ADVERSE SELECTION IN ACA EXCHANGE MARKETS: EVIDENCE FROM COLORADO*

Matthew Panhans†

Duke University

February 2017

Abstract

This study tests for adverse selection in the Affordable Care Act (ACA) health insurance exchanges established in 2014, and quantifies the consequences for consumer welfare and market efficiency. Using a new statewide dataset of medical claims from Colorado, I use plausibly exogenous premium variation generated by geographic discontinuities to test for adverse selection. In this context, a positive relationship between premiums and medical spending of the insured population indicates adverse selection, as the lowest cost individuals are the first to drop out of the market in response to rising premiums. I find evidence of adverse selection in the non-group market, where a 1% increase in premiums leads to a 0.8% increase in the average annual medical expenditures of the insured population. I then estimate insurance demand using the same geographic premium variation. The demand and cost estimates are combined in a framework to calculate the welfare loss due to selection, as well as an evaluation of policy interventions. My estimates indicate that providing additional premium subsidies would enhance welfare in this market, and moreover, due to heterogeneity across age groups in both demand and costs, I estimate that age-targeted premium subsidies would be a more cost-effective use of public funds to enhance welfare. These results offer the first quasi-experimental evidence of adverse selection in the new ACA Exchange markets, and conclusions from the policy evaluations have implications for the future effectiveness of this cornerstone of the ACA.

JEL: I11, I13, I28
Keywords: health insurance, selection, healthcare reform

*I thank my dissertation committee Juan Carlos Suárez Serrato, Frank Sloan, Allan Collard-Wexler, and James Roberts for valuable guidance. I thank the Center for Improving Value in Health Care (and especially Jonathan Mathieu, Cari Frank, Tracey Campbell, and Alton Ward) for assistance in providing and interpreting the data.

†*matthew.panhans@duke.edu*. 213 Social Sciences, Box 90097, Department of Economics, Duke University, Durham, NC 27708-0097. Word count: 15,500.

Introduction

The new health insurance marketplaces, or "Exchanges," established as part of the Affordable Care Act (ACA) have seen many insurers exit or reduce their participation. UnitedHealth stated that in 2017 it will quit offering policies in 30 of 34 states in which it was operating, and more recently Aetna and Humana have stated intentions to pull out of most state ACA Exchanges in 2017 as well.[1] At the same time, some insurers that remain in the Exchanges have been steeply increasing premiums each year. Consumers thus face a shrinking and increasingly expensive set of options, and as a result many states have had to continually revise downward their enrollment projections.[2]

Adverse selection may be part of the cause of this underlying instability.[3] As required as part of the ACA, insurers are limited in their ability to adjust premiums for risk.[4] The resulting more uniform pricing is what creates the possibility for adverse selection. That is, if there is a positive relationship between consumers' demand for insurance and individual expected costs, healthier policyholders begin to forgo coverage, leaving insurers with an increasingly unhealthy and expensive pool to cover. Theoretically, this dynamic leads to an under-provision of insurance relative to the efficient level (Rothschild and Stiglitz, 1976). If selection is present, policy responses that increase coverage towards the efficient level, such as premium subsidies or stronger mandates, may improve welfare. However, alternative stories could generate these patterns; for example, the explanation could simply be one of mis-pricing by insurers in the early years of a new market and new population gaining insurance. Knowing whether cost increases are due to adverse selection is important for how to think about consumer welfare and the evaluation of policy interventions in these markets.

This study investigates the extent to which adverse selection is present in the ACA exchanges, and quantifies the consequences for consumer welfare and market efficiency. There are two primary empirical challenges to this goal. First, it has typically been difficult to obtain individual-level data on non-group consumers. I focus on Colorado's Exchange in order to take advantage of a new dataset that provides a broader and more detailed view on costs in the non-group market than has typically been available: Colorado's All-Payer Claims Database (APCD).[5] Second, even with detailed cost data, there remains the empirical challenge of disentangling the effect of adverse selection from moral hazard. Both manifest themselves as a positive relationship between the demand for insurance and an individual's healthcare utilization, but each has very different policy implications. Plausibly exogenous variation in premiums can be useful in isolating the selection

[1]"Obamacare Insurers Are Suffering. That Won't End Well," *BloombergView*, November 19, 2015; "Aetna Joins Other Major Insurers In Pulling Back From Obamacare," *NPR*, 16 August 2016

[2]"Costs, changes led Obamacare enrollment to fall short of earlier estimates," *USA Today*, 16 February 2016.

[3]See Akerlof (1970); Rothschild and Stiglitz (1976) and the large literature that followed on how selection can cause instability. Cutler and Reber (1998) show a stark example when after a pricing reform at Harvard, within 3 years adverse selection eliminated the market for more generous insurance entirely.

[4]Specifically, they are unable to price based on each individual's health status or medical history, as was possible before the passage of the ACA.

