4) Coinsurance	(5) HMO plan
Expansion Ineligibles × NE State × Post-ACA	3.494	-3.660	-0.0000851	0.00428	0.00917
	(7.074)	(26.74)	(0.000393)	(0.00358)	(0.00682)
Non-Expansion State × Post-ACA	18.33	51.28	0.00146	-0.0492*	-0.00529
	(44.19)	(168.5)	(0.00264)	(0.0222)	(0.0562)
R^2	0.035	0.070	0.036	0.066	0.223
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	YES	YES	YES	YES	YES
Fraction Meeting ACA Requirements	YES	YES	YES	YES	YES
Demographics×year	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 29,755 silver plans across 39 states and 414 rating areas in 2011 and 2013. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.12: Robustness Check (Full Panel), Double Difference: Exclusion of Deep South States

	(1) Deductible	(2) Out-of-pocket limit	(3) AV	(4) Coinsurance	(5) HMO plan
Non-Expansion State × Post-ACA	264.4	-112.1	-0.0101	0.0147	0.0680
	(250.7)	(93.57)	(0.0149)	(0.0210)	(0.0891)
R^2	0.175	0.242	0.191	0.070	0.306
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	YES	YES	YES	YES	YES
Fraction Meeting ACA Requirements	YES	YES	YES	YES	YES
Demographics×year	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 29,828 silver plans across 33 states and 325 rating areas in 2011, 2013, and 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.13: Robustness Check (Full Panel), Triple Difference: Exclusion of Deep South States

	(1) Deductible	(2) OOP limit	(3) AV	(4) Coinsurance	(5) HMO plan
Expansion Ineligibles × NE State × Post-ACA	243.5**	8.000	-0.0125**	-0.0199*	-0.00525
	(77.87)	(36.08)	(0.00440)	(0.00793)	(0.0225)
Non-Expansion State × Post-ACA	-1052.5+	-269.2	0.0569*	0.131*	0.0630
	(544.1)	(184.3)	(0.0279)	(0.0551)	(0.140)
R^2	0.185	0.244	0.202	0.083	0.308
State FE	YES	YES	YES	YES	YES
Medical Underwriting Intensity	YES	YES	YES	YES	YES
Fraction Meeting ACA Requirements	YES	YES	YES	YES	YES
Demographics×year	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 29,828 silver plans across 33 states and 325 rating areas in 2011, 2013, and 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.14: Cross-sectional Premiums from 2014 Exchanges

	Dep var = ln(premium)	
	Contiguous RA Sample	Full Cross-Section
Silver Plans		
Non-Expansion State	0.0148	-0.0051
	(0.0186)	(0.0191)
N	2266	5243
Rating Areas	101	414
R^2	0.076	0.049
Demographics	YES	YES

Dependent variable is the log premium price for a 27 year old non-smoker. Standard errors for the contiguous RA and full cross-sectional sample, clustered at the rating area pair and rating area level, respectively, are in parentheses. All plans were offered on the insurance exchanges in 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

1.10 Acknowledgements

Chapter 1, in full, is currently being prepared for submission for publication of the material. Knepper, Matt; Niedzwiecki, Matthew. The dissertation author was a co-primary investigator and co-author of this material.

1.11 Appendix

Steps for Simulating Out-of-Pocket Spending under a Hypothetical Medicaid Expansion (similar to Pauly et al. [2014])

1. Adjust MEPS total national health spending data to account for differences in medical spending per capita between within-sample expansion and non-expansion states. To do this, we multiply total spending of each MEPS observation by .9822, as non-expansion state spending is 94.67 percent of spending in expansion states and within-sample non-expansion states account for roughly two thirds of the total sample population ($.9822 = .9467 \times (.3338) + 1 \times (.6662)$).

2. Retrieve MEPS data from 2005-2011 and inflate total health care expenditures to 2014 dollars.

3. Calculate the average total expected out-of-pocket spending under an average silver plan in a non-expansion state (The average deductible, coinsurance rate, and out-of-pocket spending limit are \$3,115.87, 18.50%, and \$5,775.90, respectively) for individuals aged 27-64, privately insured, and earning no less than 2 times the Federal Poverty Level. Repeat this exercise for our counterfactual average silver plan characteristics under a hypothetical Medicaid expansion (The average deductible is \$300.20 lower according to our regression-adjusted estimates, while the coinsurance rates and out-of-pocket spending limits are unchanged).

3. i First, calculate total expected out-of-pocket spending under an average silver plan in a non-expansion state:

Let $medspend$ = total medical spending, $deduct_{ne}$ = the average silver plan deductible in a non-expansion state, $oopmax$ = the average out-of-pocket maximum for silver plans, $medspend(oopmax)_{ne}$ = the minimum level of total medical spending at which the insuree reaches his or her out-of-pocket spending limit, $coins$ = the average coinsurance rate for silver plans

Then E[OOP costs|non-expansion] =

$$\begin{cases} medspend & \text{if } medspend < deduct_{ne} \\ deduct_{ne} + coins \times (medspend - deduct_{ne}) & \text{if } deduct_{ne} < medspend < medspend(oopmax)_{ne} \\ oopmax & \text{if } medspend \geq medspend(oopmax)_{ne} \end{cases}$$

Note further that one can solve for $medspend(oopmax)_{ne}$ explicitly by rearranging the following equation:

$$maxoop = deduct_{ne} + coins \times (medspend(oopmax)_{ne} - deduct_{ne}),$$

which yields

$$medspend(oopmax)_{ne} = \frac{(maxoop - deduct_{ne})}{coins} + deduct_{ne}$$

3. ii Last, calculate total expected out-of-pocket spending under an average silver plan under a hypothetical expansion:

Now let $deduct_e = deduct_{ne} - \$300.20$
Then E[OOP costs|expansion] =

$$\begin{cases} medspend & \text{if } medspend < deduct_{e} \\ deduct_{e} + coins \times (medspend - deduct_{e}) & \text{if } deduct_{ne} < medspend < medspend(oopmax)_{e} \\ oopmax & \text{if } medspend \geq medspend(oopmax)_{e} \end{cases}$$

Note also that $medspend(oopmax)_e > medspend(oopmax)_{ne}$

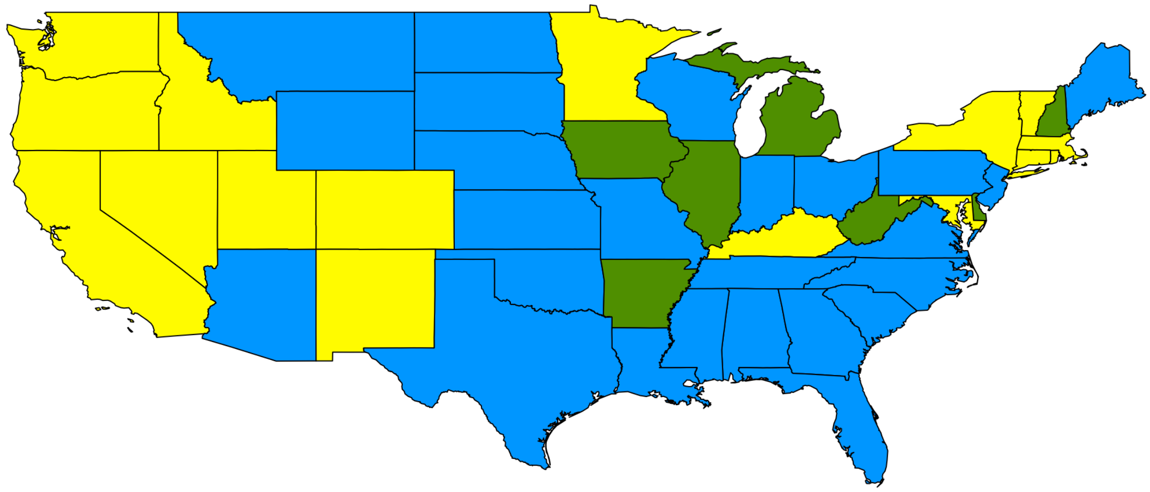


Figure 1.13: Health Insurance Exchanges
States whose exchanges are run by the federal government (default option) are colored blue. States running their own exchanges are colored yellow. States collaborating with the federal government are colored green.

Table 1.15: Medicaid Expansion and Exchange Status

	Federal Exchange	Partnership Exchange	State Exchange
Expanded Medicaid	AZ, NJ, ND, OH	AR, DE, IL, IA, WV	CA^{+}, CO, CT, DC, HI***, KY, MD, MA^{*+}, MN, NV, NM, NY^{*+}, OR^{*+}, RI^{+}, VT^{*+}, WA*
Did Not Expand Medicaid	AL AK FL GA IN KS LA ME* MS MO MT NE NC OK PA SC SD TN TX VA WI** WY	MI, NH	ID, UT

Exchanges are either run by the federal government (federal), run entirely by the state (state), or run as a partnership between the state and federal government (partnership).

*in sample but dropped due to pre-ACA community rating

$^{+}$in sample but dropped due to active purchaser status

**Wisconsin is the only non-expansion state that covers childless adults (provided that they earn no more than the Federal Poverty Level) and so we omit it from the sample

***Missing pre-ACA data

Table 1.16: Robustness Check (Post-reform Cross-section), Single Difference: Exclusion of Deep South States

	(1) Deductible	(2) Out-of-pocket limit	(3) AV	(4) Coinsurance	(5) HMO plan
Non-Expansion State	260.1	-115.6	-0.00955	-0.00379	0.130
	(255.3)	(93.69)	(0.0141)	(0.0208)	(0.0830)
R^2	0.097	0.035	0.056	0.101	0.093
Demographics	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Sample includes 4,629 silver plans across 33 states and 325 rating areas in 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.17: Robustness Check (Post-reform Cross-section), Double Difference: Exclusion of Deep South States

	(1) Deductible	(2) OOP limit	(3) AV	(4) Coinsurance	(5) HMO plan
Expansion Ineligibles × NE State	270.9**	63.32	-0.0142**	-0.0130+	-0.0343
	(73.49)	(38.70)	(0.00455)	(0.00659)	(0.0245)
Non-Expansion State	-1197.0*	-507.5*	0.0675*	0.0686	0.288+
	(507.6)	(202.0)	(0.0273)	(0.0438)	(0.166)
R^2	0.120	0.043	0.081	0.109	0.096
Demographics	YES	YES	YES	YES	YES

Dependent variables are indicated in the column headings. Regressors include all of those indicated in footnotes to Table 1.7. Standard errors, clustered at the state level, are in parentheses. Sample includes 4,629 silver plans across 325 rating areas from 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.18: Plan Deductibles (Difference-in-Differences)

	Dep Var = Deductible			
	(1)	(2)	(3)	(4)
Bronze Plans				
Non-Expansion State × Post-ACA	31.91	35.13	19.58	18.52
	(78.74)	(81.47)	(86.40)	(92.30)
R^2	0.051	0.051	0.052	0.054
Silver Plans				
Non-Expansion State × Post-ACA	300.5	249.9	320.5	296.9
	(422.5)	(398.5)	(377.2)	(258.1)
R^2	0.093	0.103	0.109	0.145
State FE	YES	YES	YES	YES
Medical Underwriting Intensity	NO	YES	YES	YES
Fraction Meeting ACA Requirements	NO	NO	YES	YES
Demographics×year	NO	NO	NO	YES

Dependent variable is the annual deductible. Regressors include non-expansion and post-ACA dummy variables (along with their pairwise interaction), percent upcharged, percent denied, state fixed effects, percent black, percent hispanic, percent graduated high school, percent graduated college, percent unemployed, metropolitan status, percent aged 0-18 , percent aged 35-49, and percent aged 50-64. Standard errors, clustered at rating area level, are in parentheses. Sample includes 16,766 bronze plans and 35,258 silver plans across 414 rating areas from 2011, 2013, and 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 1.19: Plan Deductibles (Difference-in-Difference-in-Differences)

	Dep Var = deductible			
	(1)	(2)	(3)	(4)
Bronze Plans				
ExpansionIneligibles × NE State × Post-ACA	52.64	55.40	56.16	46.67
	(47.82)	(47.77)	(46.96)	(46.02)
Non-Expansion State × Post-ACA	-188.6	-199.7	-221.4	-176.1
	(262.3)	(262.5)	(261.2)	(264.2)
R^2	0.052	0.053	0.053	0.056
Silver Plans				
ExpansionIneligibles × NE State × Post-ACA	215.2**	202.4**	202.4**	206.3**
	(70.45)	(67.86)	(64.50)	(66.25)
Non-Expansion State × Post-ACA	-934.8*	-921.4*	-846.9*	-806.3^{+}
	(409.3)	(410.4)	(413.5)	(458.1)
R^2	0.105	0.114	0.120	0.152
State FE	YES	YES	YES	YES
Medical Underwriting Intensity	YES	YES	YES	YES
Fraction Meeting ACA Requirements	YES	YES	YES	YES
Demographics×year	YES	YES	YES	YES

Dependent variable is the annual deductible. Uninsured variable represents the fraction of residents (19-64 years old) in a rating area who were uninsured or purchased private insurance not through an employer. Regressors include fraction uninsured and earning $< 138\% FPL$, post-ACA and non-expansion dummies (and all of their pairwise interactions along with the triple interaction), percent upcharged, percent denied, state fixed effects, percent black, percent hispanic, percent graduated high school, percent graduated college, percent unemployed, metropolitan status, percent aged 0-18 , percent aged 35-49, and percent aged 50-64. Standard errors, clustered at rating area level, are in parentheses. Sample includes 16,766 bronze plans and 35,258 silver plans across 414 rating areas from 2011, 2013, and 2014. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Chapter 2

When Case Settlement Rates are Unsettling: A Test for Detecting Judicial Biases with an Application to Workplace Discrimination Cases

2.1 Introduction

The Equal Employment Opportunity Commission (EEOC) is responsible for enforcing laws against workplace discrimination on the basis of race, sex, disability, national origin, and age.[1] In particular, the EEOC decides, from the set of nearly 90,000 discrimination charges filed annually, which cases to bring forth in the court of law. Though they ostensibly have the authority to punish employers that exercise taste-based discrimination, they are only effective insofar as the court appropriately punishes offending parties when merited. Where the arbiters of the law themselves are biased on the basis of race, sex, disability, etc., the ability of the EEOC to protect workers from unlawful behavior may be compromised. Moreover, failing to punish employers guilty of engaging in discriminatory practices may well propagate future unlawful behavior to the extent that successful claims have a deterrence effect.[2]

Given the high stakes of workplace discrimination cases, other researchers have aimed to detect whether justices are in fact biased against particular social groups that are different from their own. However, many of these studies limit their analysis solely to cases that reach the litigation stage. This practice drastically limits the sample size but perhaps more importantly, ignores the underlying forces that govern how a case proceeds through the legal system and leads to the accruement of a nonrandom sample of cases.

A seminal model of dispute selection developed by Priest and Klein (1984) (hereafter PK) explicitly provides the foundation for thinking about the factors that determine whether a case will bypass the settlement stage. PK show that under mild yet sensible conditions on litigation and settlement costs, a case will proceed to litigation only when there is sufficient uncertainty over the expected case outcome

[1]source: http://eeoclitigation.wustl.edu/
[2]source: This is a topic for future research that I hope to explore

across the plaintiff and defendant. This uncertainty, in turn, increases the closer is the defendant's fault level to the judge's minimum fault level threshold above which he or she will rule in favor of the plaintiff. In this paper, I extend the PK model to show that so long as cases of a particular type are randomly assigned to judges, a difference in case settlement rates by judge sex necessarily means that male and female judges have different decision rules, i.e. there is a judicial bias.

Furthermore, I show that, even under random assignment of judges to cases, a simple comparison of plaintiff victory rates across male and female judges can systematically fail to detect a judicial bias where one is in fact present. I then leverage a database of approximately 2,300 randomly selected EEOC litigation cases and find evidence consistent with the existence of a judicial gender bias; in particular, female plaintiffs filing workplace sex discrimination claims are 8-9% more likely to settle and 6-7% more likely to win compensation whenever a female justice is assigned to the case. These findings highlight the importance of considering judge sex-specific settlement rates in determining whether the set of litigated cases originate from different parts of the underlying distribution of defendant fault levels. This important omission explains why the prior literature has collectively underestimated the amount of judicial gender bias prevailing in workplace sex discrimination cases.

The following section describes the process by which judges are assigned to federal district court cases and summarizes the prior literature on judicial bias. Section 3.3 reproduces and extends the PK model in order to generate testable predictions for determining the existence of a judicial bias. Section 3.4 describes the data and how the study sample was constructed. Section 2.5 details the judge randomization tests and estimation strategy. Section 3.6 describes the specific empirical results of the test for judicial bias in workplace sex discrimination cases and section 3.8 concludes.

2.2 Institutional Background

2.2.1 Random Assignment of Judges to Cases in Federal District Courts

Prior to the early 1970s, federal district courts utilized a master calendar system whereby the chief judge of the district allocated judicial duties to judges over the course of the year. In this system, the chief judge could then subjectively assign cases to judges according to judge experience, expertise, and interest. Apart from being inefficient, the system made it easy for plaintiffs to 'judge-shop' since claimants could predict which judge would be assigned according to when the case was filed. Federal district courts have since replaced the master calendar system with an individual calendar system under which assignment is made on a rotational or otherwise random basis. Under the prevailing system, inexperienced and veteran judges are equally likely to be assigned complex cases (Bird (1975)). Critically, this allows us to conclude that any differences in sex discrimination case outcomes across female and male justices are caused by the bundle of characteristics of the assigned judges.[3]

According to the Federal Judicial Center, "Each court with more than one judge must determine a procedure for assigning cases to judges. Most district and bankruptcy courts use random assignment, which helps to ensure a fair distribution of cases and also prevents 'judge shopping,' or parties attempts to have their cases heard by the judge who they believe will act most favorably. Other courts assign cases by rotation, subject matter, or geographic division of the court."[4] The United States Courts website further remarks that the majority of courts use some form of a random drawing, simply rotating the names of available judges.[5]

[3] I perform a series of tests for judge randomization in Section 2.5.1.
[4] source: http://www.fjc.gov/federal/courts.nsf
[5] source: http://www.uscourts.gov/Common/FAQS.aspx

2.2.2 Prior Literature on Judge Characteristics and Case Outcomes

There is a substantial and growing body of literature that exposes the existence of ‘ingroup’ judicial biases. Even though judges are expected to adhere to a strong norm of non-discrimination, recent evidence has demonstrated that they are prone to offering preferential treatment to the social group to which they belong. These biases have been demonstrated along the dimensions of race (Abrams et al. (2012); Chew and Kelley (2008); Crowe (1999); Cameron and Cummings (2003)) and ethnicity (Shayo and Zussman (2011); Gazal-Ayal and Sulitzeanu-Kenan (2010)).

Analogous studies have examined the existence of a gender bias among judges with mixed results. A review of the legal literature supports the hypothesis that differences in the ways in which female and male judges rule appear most prominently in sex discrimination cases (Chew (2010)). However, much of the literature focuses on appellate rather than district court decisions. Appellate cases, unlike district court cases, represent a nonrandom subsample of federal cases. They arise only when the losing party asks the court of appeals to review the case and are only an option for cases that initially go to trial.[6] These features potentially limit the generalizability of what can be learned from judge voting behavior in appeals court cases. Nonetheless, after matching judges by age, ideology, and time period, Boyd et al. (2010) finds that the inclusion of one female judge on the three-judge panel increased the probability of plaintiff success by 20% in sex discrimination cases but not for other types of cases. A bevy of other studies find substantively similar results (Davis et al. (1993); Farhang and Wawro (2004); Songer et al. (1994); Peresie (2005); Crowe (1999)). Massie et al. (2002) also finds gender biases in the outcomes of the broader set of cases involving civil liberties.

[6] source: http://www.fjc.gov/federal/courts.nsf

Since appeals cases constitute a subsample of the set of litigated cases, they may not be representative of the unconditional distribution of cases filed with federal courts. Studies of gender biases in sex discrimination cases in federal district courts, however, have been far less conclusive. This ambiguity is due to three main factors: the restriction to published court opinions, a paucity of female justices in the data, and the analysis' exclusion of settled cases.

As Kim (2009) remarks, a common issue with many of these studies is that the data are extracted from published opinions only, which constitutes only one quarter of all briefed and submitted cases. Published and unpublished opinions also vary systematically; published opinions tend to be more politically motivated and ones for which judges are less likely to exercise personal discretion since they are subject to public scrutiny. Cases with unpublished opinions are also disproportionately difficult cases. For example, Songer et al. (1994) found that more than half of appeals cases with reversals or non-unanimous decisions were reviewing unpublished district court opinions.

An additional complication in the literature is that because female justices have been historically underrepresented, existing studies contain small sample sizes that preclude the researchers from uncovering precise effects as in Segal (2000), Manning (2004), and Kulik et al. (2003). Even though Ashenfelter et al. (1995), who also find no evidence of a judicial gender bias, study both the published and unpublished opinions of randomly assigned judges, their 1980-1981 sample contains only 5 female judges out of 47 total judges across three federal district courts. Ashenfelter et al. (1995) is also unique in that it tracks the rates with which plaintiffs win or settle. The seminal dispute selection model of Priest and Klein (1984) explores the extent to which litigated cases, which make up only 5% of all civil cases, are not representative. Thus, when all cases that are resolved prior to litigation are excluded, one misses the opportunity to learn about judicial biases and perceived judicial biases from the cases whose outcomes are far more common.

2.3 Model

In this section I reproduce the model of PK, which I then use to develop an empirical test for detecting judicial bias across a given judge background characteristic. I also describe other relevant conclusions that can be drawn from the results of such a test.

2.3.1 Disputes and the Decision Standard

As in PK, I define a dispute as a legal claim filed that states that the actions undertaken by the defendant have caused the plaintiff to incur damages. The dispute is resolved either by an out-of-court settlement reached by the dueling parties or, failing to reach such an agreement, by a verdict in the court of law. I call the former resolutions "settlements" and the latter "litigated cases." When a case is litigated, it must reach one of two possible outcomes: a plaintiff verdict or defendant verdict. When such a case reaches trial and the third-party arbiter, or judge, rules in favor of the plaintiff, the defendant is required to pay the plaintiff some amount of damages determined before the proceedings. If the judge instead rules in favor of the defendant, the defendant then pays nothing to the plaintiff and suffers only the costs of litigation. Independent of the amount by which the defendant may or may not compensate the plaintiff, both parties incur either the cost of litigation or cost of settlement depending upon the case resolution.

Of particular importance is the assumption that all judges or juries apply a specific standard for evaluating whether the defendant is at fault; I call this benchmark the decision standard, Y*. Let Y* be a scalar that is fixed for a given judge. Then let X be the set of case characteristics used by the judge to determine the level of defendant fault, Y. As in PK, let Y = $H(X)$, where $H(X)$ is a mapping from the case characteristics to a determined level of defendant fault. Whenever Y>Y*, the perceived level of defendant fault is above the minimum threshold required for the judge to rule in favor of the plaintiff. Conversely, if Y<Y*, the judge would rule in favor of the defendant. As in PK, I also assume that the plaintiffs and defendants know the decision standard Y* once assigned a judge. Figure 2.1 shows an arbitrary distribution of disputes so as to depict the relationship between Y, Y*, and the outcome of the case conditional on reaching litigation.

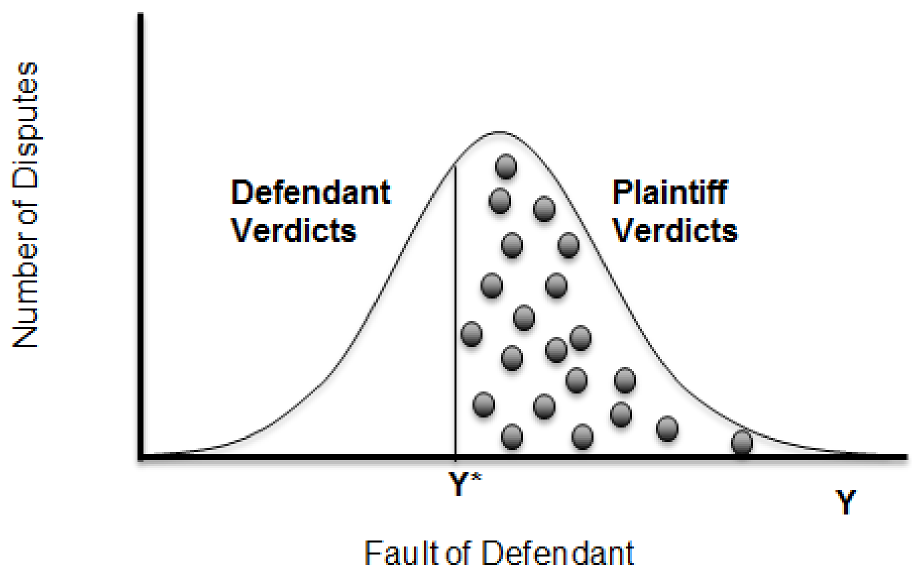


Figure 2.1: Decision Standard and Plaintiff Outcomes

2.3.2 Parties' Expectations of the Defendant Fault Level

Suppose that Y', the true level of defendant fault in a case, is randomly drawn from the distribution of disputes. Also assume that each party, the defendant and plaintiff, knows Y* for the given judge. However, the defendant and plaintiff are less interested in Y' than they are in where Y' is in relation to Y*. That is, they would like to know the probability that the judge finds a verdict in favor of the plaintiff. To do this, they must form expectations regarding the true value of Y, Y', from the case evidence and circumstances. I assume that the plaintiff and defendant estimate Y' with some error to reflect the uncertainty over the case evidence and the interpretation function, H(X), of the presiding judge.

Following the notation of PK, I denote $\hat{Y'}_p$ to be the plaintiff's estimate of Y' and $\hat{Y'}_d$ to be the defendant's estimate of Y'. Then for a given dispute:

$$\hat{Y'}_p = Y' + \epsilon_p \tag{2.1a}$$

$$\hat{Y'}_d = Y' + \epsilon_d \tag{2.1b}$$

where I assume that ϵ_p and ϵ_d are independent random variables with mean 0 and variance σ^2_ϵ. Thus, each party's estimate of Y' is independent and unbiased.

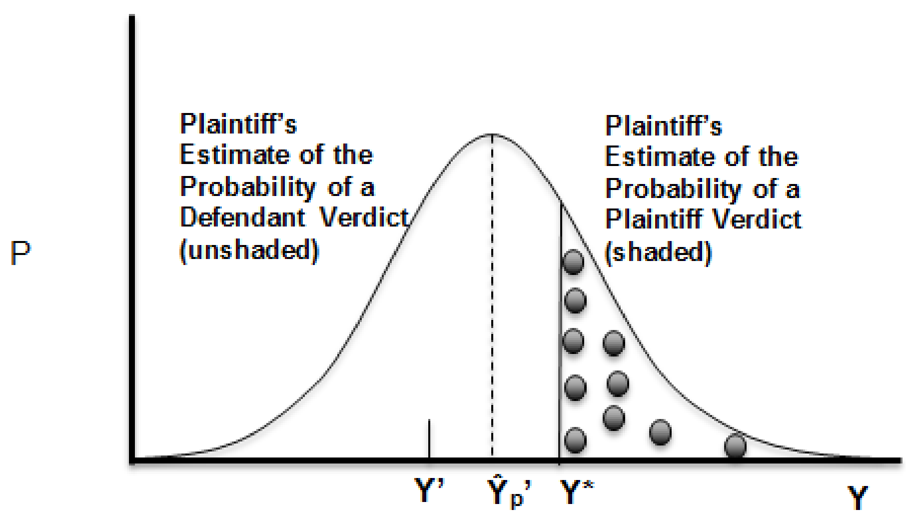


Figure 2.2: Probability Distribution around Plaintiff's Fault Estimate

Figure 2.2 shows an example of the mean value of a plaintiff's estimate of the defendant fault level for a particular dispute, given by $\hat{Y'}_p$. Such an estimate deterimines a distribution of values of the true Y, Y', which will differ from the plaintiff's estimate by the realized value of ϵ_p. The shaded region to the right of Y* indicates the plaintiff's estimated likelihood that Y'>Y*, or equivalently, that the judge would rule in favor of the plaintiff. I denote the magnitude of this region as $\hat{P'}_p$. The complementary area of the distribution for which Y'<Y* corresponds to the plaintiff's estimated likelihood that the defendant would win the trial. An analogous distribution exists for the defendant and equations (2.1a) and (2.1b) require that $E(\hat{Y'}_p) = E(\hat{Y'}_d) = Y'$.