[5]The non-group market excludes individuals with public health insurance through Medicaid or Medicare, and individuals in the employer-sponsored (group) market. Thus, this market includes everyone who purchases insurance directly from an insurer; these may be individuals who are self-employed, or working part-time or for an employer who does not offer insurance.

effect, and in this study I exploit geographic discontinuities in insurance premiums as a source of premium variation. Specifically, the discontinuities occur at boundaries of the so-called rating areas that were established in 2014 as part of the ACA. By comparing individuals living in zip codes just on either side of the boundaries but in the same local medical market, I test for adverse selection, indicated by a positive relationship between insurance premiums and the average costs of the insured population. The premium variation isolates the effect of adverse selection independently from moral hazard, offering the first quasi-experimental evidence on selection directly related to the new ACA Exchanges.

Finally, an economic framework is required to go beyond detecting adverse selection and evaluate welfare effects of changes in coverage. I describe a welfare framework that allows for salient features of the non-group market in the context of the ACA, and that can be used to evaluate alternative policies. This framework follows the sufficient statistics framework of Einav et al. (2010) (henceforth EFC), which provides a close link between the welfare costs of adverse selection and reduced-form effects of premium variation on costs and demand. There are two main features of the ACA context that motivate extensions to that framework for the present study. First, the decomposition of selection and moral hazard is typically done in a framework where consumers face only a binary decision (e.g. low or high generosity insurance). In the context of the ACA Exchanges, individuals face more than two options, and selection can occur both on the extensive margin (purchasing insurance or not) and on the intensive margin (choosing the generosity of the plan). The framework I use guides the empirical specifications, and shows how to isolate the effect of selection on the extensive margin in this context. Secondly, I extend the framework to allow for the inclusion of a social value for insurance coverage. Since one of the primary goals of the ACA was to increase the insured rate, this is likely an important consideration in this context. The source of value may come from an intrinsic value on increasing coverage, or from an externality imposed by uninsured individuals who forgo traditional health care until they end up incurring costlier services through an emergency department.

The results indicate robust evidence that consumers do adversely select into the non-group market. I find that a 1% increase in monthly premiums leads to a 0.8% increase in the average annual medical expenditures of the insured population. The selection effect is driven by individuals below age 55, and particularly ages 35-44. Placebo checks on market segments where the boundary should not matter, including the non-group market in 2013 before there were boundaries, as well as the group (employer) and Medicaid markets, provide evidence for the validity of the research design. To provide further evidence of adverse selection, I use the Healthcare Cost and Utilization Project's (HCUP) Chronic Condition Indicator (CCI) tool to measure the prevalence of chronic conditions in the population, which serves as a proxy for the population's underlying risk. I find that an increase in premiums leads to an increase in the prevalence of chronic conditions, as one would expect in an adversely selected insurance market. These results corroborate with survey evidence that the pool of uninsured individuals became healthier from 2013 to 2014 (Jacobs et al., 2016).

I estimate the demand for insurance using the same source of premium variation. Combining the demand and cost estimates allows for a welfare analysis, and I calculate that selection raises average monthly premiums by at least \$47 (from \$151 to \$198) above their level without selection. This corresponds to a lower bound estimate of welfare loss of \$26 per person per month, or \$312 annually. While this estimate is over twice as large a magnitude as found in studies in the employer market, this is only moderately larger than the estimate of \$241 annual per person welfare loss based on Massachusetts 2006 healthcare reform investigated by Hackmann et al. (2015).[6] Consequently, implementing a policy that would increase coverage, such as a stronger individual mandate or premium subsidies that increase coverage, would increase welfare and have a large scope to be welfare improving in this context.

The estimates show a large amount of heterogeneity across age groups, suggesting that age-targeted premium subsidies may be more efficient at increasing coverage and addressing selection. I explore this possibility by finding the optimal subsidy within each age group. The most effective groups to target are those whose costs will fall most quickly for each dollar of subsidy provided, and thus depend on both cost and demand factors. This exercise shows that it is effective to provide additional subsidies to the younger age groups, with the optimal policy entailing additional \$13-15 monthly subsidies to individuals in the 25-34, 35-44 and 45-54 age groups. It is not optimal to provide additional subsidies to the 55-64 age group. These differences are driven first by how quickly costs change. Selection is most acute for the middle and lower age bins, making them the most effective to target to bring in new (healthy) enrollees. On the demand side, there are differences in take-up rates, and with particularly high take-up rates in the 55-64 age group, such that the subsidy must pay a high fraction of infra-marginal consumers. Moreover, the 55-64 age group is the least price sensitive, and so it takes the most amount of additional subsidy to incentivize each additional consumer to take up insurance. In the current legislation of the Affordable Care Act, subsidies are not conditional on age. The optimal age-targeted subsidy policy can be compared to an alternative subsidy policy that spends the same amount of money on premium subsidies, but without conditioning on age in any way. My estimates indicate that the age-targeted policy can improve welfare by almost \$2 per person per month in the market. Using a conservative estimate of 400,000 individuals in Colorado's non-group market, this translates to an additional net welfare gain of over \$9 million annually compared to a blanket premium subsidy.