Now assume that I normalize the value of the decision standard so that Y*= 0. Then each party's estimate of the likelihood of a plaintiff verdict given their respective estimates of Y are as follows:

$$\hat{P'}_p = P(Y \geq 0|\hat{Y'}_p) \tag{2.2a}$$

$$\hat{P'}_d = P(Y \geq 0|\hat{Y'}_d) \tag{2.2b}$$

As in PK, I then substitute equations (2.1a) and (2.1b) in for equations (2.2a) and (2.2b), respectively. The result is that each party's estimate of the likelihood of a plaintiff verdict is equivalent to the likelihood that either party's error term is less than their estimate of a plaintiff victory, as follows:

$$\hat{P'}_p = P(\epsilon_p \leq \hat{Y'}_p) \tag{2.3a}$$

$$\hat{P'}_d = P(\epsilon_d \leq \hat{Y'}_d) \tag{2.3b}$$

Since both ϵ_p and ϵ_d have mean 0 and standard error σ_ϵ, each party's estimate can be written as:

$$\hat{P}_p = F(\hat{Y'}_p) \tag{2.4a}$$

$$\hat{P}_d = F(\hat{Y'}_d) \tag{2.4b}$$

where $F(\hat{Y'}_p)$ and $F(\hat{Y'}_d)$ are cumulative distribution functions for ϵ_p and ϵ_d, respectively. Figure 2.3 then represents P'_p as the region to the left of $\hat{Y'}_p$ in the distribution of the plaintiff's error term. An analogous relationship also holds for the defendant.

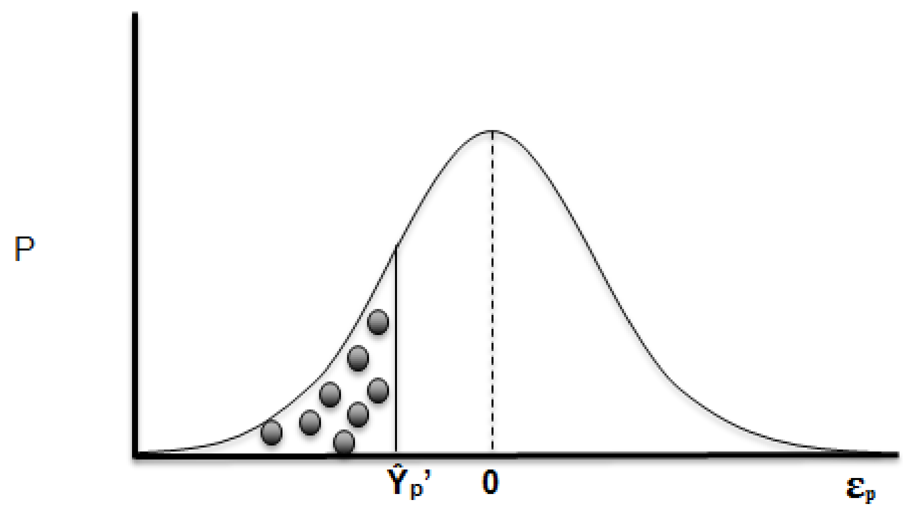


Figure 2.3: Distribution of Plaintiff's Error Term

2.3.3 The Litigation Selection Process

Next, I model the parties' decision to settle or proceed to litigation. Following the notation of PK, let A represent the plaintiff's mimimum settlement demand and B the defendant's maximum settlement offer. Then define (C_p, S_p) and (C_d, S_d) as the litigation and settlement costs to the plaintiff and defendant, respectively. Lastly, I define J as the expected judgement should a liability verdict be upheld.[7] Then I can represent A and B as follows:[8]

$$A = P_p(J) - C_p + S_p \tag{2.5a}$$

$$B = P_d(J) + C_d - S_d \tag{2.5b}$$

When the plaintiff's minimum demand (A) exceeds the defendant's maximum offer (B), the case will proceed to litigation. By rearranging equations (2.5a) and (2.5b), we can see that this condition holds whenever:

$$P_p - P_d > \frac{C - S}{J} \tag{2.6}$$

where $C = C_p + C_d$ and $S = S_p + S_d$. PK further assume that litigation costs exceed settlement costs, i.e. $\frac{C-S}{J} > 0$. Also, $\frac{C-S}{J} < 1$ as otherwise, litigation would never occur because $P_p - Pd \ngtr 1$. I set $D = \frac{C-S}{J}$ so that $0 < D < 1$.

2.3.4 The Role of Uncertainty over Outcomes in Litigation Selection

Equations (2.4a) and (2.4b) imply that $P_p - P_d = F(\hat{Y'}_p) - F(\hat{Y'}_d)$, where the F's are the CDFs of the plaintiff and defendant error terms. This equation shows that the difference in the expected outcome depends upon the difference in their estimates of the level of defendant fault, Y.

[7]The use of J for both plaintiffs and defendants presumes that the stakes are symmetric across both parties. See Priest and Klein (1984) for a framework for analyzing settlement and litigation decisions when stakes are not symmetric.

[8]Since A is the settlement demand that makes the plaintiff indifferent between settlement and litigation, it must be the case that $A - S_p = P_p(J) - C_p$. B is similarly the settlement offer for which the defendant is is indifferent between settlement and litigation. Thus, $-B - S_d = -P_d(J) - C_d$. These equations can be rearranged to form equations (2.5a) and (2.5b).

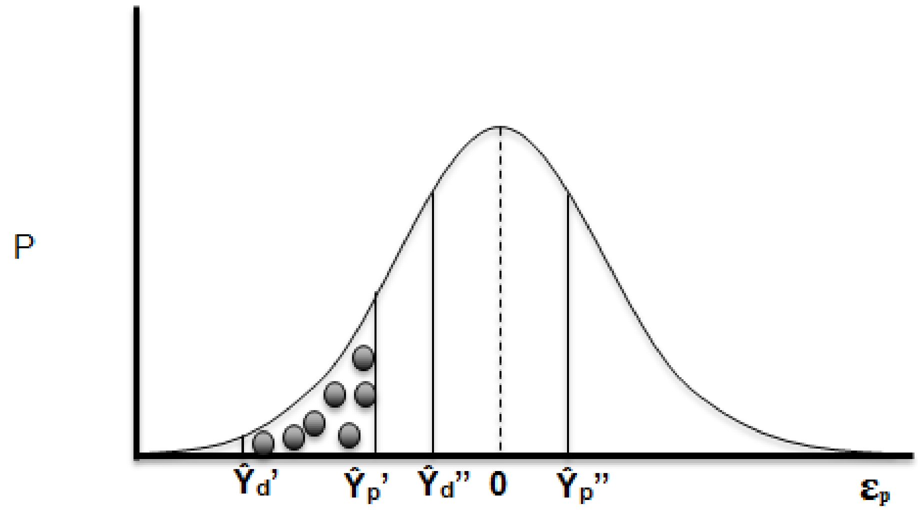


Figure 2.4: Differences between Plaintiff and Defendant Estimate of the Likelihood of a Plaintiff Victory

Let us refer to the difference in expected outcome of the trial, $P_p - P_d$, as the "disagreement area." Figure 2.4 shows the disagreement area graphically for two pairs of potential estimates of defendant fault, $(\hat{Y}'_d, \hat{Y}'_p)$ and $(\hat{Y}''_d, \hat{Y}''_p)$. Importantly, the graph demonstrates that the closer are these estimates to the decision standard, Y*, the greater will be the area of disagreement over expected outcomes (Y* is again normalized to 0). Even though the parties' estimates of the defendant fault, $\hat{Y}'$ and $\hat{Y}''$, differ by the same amount (i.e., $|\hat{Y}'_p - \hat{Y}'_d| = |\hat{Y}''_p - \hat{Y}''_d|$), the disagreement area bounded by $\hat{Y}''_p$ and $\hat{Y}''_d$ (unshaded area) is far greater than the corresponding disagreement area bounded by $\hat{Y}'_p$ and $\hat{Y}'_d$ (shaded area). Thus, the closer are the parties' estimates of the defendant fault level to the judge's decision standard, the greater is the likelihood that the case goes to trial.

Intuitively, a case will go to trial whenever the plaintiff and defendant sufficiently disagree on the expected outcome of the case, i.e. there is a large enough degree of uncertainty over the expected outcome among the two parties. Uncertainty over the outcome, in turn, increases the closer are the estimates of the defendant fault levels to the judge's switching point for ruling in favor of the plaintiff. Therefore, litigated disputes will constitute the set of those for which the true fault level, Y, is sufficiently close to the decision standard.

2.3.5 Applying the Model to Detect Judicial Bias by Judge Sex

Now suppose I would like to test for the existence of judicial bias according to a particular background characteristic of a judge. That is, I aim to detect whether judges with characteristic A have a different decision standard for determining a plaintiff verdict than do judges without characteristic A. Our particular application is to see whether female judges have lower (or higher) decision standards than do male judges in cases of sex discrimination against women. To be able to test this, I first need it to be true that cases are assigned to judges randomly by judge sex. More precisely, random assignment implies $E[\hat{Y'}_p|\text{female judge}] = E[\hat{Y'}_p|\text{male judge}]$ and $E[\hat{Y'}_d|\text{female judge}] = E[\hat{Y'}_d|\text{male judge}]$.[9] This guarantees that, on average, the distance between the plaintiff and defendant's estimates of Y are the same across judges.

If I further observe that the settlement rate is higher when a female judge presides, it must be the case that the estimates of defendant fault lie further away from the decision standard as compared to when a male judge is assigned to the case. Equivalently, one can conclude that female and male judges have different decision standards in sex discrimination cases, i.e. a judicial bias exists. In summary, a simple test for detecting the existence of judicial bias is to compare the settlement rates of randomly assigned cases according to a particular judge characteristic. Where the

[9]Additionally, all of the other relevant parameters of the cases by judge sex will be the same on average. For example, J, the size of the judgement, will be equal in expectation across cases assigned to male and female judges.

settlement rates differ, judges with background characteristic A must use a different decision standard than judges without background characteristic A.

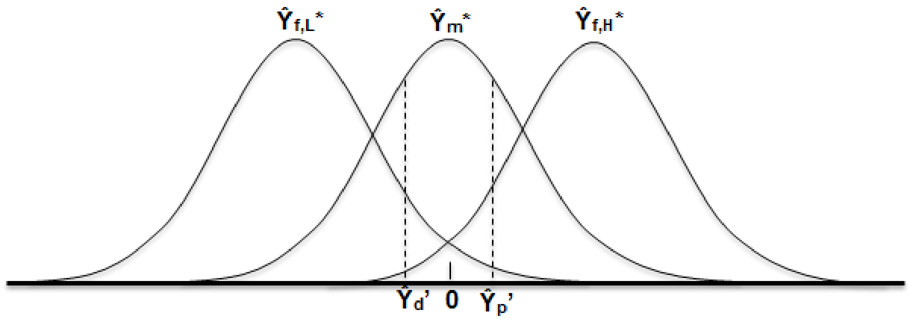


Figure 2.5: Possible Distributions of Error Terms Given Settlement Rate $_{Y_m^*}$ < Settlement Rate $_{Y_f^*}$

Figure 2.5 demonstrates the set of possible distributions of error terms across sex discrimination cases by judge gender. First, random assignment guarantees that the distance between $\hat{Y'}_p$ and $\hat{Y'}_d$ are the same on average across judge gender. The relatively smaller disagreement area for cases overseen by female judges, which is implied by the higher settlement rate, must mean that the decision standard Y_f^* is further away from $\hat{Y'}_p$ and $\hat{Y'}_d$ than is Y_m^*. While intuition and prior literature dictate that $Y_f^* < Y_m^*$ (i.e., male judges require a higher standard of proof for ruling in favor of the plaintiff), I remain agnostic by acknowledging the possibility that Y_f^* may be greater than Y_m^* so long as Y_m^* is closer to the average estimates of defendant fault by the parties in dispute.

2.3.6 Reconciling Prior Estimates of Judicial Bias with the Test

Figure 2.6 shows a distribution of Y along with values of Y_f^* and Y_m^* that is consistent with the observation that settlement rate $_{Y_f^*}$ > settlement rate $_{Y_m^*}$. The

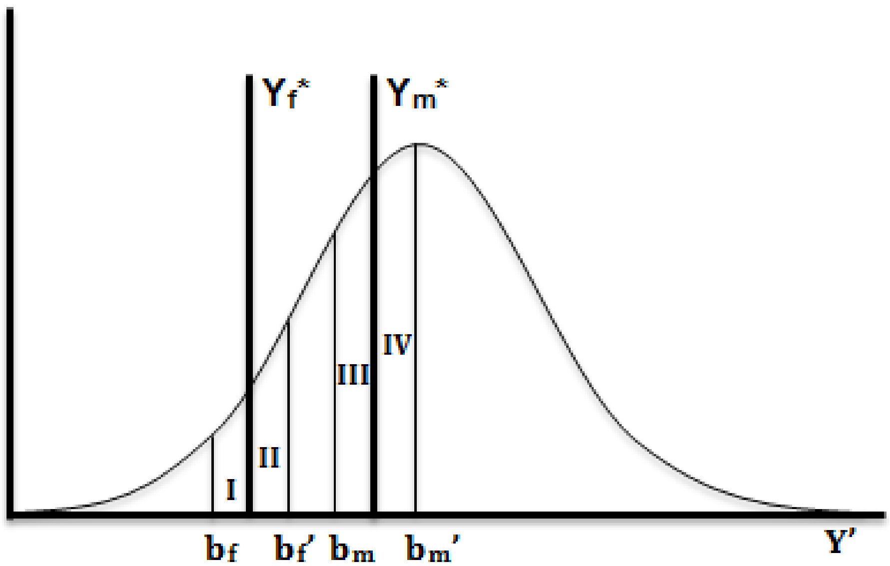


Figure 2.6: Plaintiff Victory Rates and Settlement Rates in Sex Discrimination Cases, by Judge Gender

Previous studies that aim to detect own-gender biases do so by comparing the ratios II/I with IV/III. However, settlement rates, determined by Y_f^* and Y_m^*, directly measure the difference in decision standards used by male and female judges.

set of cases bounded by b_f and b'_f correspond to those cases that would reach the litigation stage conditional on being assigned a female judge. Alternatively, the set of cases bounded by b_m and b'_m correspond to those cases that would reach the litigation stage conditional on being assigned a male judge. Furthermore, areas I and II represent the set of litigated cases overseen by a female judge that would be lost and won by the plaintiff, respectively, whereas areas III and IV represent the set of litigated cases overseen by a male judge that would be lost and won by the plaintiff, respectively.

First note that $I + II < III + IV$, or the number of litigated cases overseen by male judges exceeds the number overseen by female judges, since settlement rate $Y_f^* >$ settlement rate Y_m^*. Moreover, prior studies that aim to detect sex discrimination by judges in sex discrimination cases do so by comparing the ratios II/I with IV/III, or the plaintiff victory rates across female and male judges. While it is true that II/I > IV/III, this only tells part of the story. As is evident from the diagram, I and II correspond to different slivers of the distribution of Y than do III and IV. A more reliable and powerful representation of the bias is indicated by the distance between the decision standards for female and male judges, Y_f^* and Y_m^*, respectively. For example, the difference between the areas of I and II and the areas of III and IV will be relatively small. Thus, given a small sample of cases (and an especially small subsample of litigated cases), it is less likely that one would have sufficient power to statistically distinguish plaintiff victory rates across male and female judges. This partially explains why many prior studies' estimates of this gender bias are either small, imprecise, or null. When one considers first the potential difference in male and female judge decision standards, it becomes clear that prior estimates have a tendency to understate the extent to which male judges are biased against female plaintiffs in cases of workplace sex discrimination.

PK further demonstrate that the plaintiff victory rates tend to converge on 50% as the estimation errors decrease. In other words, as $\sigma_\epsilon \to 0$, II $\to$ I and IV $\to$ III.[10] This fact further undermines the reliability of comparing plaintiff victory rates to detect judicial bias relative to the merits of comparing settlement rates to achieve the same goal. Moreover, even if I had not made the initial assumption that each party knows the decision standard for the assigned judge, disparate settlement rates indicates that this must have been the case. In other words, when settlement rates differ, this implies not only that there is a judicial bias but also that both parties are aware of the existence of such a bias.

[10] In fact Priest and Klein (1984) devote a considerable discussion to showing the conditions under which the 50% plaintiff victory rate rule holds. See Priest and Klein (1984) for a geometric proof of this claim.

2.4 Data and Sample Construction

The data begin with approximately 2,300 federal district court cases of workplace discrimination brought forth by the Equal Employment Opportunity Commission (EEOC) between 1997 and 2006, which constitutes a 33% random sample. These data were collected by Professors Pauline Kim, Margo Schlanger, and Andrew Martin as part of the EEOC Litigation Project, whose goal was to study the litigation activities of the EEOC.[11] (Their subsequent work can be found in Kim (2009) and Schlanger and Kim (2014).)

Importantly, the EEOC only pursues litigation on roughly 0.4% of all discrimination claims filed (or 330 cases per year over 90,000 or so charges filed).[12] This means that even though the sample of litigated cases here constitute a random sample of all litigated cases, they are not random in the sense that the EEOC had selectively deemed these cases as having had enough merit to make the pursuit of litigation worthwhile. That said the data avoid many of the selection problems plaguing other legal studies by including all court decisions (published and unpublished) and all kinds of outcomes (pretrial ajudication, settlements, and post-trial judgments). These features make the data a suitable candidate for uncovering causal effects of different judge characteristics on case outcomes.

The aforementioned researchers began with a list of basic case information, such as the name of the defendant, case basis and issues, case outcome, and date filed, obtained from the EEOC. They then used each case's corresponding docket number in conjunction with an online resource for researching court cases, called "Public Access to Court Electronic Records" (PACER), in order to extract more detailed case information such as all events and motions ordered along with the demographic

[11]The master database is housed at http://eeoclitigation.wustl.edu/.

[12]source: http://www.eeoc.gov/eeoc/statistics/enforcement/index.cfm

characteristics of the judges presiding over the litigation.[13]

Since my focus was on detecting the presence of judicial bias, I isolated two types of cases that I thought would be most subject to such forces: sex discrimination and racial discrimination cases. While racial discrimination cases are a worthwhile subject of study, the sample included only 230 racial discrimination cases filed by African-American plaintiffs, 31 of which were assigned to African-American judges. Thus, there is not a sufficient amount of power to test whether the assignment of African-American judges to racial discrimination cases involving African-American plaintiffs has led to different outcomes than when white judges were assigned.

There were, however, roughly 800 cases of workplace sex discrimination against females, which constitute 37% of all federal court cases from the initial random sample and are the focus of this study. Prior studies point to the fact that female judges, who are themselves the minority in a male-dominated profession, may have had similar experiences to alleged victims of sex discrimination, which may lead them to be able to better detect evidence of this unlawful behavior than their male counterparts (Chew (2010)). Of these 800 cases, 80% were overseen by male judges, who also make up 80% of the total population of judges as would be expected given random assignment of judges to cases.[14] Many prior studies were inhibited by power issues in cases of sex discrimination given the use of historical data coupled with the relatively recent trend of greater appointments of female federal judges. The relatively small number of cases collected, while still more than many past studies of sex discrimination litigation, work against our being able to detect a precise estimate of the extent to which a judicial bias may exist.

[13]When necessary, judges' biographical information was supplemented with the records from online judge databases, such as http://www.fjc.gov/public/home.nsf/hisj, which maintains a directory of all federal judges appointed from 1789 through the present day.

[14]In the following section, I devote a considerable discussion and empirical examination of the extent to which cases are randomly assigned by judge gender.

2.5 Identification and Estimation

Absent a mechanism that ensures random assignment, certain judges may receive cases where the plaintiff is disproportionately likely to win. For example, to the extent that claimants would be able to predict which judge will be assigned to a case, they may be able to 'judge shop' by filing a claim when a judge more favorable to the plaintiff's cause is more likely to be selected. This would introduce selection bias into our estimates, thereby invalidating our efforts to identify the causal effect of being assigned a female judge on case settlement rates. If instead judges are randomly assigned to cases, then the observed difference in case outcomes by judge characteristics is the causal effect of those judge characteristics on outcomes (Rubin (1974); Rubin (1977); Holland (1986)).

Fortunately, federal district court judges are indeed assigned randomly to cases, as is documented in Section 2.2.1. Nonetheless, I perform a series of randomization tests to see whether the data are consistent with randomization by judge sex. First, I perform a χ^2 test to see whether a judge's assignment to a sex discrimination case is independent of the judge's sex. Moreover, if judges are truly randomly assigned to cases, one should expect to see that the case characteristics are balanced on average across the sex of the judge. Accordingly, I test for balance on case covariates across female and male judges as well.

2.5.1 Judge Randomization Tests

In Table 2.1, I report the results of the χ^2 test for whether judge sex is independent of the type of discrimination case assigned. Consistent with randomization, the fraction of sex discrimination cases assigned is nearly identical by judge sex, as it is equal to 36% for female judges and 37% for male judges.

In Table 3.2, I further report the means and mean differences for case outcomes, judge characteristics, and case characteristics. If judges are randomly assigned to cases, one should hope to see that case characteristics balance on average across judge sex. The absence of a difference by judge sex in the basis for the case, issues raised in the case, and location of where the employer was located provides strong evidence for random assignment. Female justices are 8 percentage points more likely to receive a class action suit at the 10% level. However, even if judges are truly randomly assigned to cases, one should expect to see one significant difference in case type covariates at the 10% level for every ten that I test. Here, I test 15 variables and observe one significant difference and so this should not raise any red flags.

2.5.2 Estimation Strategy

Because judges are randomly assigned, I could simply compare mean settlement rates by judge sex. Controlling for additional case covariates should do nothing to change the estimate of the effect of female judge assignment on settlement rates, though doing so may potentially improve the precision of the estimate. However, because judge characteristics are not randomly assigned conditional on judge sex, adding other judge-specific controls should allow me to test whether it is in fact the "sex" part of the judge characteristic bundle that is driving the difference in outcomes. Our most comprehensive regression specification will be of the form:

$$\begin{aligned} 1(Case\ settled)_{ijt} = \beta_0 &+ \beta_1(judge\ sex)_{ij} + \beta_2(judge\ age)_{jt} \\ &+ \beta_3(judge\ experience)_{jt} + \beta_4(judge\ race)_j \\ &+ \beta_5(Democratic\ appointee)_j \\ &+ \beta_6(number\ of\ complainants)_i + year_t + state_i + \epsilon_{ijt} \end{aligned} \tag{2.7}$$

where i, j, t indexes the case number, judge assigned, and year in which the discrimination claim was filed. I cluster the standard errors at the judge level since the 'treatment' is judge sex. I further control for age, experience, and political party. Especially since female appointees are more likely to have been appointed by recent presidents and by Democratic ones and consequently have fewer years of experience, I would like to make sure that judge sex is not simply proxying for other predictors of elevated settlement rates. These features are apparent in Table 3.2, which shows that the female justices in the database are 6.5 years younger, 28 percentage points more likely to have been appointed by Democratic presidents, and have 2.5 fewer years of experience than male judges. I also control for the number of complainants in the case since the greater are the number of plaintiffs, the less likely it is that the defendant will be found without fault. Thus, we should expect that the settlement rate increases with the number of complainants since uncertainty over the potential trial outcome diminishes as well. To capture the possibility that the tolerance threshold for what constitutes workplace sex discrimination has decreased over time (i.e., decision standards have become lower), I also include year fixed effects. Additionally, I allow for these attitudes to vary by location by including state fixed effects.

Still, we should expect to see that the estimated effect of female judge assignment on settlement rates are stable as I control for case covariates. Large fluctuations in our estimate would potentially cast doubt on the plausibility of random assignment. The addition of judge covariates allows us to distinguish the extent to which differences in outcomes are caused by the judge's sex independently of other judge characteristics that are correlated with sex. As a subsequent robustness check, I run both a probit and random effects model to see whether the estimate of the coefficient of interest, β_1, is invariant to the particular specification used.

In order to sign the direction of the potential judicial bias, I further test whether a female plaintiff is more likely to win compensation, in either the pre-trial or litigation stage, when a female justice is assigned to the case. I use the same regression framework as in Equation 2.7, only replacing dependent variable $1(Case\ settled)$ with $1(Plaintiff\ compensated)$. Again, random assignment of judges to cases ensures that the distribution of defendant fault levels, Y', faced by male and female justices is the same, on average. Therefore, where the distribution of case outcomes differ, one can causally attribute this difference to the set of characteristics that belong to the presiding judge. Testing for differences in the rates with which plaintiffs receive judgment payments can provide additional evidence on the direction of a potential judicial bias. In particular, if I find that female plaintiffs filing sex discrimination suits are more likely to win compensation when a female judge is assigned to a case, it must mean that female justices are more favorable to female plaintiffs than are male justices.

2.6 Results

Table 2.2 shows the distribution of final case resolutions according to whether a case was settled, litigated, or reached pre-trial ajudication. I exclude cases reaching pre-trial ajudication ($< 8\%$ of the sample) from the regression analysis of settlement rates since they are at odds with the uncertain versus certain dichotomy of litigated versus settled cases.

Summary judgments, for example, occur when a filing party 'attempts to convince the court that it is entitled to judgment as a matter of law because after looking at all the evidence in the case, there is no dispute as to the facts.'[15] These cases are not settlements in the sense that the parties do not come to a mutual agreement that involves the defendant paying the plaintiff a specified amount. Rather, if the motion for summary judgment is granted, the court decides to award the plaintiff a judgment (a plaintiff summary judgment) or dismisses the case (a defendant summary judgment).

[15]source: http://www.fjc.gov/federal/courts.nsf

Accordingly, cases that are resolved by a summary judgment should be considered especially weak or especially strong cases so that coding them as a non-settlement would belie the extent to which there is uncertainty over the outcome. The next most common type of pre-trial ajudicated outcome reached, the default judgment, occurs whenever a party fails to defend a claim that has been brought forth by another party. Again, these cases are likely to fall at either one of the extremes in terms of the defendant's estimated fault level, Y', by both parties.

Table 3.2 shows that workplace sex discrimination cases filed by female plaintiffs are 7.4% more likely to settle and 5.6% more likely to result in a judgment being awarded to the plaintiff whenever a female judge is assigned to the case. Trial victories are also 17% more common among cases overseen by female judges, though the difference is insignificant due to the fact that less than 5% of all cases that reach the litigation stage. Appeals are approximately 4% more common among male judges, which is to be expected given that the frequency of plaintiff non-compensation is higher when a male judge is assigned and that an appeal will only be filed when the plaintiff does not receive compensation.