Addressing adverse selection is important for the long-term success of the Exchange markets and, consequently, increasing insurance coverage in the U.S. As the vehicle through which individuals who meet income requirements could receive federal premium subsidies, the Exchanges played a key role in the overall ACA goal of decreasing the number of uninsured individuals, and some advocates view these markets with "managed competition" as a possible future for the delivery of health insurance (Enthoven, 1993). Selection works against this goal, however, and will lead to inefficiencies or, at worst, a complete unraveling of the insurance market. Fortunately, policy measures can mitigate the effects of selection. The Centers for Medicare and Medicaid seem to be aware of these issues,

[6] For example, Einav et al. (2010) find that selection raises monthly premiums by about \$16 in their context within the employer-sponsored insurance market.

as evidenced through a June 2016 press release proposing increased outreach to uninsured young adults as a way to strengthen the insurance marketplaces.[7] If individuals respond to prices, premium subsidies may be an additional way to increase coverage among young adults, and indeed the key policy takeaway of the present study is that age-targeted premium subsidies will offer the highest return to ameliorating the adverse effects of selection in the ACA Exchanges.[8]

The paper proceeds in six sections. Section 1 details the background of the ACA and Colorado's non-group market, as well as previous research. Section 2 describes Colorado's All-Payer Claims Database (APCD) in detail, and provides descriptive evidence of the market. Section 3 provides evidence of adverse selection, and Section 4 shows the estimates of demand and welfare implications of policy interventions. Robustness checks and extensions are discussed in Section 5. Section 6 provides additional discussion and concludes.

1 The ACA Exchanges and Literature Context

The community rating provisions of the Affordable Care Act restricted the factors on which insurers could price policies, allowing only limited pricing based on age, location, and smoking status, while entirely prohibiting any pricing on medical history or health status. This created the possibility for adverse selection in the non-group health insurance market. Using discontinuities in premiums created by the rating area boundaries, this study provides some of the first quasi-experimental evidence on selection related to the new ACA Exchanges.

1.1 Community Rating Provision of the ACA

The community rating provision both creates the possibility for adverse selection in the insurance market and provides a source of identification for my study. Under community rating, states were charged to set rating areas, within which any plan offered by an insurer had to have a uniform price. States, including Colorado, typically designated these as clusters of counties. Prior to community rating, insurers were able to price based on an individual's health status or medical history. Uniform pricing eliminated that practice, essentially giving individuals inter-temporal insurance against reclassification risk. However, this also opens the door to adverse selection. When insurers can no longer price based on risk, uniform pricing allows for the possibility that high-risk individuals will demand more insurance, leaving the market adversely selected.

The research design used in this study exploits these rating area boundaries to estimate how behavior responds to changing premiums. In 2014, Colorado's counties were clustered into 11 rating areas. Within a rating area, plans could only charge differential prices based on age and smoking status, and there were limits on ratio of prices charged to elderly and young individuals.[9] Insurers

[7] Centers for Medicare and Medicaid, "Strengthening the Marketplace by Covering Young Adults", Press Release, 21 June 2016.

[8] This is also an implication of the results from Tebaldi (2016) and Orsini and Tebaldi (2016). With data on individual level costs, the present study is able to more directly measure the selection effect, including selection within age groups.

[9] The premium for a 64 year old for any plan can be at most 3 times that faced by a 21 year old.

made county-level entry decisions, so although pricing was uniform across any particular rating area, plan availability was not necessarily uniform. However, if an insurer decided to operate in a county, it had to offer the full menu of plans, and could selectively offer some plans only in parts of the state.

1.2 Health Insurance Exchanges

The non-group health insurance market has played an increasingly important role in the U.S. health-care system since the passage of the Affordable Care Act. Under the ACA, every individual 18 or older who does not have a public insurance plan must either have health insurance or pay a fee to the federal government. To obtain health insurance, people generally have two options: obtain it through their employer, or purchase it themselves directly from an insurer in what is called the non-group market. If they purchase it themselves, they can either purchase it directly from an insurer independent of the ACA exchanges; or they can purchase it from an insurer through their state's exchange. Much of the "reform" part of the ACA was aimed at this market, including the development of the Exchanges, the availability of premium subsidies, and the standardization of plan characteristics. For these reasons, there has been increased attention and interest in the long-term viability of these markets and the Exchanges.

In practice, the Exchanges are organizations that facilitate the purchase of health insurance through a website. A potential consumer is able to log on to their state's website, and view options available to them based on the zip code of residence. Each plan is listed, along with the insurer offering the plan, the monthly premium required, and the generosity of the plan as categorized by metal level (bronze, silver, gold, and platinum). Consumers may also enter their income information so that, if eligible for income-based subsidies, they can view the monthly premiums they would need to pay after the subsidy.

Many states opted to use the federal government's Healthcare.gov site, but Colorado opted to run its own state-run exchange. Connect for Health Colorado, as it is called, experienced a smooth launch and exceeded its 2014 CMS enrollment target of 92,000. As of April 2014, Connect for Health Colorado announced that over 129,000 people had signed up for Qualified Health Plans (QHPs) through the exchange. Through special enrollment periods, QHP enrollment was 137,000 as of mid-2014.