2.6.1 Regression Analysis: Robustness and Placebo Tests

Given the assumption of random assignment, we would expect the mean differences in expected case outcomes to be preserved as I control for additional case characteristics. Furthermore, if the judge's sex is the true driver of differences in outcomes, then the addition of other judge controls should also not diminish the explanatory power of the judge's sex. In Table 2.4, I present the main results for the effect of male judge assignment on settlement rates in workplace sex discrimination cases. The estimated effect size is highly stable, ranging from -8 to -10% according to the comprehensiveness of the controls and the model used (OLS, Random Effects, or

Probit). Columns (4) and (5) show that the estimates do become slightly less precise when state and year fixed effects are added, although the effect size is still stable and significant at the 1% level. That said, adding state and year fixed effects reduces the degrees of freedom by 60, which is a non-trivial reduction when the number of observations is slightly above 630. Nonetheless, we should expect that the year the claim was filed has explanatory power in predicting the settlement rates if the decision standards for what constitutes workplace sex discrimination have been evolving over time.

Lastly, I report the results of a placebo test for whether male judge assignment is more likely to lead to settlement in non-sex based discrimination cases. A rejection of the placebo test hypothesis would indicate that male judges happen to have different decision standards on all matters rather than just those involving alleged sex discrimination against women. However, I find that male judge assignment is no more or less likely to result in a settlement in non-sex discrimination cases, which further validates our findings that male judges possess a relative pro-defendant bias in sex discrimination cases filed by female workers.

Consistent with our earlier prediction, the greater are the number of complainants, the more likely it is that a case will settle. If, for example, the parties calculate the probability that the defendant fault level exceeds the decision standard for at least one of the complainants, the likelihood of a defendant fault verdict should increase monotonically as the number of complainants increase.

In Table 2.5, I present the main results for the effect of male judge assignment on the rate with which a plaintiff receives compensation in workplace sex discrimination cases. Here, the results are substantively similar to those shown in the settlement regressions although they are slightly less precise. I find that male judge assignment reduces the likelihood that a female plaintiff receives compensation by 6-7.5% depending on the specification used. These results provide further evidence which suggests that not only do male judges possess a judicial bias but that the bias is against women filing sex discrimination claims. Combined with the results of Columns (1) - (5), the placebo test, again reported in Column (6), show that male judges are no more or less likely to compensate plaintiffs except when the plaintiff is a female filing a sex discrimination claim against her employer.

2.7 Conclusion

Although there is a large literature that explores the existence of judicial gender biases, these studies focus primarily on the approximately 5% of civil cases that reach the litigation stage of dispute resolution. In this paper, I extend the dispute selection model of Priest and Klein (1984) to show that one can more reliably learn about judicial gender biases through the remaining 95% of non-ajudicated cases. In particular, I demonstrate that so long as judges are randomly assigned to district court cases, the finding that settlement rates differ by the sex of the judge necessarily implies that female and male judges have different decision standards. In the context of the study, I find that workplace sex discrimination cases assigned to male judges are 8-9% less likely to settle and 6-7% less likely to result in the plaintiff being compensated. These results imply that, relative to female justices, male justices are less inclined to rule in favor of females who allege that they are victims of workplace sex discrimination. These findings are troubling because judgments paid by defendants are not just punitive but may also deter the employer from engaging in future unlawful discriminatory behaviors. To the extent that male justices are more prone to letting a guilty defendant prevail, societal costs associated with workplace discrimination increase in the present and future. The fact that the number of women justice appointees have been on the rise implies that the magnitude of these costs should decrease over time but proactive steps, such as increasing the frequency with which female justices are assigned to sex discrimination cases, may help further blunt the impact of judicial gender biases on case outcomes.

While a handful of prior studies similarly aim to detect judicial gender biases in sex discrimination cases, they fail to do so for three main reasons. First, much of these analyses are based on published opinions for which judges are less likely to exercise personal discretion due to the spectre of public scrutiny. Furthermore, female justice appointments were relatively rare prior to the last 15 years and so historical studies lack sufficient power to detect significant differences. A third often overlooked reason, which is the primary subject of this paper, is that prior researchers have focused almost exclusively on plaintiff victory rates in litigated cases. However, the set of litigated cases is quite small since a case will only be litigated whenever the plaintiff and defendant disagree enough over the expected outcome of a potential trial. Beyond the problem of small samples, this approach misses the fact that whenever female and male judges have different decision standards, the set of litigated cases that they face will differ in terms of the average defendant fault level. Thus, the set of sex discrimination cases seen by female and male judges may well have statistically identical plaintiff victory rates but this does not in itself guarantee that no judicial bias exists. In fact, I find that male judges disfavor female victims of sex discrimination relative to female judges in spite of the fact that plaintiff litigation success rates are indistinguishable across judge sex.

In future work, I hope to acquire access to the full sample of EEOC cases so as to test for the existence of judicial racial biases as well and to increase the precision with which I can estimate judicial biases overall. Additionally, I hope to obtain access to the full universe of discrimination charges filed with the EEOC in order to study the determinants of workplace discrimination and to test whether plaintiff victories have a deterrence effect that encourages employers to change their culture of workplace discrimination.

2.8 Tables

Table 2.1: Assignment of Cases of Sex Discrimination Against Women, by Sex of Judge

	Female Judge	Male Judge	χ^2 test of independence
Fraction of Sex Discrimination Cases	0.36	0.37	p=0.72
Observations	456	1855	n=2311

Final column tests: H_0 Sex discrimination cases assigned are independent of judge's sex. Sample is a random draw of all EEOC federal litigation cases from 1996 through 2006.

Table 2.2: Classification by Final Resolution Type, Sex Discrimination Cases only

Final Resolution	Settled	Trial Victory	Trial Defeat	Pre-trial Ajudication	In Sample
Consent Judgment	651	0	0	0	651
Voluntary Dismissal- Settlement	63	0	0	0	63
Voluntary Dismissal- Non-Settlement	0	0	0	16	16
Involuntary Dismissal	0	0	0	2	2
Default Judgment	0	0	0	20	20
Plaintiff's Summary Judgment	0	0	0	2	2
Defendant's Summary Judgment	0	0	0	22	22
Defendant's Judgment^{+}	0	0	0	1	1
Plaintiff's Jury Verdict	0	23	0	0	23
Defendant's Jury Verdict	0	0	13	0	13
Plaintiff's Bench Verdict	0	4	0	0	4
No Resolution$^{\phi}$	0	0	0	0	0
Observations	714	27	13	63	817

$^{+}$ Official classification is "Defendant's Judgment as a Matter of Law." $^{\phi}$ 43 cases without a resolution as of 04/22/08 were excluded from the study sample.

Table 2.3: Summary Statistics for Sex Discrimination Cases Against Women

	Female Judge		Male Judge		Difference
	Mean	SD	Mean	SD	
Case Outcomes					
Settled	0.907	0.291	0.833	0.373	0.074***
Plaintiff Compensated	0.937	0.244	0.881	0.323	0.056**
Litigated	0.030	0.171	0.051	0.219	- 0.021
Trial Won $^{\phi}$	0.800	0.447	0.629	0.490	0.171
Amount Awarded	578,224	4,078,038	374,112	3,003,484	204,112
Appeal Filed	0.025	0.158	0.063	0.242	-0.037*
Judge Characteristics					
Age	54.089	6.643	60.575	8.824	-6.486***
Years of experience	8.446	5.886	11.138	7.762	-2.692***
White	0.815	0.389	0.863	0.344	-0.048
Black	0.108	0.312	0.080	0.271	0.029
Hispanic	0.064	0.245	0.051	0.220	0.013
Other race	0.013	0.113	0.007	0.082	0.006
Democratic Appointee	0.648	0.480	0.367	0.483	0.280***
Case Characteristics					
>1 Complainants	0.535	0.500	0.467	0.499	0.068
# Benefited Persons	12.58	85.89	16.16	139.24	-3.578
Class action	0.650	0.479	0.567	0.496	0.083*
Basis					
Equal-pay	0.070	0.256	0.061	0.239	0.009
Pregnancy	0.089	0.286	0.114	0.316	-0.025
Retaliation	0.490	0.502	0.506	0.500	-0.015
Issue					
Harassment	0.737	0.442	0.693	0.461	0.043
Hiring	0.086	0.281	0.110	0.313	-0.024
Discipline	0.039	0.195	0.059	0.236	-0.020
Discharge/Layoff	0.599	0.492	0.625	0.484	-0.027
Location					
Northeast	0.115	0.320	0.115	0.319	-0.000
Midwest $^{\psi}$	0.229	0.422	0.168	0.374	0.062*
South $^{\psi}$	0.389	0.489	0.462	0.499	-0.074
West	0.261	0.441	0.249	0.433	0.012
Right-to-work state	0.494	0.502	0.494	0.500	-0.000
Observations	157		591		748

***, ** and * denote differences that are statistically significant at 1%, 5% and 10%, respectively. $^{\phi}$ The trial win, or plaintiff victory, rates are conditional upon the case reaching the litigation stage. 5 cases and 35 cases reached the litigation stage when the case was assigned to a female and male judge, respectively. $^{\psi}$ The female:male ratio of judges are equal conditional on the composition of judge sex within the region. The total fraction of female and male judges residing in Midwest states are .24 and .22, respectively, while the total fraction of female and male judges residing in Southern states are .41 and .44, respectively

Table 2.4: Effect of Male Judge Assignment on Settlement Rates in Female Sex Discrimination Cases

	(1)	(2)	(3)$^{\lambda}$	(4)	(5)	(6)$^{+}$
Dependent Variable = 1(Settlement reached)						
Male judge	-0.0882***	-0.0872***	-0.107***	-0.0818***	-0.0820***	0.00202
	(0.0237)	(0.0271)	(0.231)	(0.0307)	(0.0307)	(0.0370)
Age of judge		0.00445**	0.00410**	0.00479*	0.00472*	0.00168
		(0.00225)	(0.0110)	(0.00194)	(0.00257)	(0.00224)
Experience of judge		-0.00519**	-0.00460**	-0.00547**	-0.00541**	0.000524
		(0.00238)	(0.0116)	(0.00205)	(0.00255)	(0.00257)
White judge		-0.0136	-0.017	-0.00123	-0.00146	-0.0142
		(0.0568)	(0.0457)	(0.0667)	(0.0666)	(0.0581)
Black judge		0.00257	-0.00184	-0.00402	-0.00443	0.0136
		(0.0672)	(0.0638)	(0.0805)	(0.0804)	(0.0681)
Democratic Appointee		0.0222	0.0244	0.0429	0.0428	0.0422
		(0.0258)	(0.0239)	(0.0268)	(0.0268)	(0.0317)
Number of Complainants		0.0101***	0.0152**	0.00805**	0.00799**	0.00286*
		(0.00358)	(0.00646)	(0.00377)	(0.00376)	(0.00146)
Observations	634	634	634	634	634	786
R^2	0.011	0.030	0.050	0.142	0.142	0.135
Model	OLS	OLS	Probit	OLS	RE	OLS
Sex Discrimination Cases	YES	YES	YES	YES	YES	NO
State FE	NO	NO	NO	YES	YES	YES
Year FE	NO	NO	NO	YES	YES	YES

Dependent variable is an indicator for whether the parties reach a settlement. Robust standard errors, clustered at judge level, are in parentheses. Sample includes a random drawing of EEOC litigated discrimination cases spanning the period from 1996-2006 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ $^{\lambda}$ Marginal Effects reported $^{+}$ Column (6) shows the effect of being assigned a male judge on plaintiff compensation rates in all non-sex-based discrimination cases.

Table 2.5: Effect of Male Judge Assignment on Plaintiff Win Rates in Female Sex Discrimination Cases

Dependent Variable = 1(Plaintiff compensated)	(1)	(2)	(3)$^{\lambda}$	(4)	(5)	(6)$^{+}$
Male judge	-0.0730**	-0.0747**	-0.0877**	-0.0618*	-0.0618*	0.0176
	(0.0284)	(0.0309)	(0.0391)	(0.0324)	(0.0324)	(0.0369)
Age of judge		0.00312	0.00297	0.00201	0.00201	-0.00252
		(0.00199)	(0.0205)	(0.00212)	(0.00212)	(0.00238)
Experience of judge		-0.00378	-0.00344	-0.00322	-0.00322	0.000339
		(0.00253)	(0.00233)	(0.00265)	(0.00265)	(0.00269)
White judge		0.0352	0.0323	0.104*	0.104*	0.0515
		(0.0553)	(0.0567)	(0.0572)	(0.0572)	(0.0661)
Black judge		0.0924	0.0722	0.156**	0.156**	0.111
		(0.0653)	(0.0389)	(0.0659)	(0.0659)	(0.0763)
Democratic Appointee		0.00482	0.00675	0.0276	0.0276	0.0484
		(0.0256)	(0.0266)	(0.0244)	(0.0244)	(0.0325)
Number of Complainants		0.00624	0.00773	0.00312	0.00312	0.00277*
		(0.00495)	(0.00643)	(0.00509)	(0.00509)	(0.00159)
Observations	647	647	647	647	637	807
R^2	0.008	0.018	0.027	0.134	0.134	0.0944
Model	OLS	OLS	Probit	OLS	RE	OLS
Sex Discrimination Cases	YES	YES	YES	YES	YES	NO
State FE	NO	NO	NO	YES	YES	YES
Year FE	NO	NO	NO	YES	YES	YES

Dependent variable is an indicator for whether the plaintiff wins compensation (in either the settlement or litigation stage). Robust standard errors, clustered at judge level, are in parentheses. Sample includes a random drawing of EEOC litigated discrimination cases spanning the period from 1996-2006 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ $^{\lambda}$ Marginal Effects reported $^{+}$ Column (6) shows the effect of being assigned a male judge on settlement rates in all non-sex-based discrimination cases.

2.9 Acknowledgements

Chapter 2, in full, is currently being prepared for submission for publication of the material. Knepper, Matt. The dissertation author was the primary investigator and author of this material.

Chapter 3

The Effect of Access to Electronic Health Records on Throughput Efficiency and Imaging Utilization in the Emergency Department

3.1 Introduction

The Health Information Technology for Economic and Clinical Health (HITECH) Act passed in 2009, authorizing the use of $19.2 billion in incentives for providers who adopt Health Information Technology (HIT) that satisfies meaningful use criteria (Blumenthal (2010); Steinbrook (2009)). This act accelerated the deployment of electronic health record (EHR) systems in hospitals and medical practices. It also financed the development of health information exchanges (HIEs) to facilitate the connectivity of individual EHRs.