Connect for Health Colorado serves as a clearinghouse, which means it does not actively manage (via bids or direct purchases) which insurers offer plans on the exchange. Any company that is authorized to sell insurance in the state is eligible to offer plans through the Exchange.[10] In 2014, ten carriers offered plans on the Exchange; all metal levels (Catastrophic, Bronze, Silver, Gold, Platinum) were offered, but not every company offered plans of every metal level. Most of the large insurers in Colorado participated in the Exchange, with Time Insurance as an exception. There was also a new entrant in 2014, one of the health care CO-OPs that received start-up loans from

[10] A company must obtain a license from the state the sell insurance, and is subject to the state's insurance regulations including reviews of proposed premium rate increases.

the ACA.

Colorado was also one of 31 states that expanded Medicaid, making individuals who earn up to 150% of the Federal poverty level (FPL) eligible starting in 2014. The Exchange subsidies were thus available for individuals earning from 150%-400% of the FPL, and individuals earning above 400% of the FPL were able to enroll in an exchange plan but would not receive any subsidies. Nationally, only 15% of Exchange enrollees did not receive financial assistance, while in Colorado this number was 40% in 2014. This may also explain why compared to the national average, Coloradans were opting most often for the lower priced bronze plans than anything else - nationally 20% of enrollment was in bronze plans, but in Colorado bronze plans had 40% of the enrollment.

1.3 Relation to the Literature

Previous research on selection in commercial insurance markets has typically used medical claims from large self-insured employers and the commercial group (employer-sponsored) market. There has been less of a focus on the non-group market, the site of many of the ACA reforms. Some studies look at selection in Massachusetts, which implemented ACA-style reforms in 2006; but this study is among the first to offer quasi-experimental evidence on selection directly related to the new ACA Exchanges.

1.3.1 Adverse Selection in Insurance Markets

After the important theoretical contributions made by Akerlof (1970) and Rothschild and Stiglitz (1976), early empirical work on adverse selection can be characterized as testing for a relationship between the demand for insurance and indicators of health status or utilization, in the spirit of Chiappori and Salanie (2000) who posed positive correlation tests for asymmetric information. Studies in this context generally indicate a large quantitative importance of adverse selection in health insurance markets (Cutler and Zeckhauser, 2000). Studies that investigate the extensive margin of selection into insurance markets include Long et al. (1988),Wrightson et al. (1987), and Farley and Monheit (1985), which offer mixed evidence on whether health care expenditures affect insurance purchases.

In the decade the followed Chiappori and Salanie (2000), the literature made enormous progress in moving from simply detecting asymmetric information, to disentangling moral hazard from selection and quantifying welfare costs of adversely selected markets by taking more structured approaches. How this is done has taken a variety of forms (Bajari et al. (2014), Einav et al. (2010), Handel (2013), Bundorf et al. (2012), Keane and Stavrunova (2016)), and several general approaches are reviewed and described in Einav et al. (2010). By estimating economic models that specified consumer preferences, this second wave of literature sought to quantify the welfare loss of asymmetric information and choice frictions, and then evaluating the potential impact of subsidies, mandates, and medical underwriting (Handel and Kolstad, 2015). Interestingly, this literature has indicated that even when adverse selection is substantial, estimated welfare losses tend not to be that large.[11]

[11] Einav et al. (2010) suggest this may be due to the literature focusing on the mis-pricing of existing contracts,

One important study in this line is Einav et al. (2010), whose identification strategy relies on premium variation in order to trace out the demand curve, and data on costs from medical insurance claims to estimate the plan's cost curve. My approach uses exogenous variation in premiums and data on costs from medical claims to detect and quantify the welfare effects of adverse selection. The premium variation I use comes from a boundary discontinuity, much like is used in Cabral and Mahoney (2014) to study the effects of Medigap premium variation on utilization in the Medicare market.

Although the decade since 2000 also saw an expansion in the contexts investigated which yielded important insights into selection in insurance markets, little attention has been focused on the non-group health insurance market.[12] Hendren (2013) investigates adverse selection in three non-group markets: long-term care, disability, and life insurance. However, there remains little direct evidence on asymmetric information in the commercial non-group (individual) health insurance market in the U.S. With increased concerns about expanding coverage in the U.S. policy context, and the key role that this market segment plays in the ACA for covering the remaining insured populations, this paper fills an important gap in this literature.

1.3.2 Insurance Markets and the Healthcare Reform

Studies have investigated the 2006 Massachusetts healthcare reform to shed light on the possible effects of various ACA reforms, and their interactions with adverse selection (Shi, 2016). Hackmann et al. (2015) investigate how the individual mandate interacts with adverse selection, by comparing coverage, premiums, and claim expenditure before and after reform in Massachusetts to other states. The authors find that the insurance market in Massachusetts was adversely selected prior to reform, and the mandate reduced adverse selection leading to a welfare gain of $335 dollars per person. Shepard (2014) finds that in the Massachusetts Health Insurance Exchange, insurers were responding to adverse selection by limiting network breadth and excluding star hospitals from their networks. Other studies that use the Massachusetts reform to offer insights into the ACA include effects of reform on hospital and preventative care (Kolstad and Kowalski, 2010), the labor market (Kolstad and Kowalski, 2012), and the effects of pricing regulations on consumer surplus (Ericson and Starc, 2015).

There has simultaneously been increasing interest in modeling insurance Exchanges in light of the ACA. Using data from the employer market to calibrate their model, Handel et al. (2015) study equilibria in Health Exchanges and estimate that community rating leads to substantial adverse selection, but pricing health status can lead to lower welfare in some cases because of the intertemporal uncertainty of risk reclassification. Scheuer and Smetters (2014) explore the potential consequences of acute adverse selection in an initial period, and applying their model to the initial open enrollment period of the Exchanges, suggest that initial adverse selection could cause the

and neglecting the welfare loss from contracts that are not offered at all.

[12] Recently investigated contexts include studies in employer sponsored insurance (Cardon and Hendel, 2001; Geruso, 2013), annuities (Finkelstein and Poterba, 2004), auto insurance (Cohen, 2005), and long-term care insurance (Finkelstein and McGarry, 2006).

market to get stuck in a Pareto-inferior "bad" equilibrium. Tebaldi (2016) uses enrollment data from the first year of California's ACA exchange to develop and estimate model of insurance demand and imperfect competition.

Other studies use research designs from other settings to infer the effects of the ACA (Moriya and Simon, 2016). One study that does use a quasi-experimental design to investigate the ACA is Knepper and Niedzwiecki (2014), which looks at selection and other effects of the 2014 Medicaid expansion. But at this point, direct evidence on the ACA Exchange markets has largely been descriptive. This study builds on the previous literature on healthcare reforms in the U.S. and their interaction with adverse selection by introducing the first quasi-experimental evidence on selection in the ACA Exchange markets. Moreover, the framework allows for a quantification of the welfare consequences due to selection and the welfare gains from potential policy interventions.

2 Data and Descriptive Evidence of Selection

Using Colorado's newly available All-Payer Claims Database (APCD), I am able to study outcomes with a broad view of the non-group health insurance market that includes all insurers operating in the state. This is important for the research design for three reasons. First, the medical claims provide data on healthcare utilization and spending, which are the indicators of selection in insurance markets. Second, with data from all insurers in the state, there is sufficiently robust coverage across the state including areas where any single insurer might not have a large enrollment, or where a dataset of only one insurer might not capture individuals substituting to a similar product offered by another insurer. Finally, the claims database includes the subscriber's zip code of residence, which allows me to identify the individuals who live near the boundary of a rating area and their local medical market.

2.1 Colorado's APCD

The primary source of data consist of medical claims submitted from insurance companies in the state of Colorado over the full 2014 calendar year, as well as the coverage eligibility lists for these insurance companies; thus, insured individuals are in the dataset even if they do not use healthcare. The medical claims provide, for each individual's encounter with the healthcare system, the medical diagnosis and procedure provided, the amount paid by the insurance company, and the amount paid by the patient. The data contain some limited demographic characteristics of the individuals: age, gender, and importantly the five-digit zip code, which will be used to identify individuals who live just one either side of rating area boundaries, and thus serve as comparison groups of similar populations living in the same medical market, but facing different insurance premiums.

The market segment of primary interest in this study is the non-group (individual) market. That market segment does not include individuals with public insurance coverage such as Medicaid or Medicare, and it also does not include individuals with employer-sponsored health insurance. Several variables were used to isolate the non-group market segment, because there were some inconsistencies across payers in how the market segment was indicated. Data Appendix C describes in detail the construction of the sample and creating pairs of neighboring Zip codes that lie across rating area boundaries.

2.2 Descriptive Evidence

This section provides a descriptive overview of Colorado's insurance market segments in 2014. The stability of the insurance Exchanges in terms of premiums and insurer entry and exit depends on the health utilization of the individuals enrolling in the non-group market, and in particular those newly enrolled in 2014 in ACA-compliant health plans. I use the medical claims to document the patterns of healthcare utilization and expenditures of individuals across market segments, allowing for comparisons of the non-group market to the employer-sponsored group market.

The ACA aimed to reduce the uninsured rate in the U.S., and heading into 2014 there was a lot of uncertainty about the healthcare utilization patterns of the newly insured. Were the newly insured healthy individuals, previously uninsured because they were simply not using the healthcare system? Or were they unhealthy people, who previously could not afford insurance due to medical underwriting or a pre-existing condition? New enrollees may have benefitted from community rating and guaranteed issue, which prohibits discrimination from insurers based on any pre-existing conditions and underwriting based on health history. There were also provisions aimed to expand coverage to otherwise uninsured individuals, including tax penalties for remaining uninsured and federal premium subsidies that made insurance more affordable. Data from Colorado Health Institute suggest that 60% of enrollees in 2014 received some financial assistance. Because there were many reasons that individuals were becoming newly insured, there was a great deal of uncertainty about who these individuals were, and what their health status was like.

Descriptive evidence shows that newly insured individuals were using a great deal of healthcare services. Table A1 shows the inpatient admission rates for newly insured individuals in 2014, those who were in the non-group market but were also insured in 2013, and those insured through employer-sponsored coverage. Inpatient admission rates are highest for the newly insured in the non-group market, and about 20% higher than the rate in employer sponsored plans. However, previously enrolled individuals in the non-group market had the lowest rates, suggesting that those previously able to obtain and afford insurance in that market were particularly healthy.