Considering the high cost of HIT deployment and its impact on practice patterns, there is relatively limited evidence of how this technology impacts patient care, health care costs, and productivity (Dranove et al. (2014)). In an Emergency Department (ED) setting, the efficiency of information transmission should impact the speed with which physicians are able to diagnose and treat patients. Prompt availability of prior diagnoses, current medications, prior treatments, and test results should accordingly improve the efficiency of care. The ability to query prior diagnostic test results should also lead to reductions in repeat testing, including expensive imaging studies (Bailey et al. (2013); Chaudhry et al. (2006); Tzeel et al. (2011); Frisse et al. (2012); Lammers et al. (2014)). However, many studies of EHRs have indicated an association with increased cost and reduced efficiency. A study by Furukawa et al. (2010) revealed that EHR adoption was associated with an increase in costs per discharge of roughly six to ten percent. A study by Agha (2014), which included a panel of 3900 hospitals between 1998 and 2005, found that Health IT increased billable charges by 1.3% without any evidence of cost savings, reductions in patient mortality, adverse drug reactions, or reductions in readmission rates.

Deploying an EHR is expensive and has major impact on established work flow. These factors are of particular importance in an ED. Analyzing the impact of the availability of an EHR on throughput and costs are essential to fostering ongoing adoption and justifying the expenditure of such systems. An analysis of the impact of the availability of a prior EHR on a patient seen in the ED can potentially serve as a surrogate for the impact of a Health Information Exchange (HIE) in a community. Study objectives: The objective of this study was to assess the whether access to the emergency department EHR was associated with reduced throughput and advanced imaging utilization. The city in which the study institution resides is in the process of implementing a community-wide HIE. Considering that an institutional EHR is not dissimilar in concept to a community HIE with respect to promoting rapid and efficient access to important health information, it was anticipated that this study could provide insight into the potential impact of the HIE in the community.

3.2 Institutional Background

Retrospective multicenter study of encounters at two EDs from June 2011 to June 2012. One hospital is an urban, academic teaching hospital (level 1 trauma center) with an annual census of approximately 37,000 visits. The other hospital is a suburban community hospital with an annual census of approximately 23,000 visits. This period was selected because it coincided with the period of interaction between the specialized ED EHR (WEBCHARTS) and the Health System EHR (EPIC). Study approval was obtained through the Human Research Protections Program (HRPP) at the affiliated university.

The Health System under study has had a robust EHR in place for over five years and the EDs had initially utilized a specialized EHR that connected to the institutional EHR. Prior study of the ED system supported improved operational efficiency of the electronic order entry component of the EHR but no objective analysis has been conducted to assess the utility of the system on other parameters of performance.

3.3 Theory

Considering how many resources have been allocated to the adoption of health information technology, it is surprising that there is still little understanding of the mechanisms by which HIT reduces costs and improves productivity (Dranove et al. (2014)). In an Emergency Department (ED) settingwhere the speed of information transmission is directly tied to the speed with which physicians are able to diagnose and treat patientswe hypothesize that the availability of prior test results, ailment histories, and medication histories in the EHR would increase treatment efficiency. Our main treatment efficiency outcome is total patient throughput time. Moreover, the ability to query prior diagnostic test results could lead to reductions in new imaging exams performed at the point of care and so we test whether EHR availability is associated with reduced imaging as well. Both of these outcomes can be converted into cost savings for the ED and have important implications for patient welfare as well. We confine our analysis to the subset of patients arriving at the ED by way of ambulance in order to eliminate the influence of both patient ED site selection and patient arrival time selection on the relationship between EHR, throughput efficiency, and radiology exams performed. Importantly, an EHR will be available at the point of care only if a given patient has had an encounter with the health system being studied through the affiliated outpatient, ambulatory, or inpatient setting within 3 years prior to the beginning of the sample period. This feature leads to an adversely selected treatment population since utilization of health services is more common among chronically ill individuals. Thus, we would expect that all else equal, patients with EHRs are less likely to experience better outcomes in terms of throughput efficiency and imaging utilization.

To address this selection bias, we first estimate the relationship between the availability of EHRs and our outcome measures and then add a thorough set of patient-specific demographic, initial diagnosis, and acuity variables. Because unobservable differences in patient characteristics are likely correlated with observed ones, we examine how the relationship between EHR status and treatment efficiency outcomes vary as we control for a rich set of patient characteristics. If EHRs truly reduce throughput time and imaging utilization, we should observe that the inclusion of such control variables makes the association between EHR (+) and outcomes more negative.

3.4 Data and Sample Selection

Patient enrollment was limited to those arriving by ambulance because these visits tend to be less discretionary, assuring a combination of patients with prior encounters and thereby likely to have some entry into the systems EHR along with those without prior contact at the study institution and no health information in the EHR. We only included patients with disposition admitted or discharged, excluding those that were transferred to another facility in order to avoid patients whose evaluation was not yet complete. Patients with clearly erroneous data entry into the system such as wrong day of arrival, which could result in negative length of stay or exceptionally long lengths of stay at odds with the documentation in the record, were also omitted from the study sample.

We further omit patients without a primary care physician. The majority of these patients is uninsured and a substantial fraction is homeless. Thus, these patients have limited access to preventative care and disproportionately suffer from chronic and otherwise difficult-to-treat illnesses.

When all of the aforementioned patient groups are omitted, 9,970 patient encounters remain, 6,734 (68.5%) of which are in the treated [or EHR(+)] group to go along with 3,236 (31.5%) control [or EHR(-)] patients.

3.5 Methods

The primary objective of this study is to compare EHR (+) patients to EHR (-) patients in order to determine the impact of EHR access on ED patient care. Data collected includes population demographics of age and gender, time of arrival and time the patient left the ED, all imaging studies obtained, chief complaint, and discharge diagnosis (ICD9 codes). The primary outcomes variables were throughput time (arrival to left the ED), advanced imaging studies obtained (CT, MRI, ultrasound), and charges for imaging studies. Patients were categorized as EHR (+) based upon medical record identifier number and evidence of prior visits to either the study institution or a health provider affiliated with the study institution in the preceding three years. All others were categorized as EHR (-). Our empirical strategy is to estimate the relationship between outcomes and EHR availability in an ordinary least squares (OLS) regression while controlling for a comprehensive set of patient and encounter-specific factors.

Our basic strategy is to estimate the association between EHR(+) and length of stay (LOS) in an ordinary least squares (OLS) regression, while controlling flexibly for other patient and encounter-specific covariates that independently affect the outcome of interest. We use an identical strategy to separately estimate the association between EHR(+) and the use of specific advanced imaging exams, such as CT scans, MRIs, ultrasounds, and all advanced imaging exams.

Our covariates include ones commonly used in the existing ED LOS literature, such as age, sex, insurance provider type, triage acuity, patient disposition, and patient ICD9 codes. Although we control for gender and age, we unfortunately lack data on patient race or ethnicity. However, we proxy for differences in patients socioeconomic statuses by including zip code dummies for each patients primary care provider (This requires the assumption that the patients tend to visit a primary care provider that is close to where they live).

Additionally, we are able to control for more precise sources of variation in the outcome measures, such as unique identifiers for the attending physician, the time of day during which the visit took place, and the patient caseload upon arrival. Physician identifiers, for example, help reduce biases that may result from unobserved idiosyncratic differences in treatment practices that can influence treatment times and the level of testing.

In accordance with current best practice, we cluster our OLS standard errors at the day-site level to allow for arbitrary correlation in unobserved staffing patterns across each ED site in the sample.

We also perform a series of sensitivity checks. We alter both the way in which diagnostic codes are grouped and the study sample used and then re-estimate the relationship between EHR (+) and our outcome measures. First, we re-estimate the relationship between EHR (+) and throughput time (imaging exams) after having replaced ICD9 codes with a more coarse diagnosis classification system. This alternative diagnosis classification system is the NYU ED algorithm, which maps diagnostic discharge codes (ICD9 codes) associated with each patient visit to seven easily interpretable classifications on the basis of whether care was required within 12 hours, whether the ailment was primary care treatable, and whether the condition was preventable or avoidable (Billings et al. (2000b); Billings et al. (2000a)) The algorithm, which was developed nearly 15 years ago, was created by a panel of primary care and ED physicians who had analyzed the full records of nearly 6,000 patients. Based on the initial ailment, demographic factors, and treatment procedures provided in the record, the algorithm weights the probability that the patient visit falls into one of the following seven categories: injuries, psychiatric, substance abuse, non-emergent, emergent but primary care treatable, ED care needed but preventable or avoidable, and ED care needed and the condition was unavoidable. Figure 3.1 provides both a schematic diagram of how the algorithm operates while Table 3.1 lists seven examples of common discharge diagnoses with the associated probabilities across the seven NYU ED categories.

Next, we replace primary ICD9 codes with secondary ICD9 codes (when available). This modification results in reassigned diagnostic codes for nearly 20% of the study sample but should not alter our estimated relationship if the baseline estimates are valid. Also, one might think that patients who come in for psychiatric or substance abuse issues are quite different from the typical patient and that the availability of an EHR should have less influence over throughput time in these cases. Furthermore, the EHR (-) group does contain a slightly larger fraction of these types of patients. Thus, we might worry that if it takes longer to treat psychiatric and substance abuse patients, we would spuriously attribute part of the throughput reduction to the availability of EHRs. To address this concern, we re-run our estimating equation on the subsample that excludes patient visits that are motivated by mental illness. Lastly, the subset of patients suffering from injuries is expected to be the most random of all types of ED visits. Thus, we assess the relationship between EHR (+) and our outcomes as a further check on the credibility of our main results.

3.6 Results

Table 3.2 reports the means and standard deviations for the full set of observable characteristics across EHR (+) and EHR (-) patients. The EHR (+) patients are more than ten years older. EHR (+) patients are also more likely to be admitted to the hospital than EHR (-) patients. Table 3.3 reports the distribution across the NYU ED categories within the two samples. EHR (-) patients are 3-5% more likely to visit for complications associated with drug or alcohol abuse. While this could be concerning, these cases account for fewer than 9% of all visits among paramedic transport patients and we later show that our estimation results are robust to the exclusion of these patient types. Furthermore, EHR (+) patients are 12-15% less likely to be visiting for an injury, 6% more likely to be visiting for a non-emergent reason, and roughly 6% more likely to be visiting for unavoidable emergencies. The latter group of patients are victims of the most severe types of emergencies, such as strokes and heart attacks, and experience average treatment times significantly above the mean (418 minutes versus 386 minutes) and receive nearly 0.6 more imaging exams than the average patient (2.26 versus 1.66). However, patients with non-emergent conditions are less likely to require intensive care. This disparity in the composition of EHR (+) and EHR (-) patients across these categories is a reflection of their differences in age and presenting conditions. Lastly, the EHR (+) group experiences an average length of stay roughly 11% higher than that of the EHR (-) group (399 minutes versus 358 minutes). On the contrary, the EHR (+) group receives 19% fewer CT scans (.50 scans per person versus .62 scans per person).

3.6.1 Throughput Time

Table 3.4 shows the main results for the association between EHR (+) and throughput time. When we control only for time and location effects, EHR (+) predicts an increase in total throughput time of 50 minutes. However, once we control for patient demographics, other predetermined characteristics, and encounter-specific variables, the association falls to a positive 23 minutes. Lastly, when we include current measures of patient conditions and acuity, the relationship between EHR (+) and throughput time vanishes to a statistical 0% ([-0.7, 6.0]). We interpret this as evidence that EHR (+) patients are more chronically ill than EHR (-) patients since controlling for characteristics that proxy for health status erases the positive relationship between access to EHRs and total throughput time. We then interpret the coefficient on EHR status in the most comprehensive specification0 minutesas an upper bound for the treatment effect.

In Table 3.5, we further assess if our estimate is insensitive to the way in which the diagnostic codes are grouped. When replacing ICD9 codes with NYU categories or replacing primary ICD9 codes with secondary ICD9 codes, the relationship between EHR (+) and throughput time hovered around a statistically insignificant positive 2%. When the subsample was further restricted to injured patients or excluded psychiatric or substance abuse patients from the sample, the association between EHR (+) and throughput time is similarly indistinguishable from 0. These consistently null results corroborate the estimates in the main sample, which suggest that the availability of EHRs has no effect on throughput time for the study sample.

3.6.2 Imaging

Table 3.6 presents analogous results for the association between EHR (+) and the use of advanced imaging exams. The OLS estimates indicate that EHR(+) patients experience a reduction in CT scans of .075 per patient, a decline of 13.9% ([4.9, 23.0], p¡0.01) of the mean. We do not observe any statistically significant relationship between EHR status and ultrasounds or MRIs. However, EHR(+) patients also experience a suggestive decrease in total imaging exams by .076, which suggests that all the imaging exam reductions are driven by reduced use of CT scans. CT scans are run most frequently in the ED. To the extent that bottlenecking occurs, this finding is consistent with the hypothesis that physicians are more likely to ration the use of

CTs according to whether they can access prior test results. Lastly, Table 3.7 provides a sensitivity analysis for the use of CT scans. We find that the estimated reduction in CT scans is similarly invariant to the use of the NYU ED algorithm, secondary ICD9 codes, restrictions to injured patients only, and the exclusion of psychiatric and substance abuse patients from the study sample. The estimated reduction in CT scans for EHR(+) patients is bounded between 8.8% and 18.4% for all of these specifications and samples.

3.7 Discussion and Limitations

These results are consistent with prior evidence that access to patient EHRs through a health information exchange system led to reductions in CT scans but no measurable reduction in throughput time (Frisse et al. (2012); Lammers et al. (2014); Daniel et al. (2010)). Weiner et al. (2003) similarly demonstrates that access to prior test results can reduce future orders of such tests. Our primary contribution is to show that these results are also present when including a thorough set of patient-level controls and alternate groupings of such controls, which reduces concerns that the relationship is purely associational. For example, unlike other studies evaluating the utility of health information technology, we were fortunate enough to have detailed enough data to control for physician-specific effects and the severity of the presenting conditions.