Table A1 also shows the same rates for outpatient visits for healthcare, which typically involve an individual receiving healthcare but without staying overnight at a hospital. A similar pattern emerges: newly enrolled individuals use outpatient services at the highest rate, and it is close to 15% higher than the employer market, but the lowest rate is for individuals previously enrolled in the non-group market. This same pattern holds when looking at the number of claims filed for professional medical services.

The medical claims also show that actual medical spending follows similar patterns to healthcare utilization. I use the "allowed amount" on the claim, which is the total reimbursement due to the provider including payments from both the insurer and the patient. This allows for a more even comparison across individuals in plans with different generosity. Figure I shows the average annual medical expenditure for an individual, and broken down by each type of service (Inpatient, Outpatient, Professional, or Pharmacy claim). Consistent with the utilization patterns, newly enrolled individuals are spending a great deal more on healthcare per year. Using a regression framework, these difference can be measured while controlling for the different location and age distributions across the market segments. These results are shown in Table A2, and suggest that even when controlling for age and location, newly enrolled individuals are spending between $800-$900 more per year than those who were previously insured and those in the group market. This gap may simply reflect pent-up demand from newly insured individuals, and it could possibly dissipate over time. The differences across market segments also suggest that medical underwriting was an important factor in why some individuals remained uninsured before 2014.

These patterns raise the question of to what extent the high costs of the new market are due to adverse selection. It could be the case that this new insurance market is simply a different, higher-utilizing population, with corresponding higher healthcare expenditures. Or, it could be the case that the high levels of spending reflect that only the relatively high cost individuals entered the market and thus show up in the claims data, that is, that the market is adversely selection. In this case, there could be healthy individuals potentially in the insurance pool, and if they were to enter the market, the spending levels could be more in line with the numbers from other market segments. Importantly, there is a large scope for selection. Table I shows the take up rates for individuals in the non-group market in Colorado in 2014. Overall, only 57% of individuals potentially in the non-group insurance market are insured, and these rates are particularly low for the younger population, with only 36% of 25-34 year olds insured. Thus the descriptive evidence shows that there is scope for adverse selection, and the utilization patterns are consistent with such a story. Alternatively, this newly insured population may simply be different in terms of utilization than individuals in other market segments. Detecting selection requires going beyond the descriptive evidence.

A first step beyond descriptives is to conduct a positive correlation test for asymmetric information. Table A3 shows the average annual medical expenditures for individuals in the non-group market in 2014, by metal level. This positive correlation between plan generosity and spending is in the spirit of Chiappori and Salanie (2000)'s test for asymmetric information. These patterns, however, cannot disentangle moral hazard from adverse selection. For the purpose of welfare analysis and policy interventions, it is necessary to be able to isolate the underlying cause of the positive correlation between demand for insurance and expected costs, and it is for that reason that I use a research design that allows for estimation of the isolated effect of selection from moral hazard.

3 Main Results: Are the Exchanges Adversely Selected?

Adverse selection implies that annual medical expenditures of the insured population increase in response to rising premiums, as relatively healthy individuals drop out of the market first when premiums rise. However, there are two primary challenges to empirically detecting adverse selection.

The first is the classic problem of disentangling moral hazard from adverse selection, as both produce the same types of patterns in data. Secondly, there is a broader endogeneity problem due to insurers setting an area's premiums to reflect the local population; as a consequence, naively looking at the relationship between costs and premiums in the cross section would yield a biased estimate of the selection effect.[13] In particular, insurers set premiums in response to an area's prices for healthcare services, treatment intensity by local physicians and hospitals, and the health status of the population. To isolate the effect of selection from moral hazard and address these broader endogeneity concerns, a source of exogenous premium variation is necessary.

I use geographic boundary discontinuities as a source of plausibly exogenous premium variation, and compare individuals living in zip codes just on either side of boundaries of rating areas established as part of the ACA beginning in 2014. The key assumption is that the populations on either side of the boundary are comparable except for the different premiums they face. Specifically, unobserved drivers of cost should be uncorrelated with the premium variation for the assumption to hold. Because I compare only individuals in neighboring zip codes that are also within the same local medical market, I argue and present evidence that the populations in my sample are comparable across the boundary. The reason that similar populations face different premiums is because of the rating area designation, and because zip codes are small relative to the entire rating area, the comparison populations face premiums that were largely determined by the rest of the rating area. Consequently, the premium level can be thought of as exogenous to the individuals in each boundary zip code. As an additional check, I instrument for premiums using a calculated "leave-out cost" instrument to use only the portion of premiums determined by the rating area when excluding each zip code of interest.[14] I show in Appendix E formally how using this identification strategy yields an estimate of adverse selection on the extensive margin, free from a moral hazard effect, and additionally how sorting across plans on the intensive margin will lead to an underestimate of the selection effect.

3.1 Boundary Discontinuity

The designation of rating areas in Colorado for 2014 are shown in Figure II at the zip code level.[15] Individuals living in zip codes along the rating area boundary, despite living only a short distance away from each other and facing the same healthcare provider markets, can face potentially very

[13] I present results from such a regression in Appendix Table A4, which shows a positive correlation between the premiums in an area and the medical spending of the insured population. However, these estimates will not reflect solely selection because insurers are able to endogenously set premiums in response to costs. Thus this regression does not serve as good evidence of adverse selection.

[14] A "leave-out cost" instrument is also used in Cabral and Mahoney (2014), who use Medigap premium discontinuities in local medical markets that span state boundaries to investigate the Medigap enrollment decision. Their research design, which investigates the effects of supplemental insurance on traditional Medicare spending, is however an estimate of a moral hazard effect.

[15] The rating areas were defined based on counties. In the claims dataset, however, I only observe zip codes, and there are some zip codes that cross county boundaries. For the main analysis, I assign these ambiguous zip codes to the county with the largest fraction of the zip code's population, and this is what is displayed in the map. As a robustness check, I repeat the analysis dropping these ambiguous zip codes and generally find similar patterns.

different premiums. To exploit this discontinuity, for each zip code on a rating area boundary, all of the neighboring zip codes that were in a different rating area were identified. Zip codes were then paired with a neighboring zip code if one was found the met the following criteria: (i) was in a different rating area, (ii) but the same local medical market, and (iii) the two zip codes mutually shared the longest border with each other.

In the main specifications, I use hospital referral regions (HHR's) as the definition of the medical market. This definition comes from the Dartmouth Atlas, and Figure II shows a map of the zip codes in Colorado assigned to HRRs. With this definition, the zip code pairing algorithm yields 32 pairs of zip codes. For robustness, I also consider other market definitions, such as hospital service areas (HSA's), which are depicted in Figure III. Because HSA's are smaller areas, this leaves less candidate zip codes for the boundary, as made clear through the figure.

Within each pair, individuals who resided across the boundary would face different premiums because of the way the community rating was designed. However, the difference varies across groups. Figure IV shows the change in monthly premium that a 30 year old non-smoker would face for a standard silver plan from HMO Colorado (Blue Cross Blue Shield). In some zip code groups, the difference amounts to around a 1% increase in monthly premiums, while in others it can mean an increase of over 40%. The median difference is a 15% increase in monthly premiums across the boundary.

Beginning in 2014, as a consequence of the community rating provisions of the ACA, insurers submit rate tables with age and area factors that will determine an individual's monthly premium. These factors are multiplied by a plan's base rate to determine the final premium. For example, a standard silver plan from HMO Colorado has a base rate of \$262.13 as the monthly premium. The monthly premium an individual i residing in zip code k would have to pay for the plan depends on the insurer's area factor $AREA_k$ and age factor AGE_i, by the following formula:

$$prem_{ik} = 262.13 \times AREA_k \times AGE_i$$

Denote by $g(k)$ the zip code pair to which zip code k has been assigned, and c_i the annual medical spending of individual i in 2014. Then the estimating equation to detect adverse selection is:

$$\log(1 + c_i) = \gamma + \delta \left(\frac{prem_{ik} - prem^L_{ik}}{prem^L_{ik}} \right) + \phi_{g(k)} + \mu_i \tag{1}$$

where $prem_{ik}$ is the premium that individual i residing in zip code k faces for insurance, and $prem^L_{ik}$ is the premium faced by that individual when residing in the less expensive side of the zip code pair.[16] The $\phi_{g(k)}$ denotes a fixed effect for each group of zip codes that have been matched, such

[16] In this formulation, the term on the δ coefficient equals zero for individuals in the less expensive zip code. A more standard way to run this regression would be to simply include the log(premium) for each i in the regression. However, because log differences approximate % changes best for small % changes, and the % increases across some of the boundaries are >20%, and up to a 50% increase, the interpretation of the log specification deviates compared to using the mathematical definition of the % change.

that the identifying variation comes only from individuals within matched zip codes. A positive estimate of δ suggests that as the relative premiums across the boundaries increase, the relative costs for those individuals on the high premium side increase. This indicates adverse selection.

Due to the complexity in accounting for consumer's choice sets, which involve many plans of different generosities, premiums, and contract structures (in terms of copays, co-insurance, deductibles), I collapse the choice set available in each zip code down to a single index of how expensive insurance is for individuals residing in that zip code. For robustness, I estimate Equation 1 using several different methods of constructing the insurance premium index, including the average silver plan available, the lowest cost silver plan, as well as using specific plans available statewide (e.g. BlueCross BlueShield's Silver Plan) as the index. The results are similar across these various measures of the index for how expensive insurance premiums are in a zip code.

The key identification assumption is that the populations residing in paired zip codes on either side of the boundary are comparable. The distributions of both observed and unobserved variables should be the same for populations living on either side of the rating area boundary, such that they are otherwise identical individuals facing different premiums only because they are part of different rating areas. Moreover, because zip codes are small relative to entire rating areas, the premium in a zip code is determined mostly by the individuals living in other parts of the rating area, and thus there is a component to the premium that is exogenous to individuals living in the zip code. In practice, one can test the extent to which the observables are balanced in populations across the rating areas.