Moreover, our treatment variable does not consider whether a physician actually queries the EHRs but instead only considers whether such a record is available. This is a more conservative approach as compared to many other studies since physicians are more likely to query an EHR whenever they think it would be helpful (Carr et al. (2014)).

As is the case with many studies that are without a source of randomly assigned variation in treatment, we cannot conclude that the relationship between the availability of EHRs and our outcome measures is causal. Though we control for other patient-level covariates and comorbidities that can plausibly influence throughput times and number of imaging exams administered, we are still merely making a cross-sectional comparison. A more convincing study design would be a before-and-after study that introduces EHR and compares change in outcomes for those before and after the change. However, even this approach can be hampered by other changes that may occur in the before and after time frame.

Another limitation of the study is that EHR (+) patients tend to be less healthy than EHR(-) patients: since an EHR could have only been generated provided that the patient had a prior encounter at one of the clinics or EDs affiliated with the study institution, we expect that the EHR (+) group is more chronically ill than is the EHR (-) group. This biases our results toward finding a positive association between EHR availability and throughput time We do partially correct for this confound by controlling for case severity and diagnostic codes but the residual variation in treatment efficiency may not be entirely purged of this bias. Another bias that works in the same direction is that all patients with at least one prior encounter are assigned to the EHR (+) group, even if the depth of information provided in the record is minimal and of little utility. Because these biases work against our finding a reduction in throughput time and imaging exams among EHR (+) patients, we interpret our results as a conservative estimate of the true effect.

Lastly, the findings in this study may be applicable to only this single Health System ED, ambulance patients, and to this particular EHR.

3.8 Conclusion

The main objective in this study is to characterize the relationship between the availability electronic health records and treatment efficiency in the emergency department, as measured by total throughput time and diagnostic imaging exams performed. What makes this difficult is that the availability of EHRs is contingent upon whether a patient has had a prior encounter with the healthcare system being studied within the three years prior to the beginning of the sample. Thus, the treated group contains the subset of patients who may be chronic users of inpatient, outpatient, or ED facilities. Absent a source of identifying variation, we estimate the relationship between our outcome variables and the EHR treatment while flexibly controlling for patient characteristics that are likely correlated with health status. We then perform sensitivity checks on these estimates by altering the composition of patients included in the samples by ailment status and by using alternate classifications for the initial presenting conditions of patients.

Our robust estimates of the effect of availability of EHR on throughput time remained small and positive or null across all specifications. We have also found that the availability of EHRs is associated with a reduction in CT scans by 13.9% [resulting in a charge cost saving of $22.52 per patient encountered], but has no relationship with the number of MRI scans, ultrasounds, or other forms of imaging exams. This finding makes sense in light of the fact that CT scans are among the most common type of advanced imaging technology used in EDs.

Beyond the cost savings associated with reduced utilization of CT scans there is the potential added benefit of reduced exposure to ionizing radiation (Brenner et al. (2003); Brenner and Hall (2007); Tubiana et al. (2008)). Hence, the 13.9% reduction in CT scans highlights a potential long-term health benefit attributable to EHR availability.

The possible time savings and demonstrated imaging exam reductions are just two benefits associated with the availability of Electronic Health Records. An integrated EHR system, like the Health Information Exchange (HIE), will allow for independent providers to exchange patient information at the point of care. While this study provides a glimpse into the potential benefits that can arise from an integrated EHR system, future studies of the impact of a fully integrated system on costs and outcomes are certainly merited as the American Reinvestment and Recovery Acts and Affordable Care Acts continue to provide large sums of money to finance the establishment of such systems.

3.9 Figures

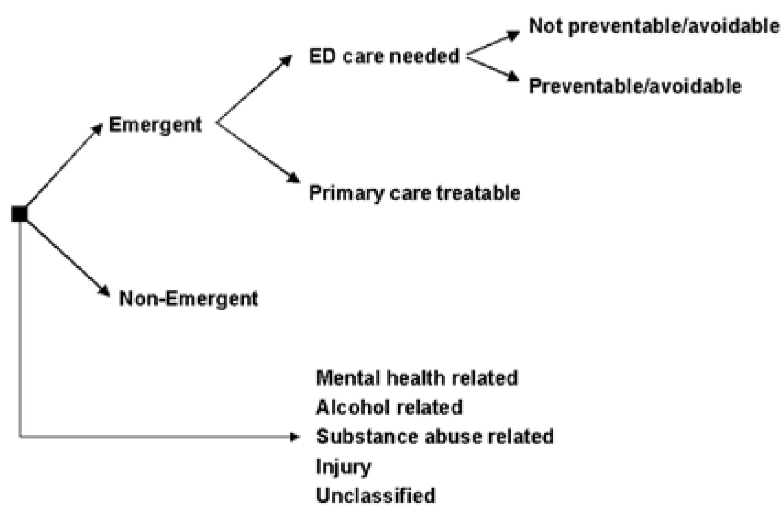


Figure 3.1: NYU ED Algorithm Schematic
The above schematic diagram was reproduced from John Billings NYU Wagner webpage, http://wagner.nyu.edu/faculty/billings/nyued-background.

3.10 Tables

Table 3.1: Presenting Conditions and Associated NYU ED Algorithm Categorization

	Injury	Drugs/alcohol	Mental health	Non-emergent	Emergent, Primary care treatable	ED care needed, preventable	ED care needed, not preventable
Arm laceration	1	0	0	0	0	0	0
ETOH	0	1	0	0	0	0	0
Observed mental condition	0	0	1	0	0	0	0
Cardiac arrest	0	0	0	0	0	0	1
UTI	0	0	0	0.462	0.297	0	0.241
Hypoglycemia	0	0	0	0	0.22	0.78	0
Asthma attack	0	0	0	0	0.02	0.98	0

Table 3.2: Table of Means

	EHR (-)		EHR (+)		Difference
	Mean	SD	Mean	SD	
acuity	1.933	0.497	1.932	0.446	0.001
age	48.81	20.86	57.40	18.23	-8.59***
male	0.538	0.499	0.517	0.500	-0.011*
fraction discharged	0.727	0.446	0.603	0.489	0.124***
total diagnoses	1.296	0.621	1.349	0.663	-0.053***
length of stay (minutes)	358.14	229.20	399.18	225.18	-41.04***
total imaging studies	1.802	2.169	1.650	1.804	0.152***
total CT scans	0.615	1.019	0.498	0.845	0.117***
total ultrasound scans	0.091	0.316	0.085	0.304	0.006
total MRI scans	0.045	0.315	0.048	0.310	-0.003
Observations	3,236		6,734		9,970

***, ** and * denote differences that are statistically significant at 1%, 5% and 10%, respectively.

Table 3.3: Classification by NYU ED Categories

	EHR (-)		EHR (+)		Difference
	Mean	SD	Mean	SD	
Psychiatric	0.041	0.199	0.036	0.186	0.005
Drugs/alcohol	0.077	0.267	0.044	0.206	0.033***
Injury	0.257	0.437	0.102	0.303	0.155***
Non-emergent	0.156	0.298	0.204	0.323	-0.048***
Emergent, primary care treatable	0.155	0.204	0.218	0.218	-0.063***
ED care needed, avoidable	0.043	0.171	0.063	0.205	-0.020***
ED care needed, unavoidable	0.223	0.322	0.277	0.324	-0.054***
Observations	3,236		6,734		9,970

***, ** and * denote differences that are statistically significant at 1%, 5% and 10%, respectively.

Table 3.4: Association between EHR (+) and ED Length of Stay (LOS)

Dependent Variable = LOS (in minutes)	(1)	(2)	(3)
EHR (+)	50.2***	23.3***	10.2
	(5.3)	(6.7)	(6.9)
Observations	9,933	9,933	9,933
R^2	0.044	0.142	0.285
Day, month, hour	YES	YES	YES
Location	YES	YES	YES
Age	NO	YES	YES
Sex	NO	YES	YES
Caseload	NO	YES	YES
Insurance provider	NO	YES	YES
Primary Care zip codes	NO	YES	YES
Attending Physician	NO	YES	YES
ICD9 Codes	NO	NO	YES
Acuity	NO	NO	YES
Procedure	NO	NO	YES
Admission Status	NO	NO	YES

Dependent variable is the length of stay in minutes. Robust standard errors, clustered at the day-site level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Sensitivity Analysis: Association between EHR (+) and ED LOS

Dependent Variable = LOS (in minutes)	(1)	(2)	(3)	(4)
EHR (+)	13.0*	9.7	2.2	10.9
	(6.3)	(6.6)	(16.9)	(6.8)
Observations	9,933	9,933	1,513	9,015
R^2	0.238	0.283	0.479	0.2953
NYU ED Algorithm	YES	NO	NO	NO
Secondary ICD9 codes	NO	YES	NO	NO
Injuries only	NO	NO	YES	NO
Psych./Substance abuse pts.	YES	YES	NO	NO

Dependent variable is the length of stay in minutes. Robust standard errors, clustered at the day-site level, are in parentheses. All models are adjusted for time of visit, location, age, sex, caseload, insurance codes, zip codes of primary care physicians, attending physician, acuity, procedures, and admission status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Association between EHR (+) and Advanced Imaging

Dependent Variable = Number of Exams	Total Imaging	CT scans	Ultrasound	MRI scans
EHR (+)	-0.0764	-0.0747***	-0.0130	-0.0019
	(0.0478)	(0.0248)	(0.0094)	(0.0094)
Observations	9,933	9,933	9,933	9,933
R^2	0.367	0.334	0.204	0.151

Dependent variable is the number of imaging exams run per patient. Robust standard errors, clustered at the day-site level, are in parentheses. All models are adjusted for time of visit, location, age, sex, caseload, insurance codes, zip codes of primary care physicians, attending physician, acuity, procedures, and admission status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Sensitivity Analysis: Association between EHR (+) and CT scans

Dependent Variable = CT scans	(1)	(2)	(3)	(4)
EHR (+)	-0.0986***	-0.0797***	-0.0475	-0.0858***
	(0.0251)	(0.0244)	(0.0864)	(0.0268)
Observations	9,933	9,933	1,513	9,015
R^2	0.211	0.346	0.515	0.358
NYU ED Algorithm	YES	NO	NO	NO
Secondary ICD9 codes	NO	YES	NO	NO
Injuries only	NO	NO	YES	NO
Psych./Substance abuse pts.	YES	YES	NO	NO

Dependent variable is the number of CT scans run per patient. Robust standard errors, clustered at the day-site level, are in parentheses. All models are adjusted for time of visit, location, age, sex, caseload, insurance codes, zip codes of primary care physicians, attending physician, acuity, procedures, and admission status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.11 Acknowledgements

Chapter 3, in full, is currently being prepared for submission for publication of the material. Knepper, Matt; Castillo, Edward; Chan, Theodore; Guss, David. The dissertation author was the primary investigator and author of this material.

Bibliography

Abrams, David, Marianne Bertrand, and Sendhil Mullainathan, "Do Judges Vary in Their Treatment of Race?," *Journal of Legal Studies*, 2012, *41* (2), 347–383.

Agha, Leila, "The effects of health information technology on the costs and quality of medical care," *Journal of health economics*, 2014, *34*, 19–30.

Ashenfelter, Orley and Joel Waldfogel, "Bargaining in the shadow of the judge: Empirical tests," *Mimeographed. Princeton, NJ: Princeton University*, 1995.

_ , Theodore Eisenberg, and Stewart J Schwab, "Politics and the judiciary: The influence of judicial background on case outcomes," *The Journal of Legal Studies*, 1995, pp. 257–281.

Bailey, James E, Jim Y Wan, Lisa M Mabry, Stephen H Landy, Rebecca A Pope, Teresa M Waters, and Mark E Frisse, "Does health information exchange reduce unnecessary neuroimaging and improve quality of headache care in the emergency department?," *Journal of general internal medicine*, 2013, *28* (2), 176–183.

Billings, J, N Parikh, and T Mijanovich, "Issue brief no. 434emergency department use: the New York story," *New York: The Commonwealth Fund, November*, 2000.

Billings, John, Nina Parikh, and Tod Mijanovich, "Emergency department use in New York City: a substitute for primary care?," *Issue brief (Commonwealth Fund)*, 2000, (433), 1–5.

Bird, Susan Willett, "The Assignment of Cases to Federal District Court Judges," *Stanford Law Review*, 1975, *27* (2), pp. 475–487.

Blumenthal, David, "Launching HIteCH," *New England Journal of Medicine*, 2010, *362* (5), 382–385.

Borzekowski, Ron, "Measuring the cost impact of hospital information systems: 1987–1994," *Journal of Health Economics*, 2009, *28* (5), 938–949.

Boyd, Christina L, Lee Epstein, and Andrew D Martin, "Untangling the causal effects of sex on judging," *American journal of political science*, 2010, *54* (2), 389–411.

Braveman, Paula, Susan Egerter, and David R Williams, "The Social Determinants of Health: Coming of Age," *Annual Review of Public Health*, 2011, *32*, 381–398.

Brenner, David J and Eric J Hall, "Computed tomographyan increasing source of radiation exposure," *New England Journal of Medicine*, 2007, *357* (22), 2277–2284.

_ **, Richard Doll, Dudley T Goodhead, Eric J Hall, Charles E Land, John B Little, Jay H Lubin, Dale L Preston, R Julian Preston, and Jerome S Puskin**, "Cancer risks attributable to low doses of ionizing radiation: assessing what we really know," *Proceedings of the National Academy of Sciences*, 2003, *100* (24), 13761–13766.

Buchmueller, Thomas and John Dinardo, "Did Community Rating Induce an Adverse Selection Death Spiral? Evidence from New York, Pennsylvania, and Connecticut," *American Economic Review*, 2002, *92* (1), 280–294.

Buchmueller, Thomas C and Su Liu, "Health Insurance Reform and HMO Penetration in the Small Group Market," *Inquiry: The Journal of Health Care Organization, Provision, and Financing*, 2005, *42* (4), 367–380.

Cameron, Charles M and Craig Cummings, "Diversity and judicial decision making: Evidence from affirmative action cases in the federal courts of appeals, 1971-1999," in "annual meeting of the Midwest Political Science Association" 2003.

Carr, Christine Marie, Charles Samuel Gilman, Diann Marie Krywko, Haley Elizabeth Moore, Brenda J Walker, and Steven Howard Saef, "Observational Study and Estimate of Cost Savings from Use of a Health Information Exchange in an Academic Emergency Department," *The Journal of emergency medicine*, 2014, *46* (2), 250–256.

Chaudhry, Basit, Jerome Wang, Shinyi Wu, Margaret Maglione, Walter Mojica, Elizabeth Roth, Sally C Morton, and Paul G Shekelle, "Systematic review: impact of health information technology on quality, efficiency, and costs of medical care," *Annals of internal medicine*, 2006, *144* (10), 742–752.

Chew, Pat K, "Judges' Gender and Employment Discrimination Cases: Emerging Evidence-Based Empirical Conclusions," *J. Gender Race & Just.*, 2010, *14*, 359.

_ **and Robert E Kelley**, "Myth of the color-blind judge: An empirical analysis of racial harassment cases," *Wash. UL Rev.*, 2008, *86*, 1117.

Clemens, Jeffrey, "Regulatory Redistribution in the Market for Health Insurance," Working Paper 19904, National Bureau of Economic Research February 2014.

Clermont, Kevin M and Stewart J Schwab, "Employment discrimination plaintiffs in federal court: From bad to worse," *Harv. L. & Pol'y Rev.*, 2009, *3*, 103.

Cohen, Alma and Liran Einav, "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, 2007.

Collins, J R, Kenneth L Manning, and Robert A Carp, "Gender, critical mass, and judicial decision making," *Law & Policy*, 2010, *32* (2), 260–281.

Crowe, Nancy E, "The Effects of Judges' Sex and Race on Judicial Decision Making on the US Courts of Appeals, 1981-1996." PhD dissertation, University of Chicago, 1999. Department of Political Science. 1999.

Currie, Janet and Jonathan Gruber, "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *The Quarterly Journal of Economics*, 1996, *111* (2), 431–466.

Cutler, David M, Adriana Lleras-Muney, and Tom Vogl, "Socioeconomic Status and Health: Dimensions and Mechanisms," Technical Report, National Bureau of Economic Research 2008.

Daniel, Gregory W, Edward Ewen, Vincent J Willey, IV Reese, L Charles, Farshad Shirazi, and Daniel C Malone, "Efficiency and Economic Benefits of a Payer-based Electronic Health Record in an Emergency Department," *Academic Emergency Medicine*, 2010, *17* (8), 824–833.

Davis, Sue, Susan Haire, and Donald R Songer, "Voting behavior and gender on the US Courts of Appeals," *Judicature*, 1993, *77*, 129.

Dranove, David, Chris Forman, Avi Goldfarb, and Shane Greenstein, "The Trillion Dollar Conundrum: Complementarities and Health Information Technology," *American Economic Journal: Economic Policy*, 2014, *6* (4), 239–270.

Dube, Arindrajit, T William Lester, and Michael Reich, "Minimum Wage Effects Across State Borders: Estimates using Contiguous Counties," *The Review of Economics and Statistics*, 2010, *92* (4), 945–964.

Eibner, Christine, "Spillover Effects of State Medicaid Expansion Decisions on Private Health Insurance Markets," in "Health & Healthcare in America: From Economics to Policy" ASHEcon 2014.

Elo, Irma T, "Social Class Differentials in Health and Mortality: Patterns and Explanations in Comparative Perspective," *Annual Review of Sociology*, 2009, *35*, 553–572.

Fang, Hanming, Michael P Keane, and Dan Silverman, "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy*, 2008, *116* (2), 303–350.

Farhang, Sean and Gregory Wawro, "Institutional dynamics on the US court of appeals: Minority representation under panel decision making," *Journal of Law, Economics, and Organization*, 2004, *20* (2), 299–330.

Feldman, Roger and Bryan Dowd, "Risk Segmentation: Goal or Problem?," *Journal of Health Economics*, 2000, *19* (4), 499–512.

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

Friedl, John and Andre Honoree, "Is Justice Blind-Examining the Relationship between Presidential Appointments of Judges and outcomes in Employment Discrimination Cases," *Cumb. L. Rev.*, 2007, *38*, 89.

Frisse, Mark E, Kevin B Johnson, Hui Nian, Coda L Davison, Cynthia S Gadd, Kim M Unertl, Pat A Turri, and Qingxia Chen, "The financial impact of health information exchange on emergency department care," *Journal of the American Medical Informatics Association*, 2012, *19* (3), 328–333.

Furukawa, Michael F, TS Raghu, and Benjamin Shao, "Electronic medical records, nurse staffing, and nurse-sensitive patient outcomes: Evidence from California hospitals, 1998–2007," *Health services research*, 2010, *45* (4), 941–962.

Gabel, Jon, "Ten Ways HMOs have Changed During the 1990s," *Health Affairs*, 1997, *16* (3), 134–145.

Gazal-Ayal, Oren and Raanan Sulitzeanu-Kenan, "Let My People Go: Ethnic In-Group Bias in Judicial DecisionsEvidence from a Randomized Natural Experiment," *Journal of Empirical Legal Studies*, 2010, *7* (3), 403–428.

Grassi, Simona, "Optimal Public Rationing: Price Response and Cost Effectiveness," 2009.

Gruber, Jonathan and Robin McKnight, "Controlling Health Care Costs Through Limited Network Insurance Plans: Evidence from Massachusetts State Employees," Technical Report, National Bureau of Economic Research 2014.

_, Leemore Dafny, and Christopher Ody, "More Insurers Lower Premiums: Evidence from Initial Pricing in the Health Insurance Marketplaces," Technical Report, National Bureau of Economic Research 2014.

Hall, MA, "The Impact of Health Insurance Market Reforms on Market Competition.," *The American Journal of Managed Care*, 2000, *6* (1), 57–67.

Hamersma, Sarah and Matthew Kim, "Participation and Crowd Out: Assessing the Effects of Parental Medicaid Expansions," *Journal of Health Economics*, 2013, *32* (1), 160–171.

Herring, Bradley and Mark V Pauly, "The Effect of State Community Rating Regulations on Premiums and Coverage in the Individual Health Insurance Market," Technical Report, National Bureau of Economic Research 2006.

Hillestad, Richard and James Bigelow, "Health information technology: can HIT lower costs and improve quality?," 2005.

Holland, Paul W, "Statistics and causal inference," *Journal of the American statistical Association*, 1986, *81* (396), 945–960.

Holmes, Thomas J, "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 1998, *106* (4), 667–705.

Huang, Rocco R, "Evaluating the Real Effect of Bank Branching Deregulation: Comparing Contiguous Counties Across US State Borders," *Journal of Financial Economics*, 2008, *87* (3), 678–705.

Johar, Meliyanni, Glenn Stewart Jones, and Elizabeth Savage, "Emergency admissions and elective surgery waiting times," *Health economics*, 2013, *22* (6), 749–756.

Jr, Joseph J Doyle, John A Graves, Jonathan Gruber, and Samuel Kleiner, "Do high-cost hospitals deliver better care? evidence from ambulance referral patterns," Technical Report, National Bureau of Economic Research 2012.

Kim, Pauline T, "How Should We Study District Judge Decision-Making?," *Washington University Journal of Law and Policy*, 2009, *29* (83).

Kowalski, Amanda, "The Early Impact of the Affordable Care Act State-by-State," *Brookings Papers on Economic Activity*, 2014.

Kulik, Carol T, Elissa L Perry, and Molly B Pepper, "Here comes the judge: the influence of judge personal characteristics on federal sexual harassment case outcomes.," *Law and Human Behavior*, 2003, *27* (1), 69.

Lammers, Eric J, Julia Adler-Milstein, and Keith E Kocher, "Does health information exchange reduce redundant imaging? Evidence from emergency departments," *Medical care*, 2014, *52* (3), 227–234.

Manning, Kenneth L, "¿ Cómo Decide? Decision-Making by Latino Judges in the Federal Courts," in "annual meeting of the Midwest Political Science Association, Chicago IL" 2004.

Marmot, Michael, "The Influence of Income on Health: Views of an Epidemiologist," *Health affairs*, 2002, *21* (2), 31–46.

Massie, Tajuana, Susan W Johnson, and Sara Margaret Gubala, "The impact of gender and race in the decisions of judges on the United States courts of appeals," in "annual meeting of Midwest Political Science Association, Chicago IL," Vol. 5 2002.

McCormick, Danny, David H Bor, Stephanie Woolhandler, and David U Himmelstein, "Giving office-based physicians electronic access to patients prior imaging and lab results did not deter ordering of tests," *Health Affairs*, 2012, *31* (3), 488–496.

McCullough, Jeffrey S, Stephen Parente, and Robert Town, "Health information technology and patient outcomes: The role of organizational and informational complementarities," Technical Report, National Bureau of Economic Research 2013.

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

Morin, Jason L, "The Voting Behavior of Minority Judges in the US Courts of Appeals: Does the Race of the Claimant Matter?," *American Politics Research*, 2013, p. 1532673X13486925.

Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration," 2015.

Pauly, Mark, Scott Harrington, and Adam Leive, ""Sticker Shock" in Individual Insurance under Health Reform," Working Paper 20223, National Bureau of Economic Research June 2014.

Peresie, Jennifer L, "Female judges matter: Gender and collegial decisionmaking in the federal appellate courts," *Yale Law Journal*, 2005, pp. 1759–1790.

Price, Joseph and Justin Wolfers, "Racial Discrimination Among NBA Referees," *The Quarterly Journal of Economics*, 2010, *125* (4), 1859–1887.

Priest, George L and Benjamin Klein, "The selection of disputes for litigation," *The Journal of Legal Studies*, 1984, pp. 1–55.

Rachlinski, Jeffrey J, Sheri Lynn Johnson, Andrew J Wistrich, and Chris Guthrie, "Does unconscious racial bias affect trial judges?," *notre dame law review*, 2009, *84* (3), 09–11.

Ross, Stephen E, Tiffany A Radcliff, William G LeBlanc, L Miriam Dickinson, Anne M Libby, and Donald E Nease, "Effects of health information exchange adoption on ambulatory testing rates," *Journal of the American Medical Informatics Association*, 2013, *20* (6), 1137–1142.

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

Rubin, Donald B, "Estimating causal effects of treatments in randomized and nonrandomized studies.," *Journal of educational Psychology*, 1974, *66* (5), 688.

_ , "Assignment to Treatment Group on the Basis of a Covariate," *Journal of Educational and Behavioral statistics*, 1977, *2* (1), 1–26.

Sasso, Anthony T Lo and Ithai Z Lurie, "Community Rating and the Market for Private Non-group Health Insurance," *Journal of Public Economics*, 2009, *93* (1), 264–279.

Schlanger, Margo and Pauline Kim, "The Equal Employment Opportunity Commission and Structural Reform of the American Workplace," *Wash. UL Rev.*, 2014, *91*, 1519–1663.

Schwartz, Aaron L and Benjamin D Sommers, "Moving for Medicaid? Recent Eligibility Expansions did not Induce Migration from other States," *Health Affairs*, 2014, *33* (1), 88–94.

Segal, Jennifer A, "Representative decision making on the federal bench: Clinton's district court appointees," *Political Research Quarterly*, 2000, *53* (1), 137–150.

Shayo, Moses and Asaf Zussman, "Judicial Ingroup Bias in the Shadow of Terrorism," *The Quarterly Journal of Economics*, 2011, *126*, 1447–1484.

Shor, Boris, "ACA Implementation Votes in the States," *Available at SSRN 2271783*, 2013.

Sidorov, Jaan, "It aint necessarily so: the electronic health record and the unlikely prospect of reducing health care costs," *Health Affairs*, 2006, *25* (4), 1079–1085.

Songer, Donald R, Sue Davis, and Susan Haire, "A reappraisal of diversification in the federal courts: Gender effects in the courts of appeals," *The Journal of Politics*, 1994, *56* (02), 425–439.

Steinbrook, Robert, "Health care and the American recovery and reinvestment act," *New England Journal of Medicine*, 2009, *360* (11), 1057–1060.

Tian, Liu, "The Effect of Public Health Insurance Expansion on Private Insurance Premiums: Evidence from SCHIP," 2013.

Tubiana, Maurice, S Nagataki, and L Feinendegen, "Computed tomography and radiation exposure," *N Engl J Med*, 2008, *358* (8), 850.

Tzeel, Albert, Victor Lawnicki, and Kim R Pemble, "The business case for payer support of a community-based health information exchange: a Humana pilot evaluating its effectiveness in cost control for plan members seeking emergency department care," *American health & drug benefits*, 2011, *4* (4), 207.

Vaithianathan, Rhema, "A Critique of the Private Health Insurance Regulations," *Australian Economic Review*, 2004, *37* (3), 257–270.

Veiga, André and E Glen Weyl, "Product Design in Selection Markets," *Available at SSRN 1935912*, 2014.

Weiner, Michael, Christopher M Callahan, William M Tierney, J Marc Overhage, Burke Mamlin, Paul R Dexter, and Clement J McDonald, "Using information technology to improve the health care of older adults," *Annals of internal medicine*, 2003, *139* (5_Part_2), 430–436.

Westergren, Sarah, "Gender Effects in the Courts of Appeals Revisited: The Data Since 1994," *Geo. LJ*, 2003, *92*, 689.

Williams, Claudia, Farzad Mostashari, Kory Mertz, Emily Hogin, and Parmeeth Atwal, "From the Office of the National Coordinator: the strategy for advancing the exchange of health information," *Health affairs*, 2012, *31* (3), 527–536.