More concretely for this context, comparability of the cross-boundary populations requires that the distribution of health status, risk factors, and preferences do not change discretely at the boundary. In practice, observable risk factors such as age can be conditioned on, requiring then that the distribution of health status is the same after conditioning on observables. Similarly, the populations living in the neighboring zip codes are assumed to have the same preferences in terms of demand for insurance and risk aversion given observables. Finally, the identification requires that the populations be in similar healthcare markets. This means that the providers that individuals in each population see for care should make similar types of treatment decisions, and the prices for healthcare services should be similar. Otherwise, higher cost individuals may not be sicker, but might just be in an expensive healthcare service market, or see doctors who always choose more aggressive and expensive treatment options. I require paired comparison zip codes to be in the same local medical market in order to control for these healthcare market considerations. I provide evidence in Appendix Table A5 that many individuals face offerings from the same insurers across the boundary. Importantly, individuals always have options from at least four insurers, and thus have an offering of many of the same provider networks.

There remains one final potential concern, that because a rating area's rates are determined to some extent by all zip codes it encompasses, part of any zip code's premiums can reflect that zip code's population. Because any particular zip code is only a small fraction of a rating area, this effect is likely to be small, but to address any remaining mechanical correlation between a zip code and the premiums in its rating area I use a "leave-out cost" instrument.[17] Define the instrument as the average spending in all of zip code k's rating area, excluding zip code k itself:

$$\text{Leave-out cost}_k = \mathbb{E}[c_i \,|\, i \in I_{r,-k}]$$

where $I_{r,z}$ denotes the set of individuals in rating area r and zip code z. The leave-out cost can then be used to instrument for the premiums in a zip code. In practice, I find that using the instrument makes little difference for the main results, which would be expected if each zip code had little or no effect on the rating area's overall premium level. However, this remains a useful check of the assumptions from the research design.

Although it is not possible to check for balance in unobserved variables, I can check whether leave-out costs and premiums are uncorrelated with observable demographic characteristics of the populations in the zip codes. To investigate this, I regress tract level demographic variables $\omega_{t,k}$ on the zip code leave-out costs:

$$\omega_{t,k} = \alpha + \beta \cdot \text{Leave-out cost}_k + \phi_{g(k)} + \mu_{t,k}$$

where $\phi_{g(k)}$ are zip code pair fixed effects. The results of this regression are given in Table II. While the instrument is not correlated with most of the demographic characteristics, it does appear to be correlated to age distributions, and the proportion of the hispanic population in an area. Because I control flexibly for age in the regressions by either including age fixed effects or by using each individual's residual costs after controlling for age and gender, the age distribution is not too concerning. To address the relationship with the % Hispanic, I run regressions with these and other demographic variables included as controls. I find that in the regressions, these demographic variables have very little explanatory power and do not change the results in a meaningful way.

One concern with this approach is the external validity of the results, given that the estimates are from a population living only along the boundaries of rating areas. Table III compares the boundary population used for estimation in the preferred specifications to the statewide sample. Though the boundary sample is not drastically different, it does reflect more rural population characteristics. This is to be expected, as the rating areas were designed so as to not cut through the center of metropolitan areas. There is thus the usual caveat of external validity, with the addition that the results may be particularly relevant for many areas in the U.S. that are relatively more rural.

3.2 Boundary Discontinuity Results

I estimate the main specification given in equation (1) using various measures of change in premiums across the pair of zip codes. In the preferred specification, I compare the change in the premium for the average silver plan in each zip code. I also estimate the regression using the change in premium for the second lowest cost silver (2LCS), as well as the change in premium for the silver plans offered by three statewide insurers, BlueCross BlueShield, Rocky Mountain Health Plans

[17] As also used in Cabral and Mahoney (2014).

(RMHP), and New Health Ventures (Colorado Access). The estimates from these regressions using the boundary discontinuity are given in the top row of Table IV. The positive and statistically significant coefficients on the premium increase terms show evidence of adverse selection in the non-group insurance market. In the preferred specification using the premium of the average silver plans, the estimate implies that a 1% increase in the premium of the average silver plan in an area leads to a 0.8% increase in the average medical expenditures of the insured population. This suggests adverse selection, because as premiums rise, relatively healthy individuals dropping out of the insurance market will lead to an increase in the average costs of the insured.

The second row of Table IV shows the estimates of the regression using the leave-out cost instrument for premiums. The results are similar to the OLS results, and with less variation across the different premium measures, again with the estimates suggesting that a 1% increase in premiums led to an approximately 0.8% increase in the average costs of the insured population.

As a placebo test, I estimate the exact same regression, but now using only individuals insured in the non-group market in 2013. Since there were no rating area boundaries in 2013, there should not be an effect of crossing the boundary for the insured person in 2013. Crossing a rating area boundary in 2013 should not cause a person to change their behavior. The results are given in Table V, and indeed the coefficients are small in magnitude and statistically swamped by the standard errors. As additional placebo checks, I run the same regression also on individuals in the employer-sponsored group market in 2014, as well as the 2013 Medicaid population, neither group of which should be affected by the boundaries. Columns (3)-(4) of Table V show the results for the group market, where there is no effect as expected, and Columns (5)-(6) show the Medicaid results. These placebo tests give confidence to the validity of the research design.

These is also heterogeneity in the acuteness of selection across different age segments of the market. Differences in take-up rates suggest this may be the case, with 60% of 55-64 year olds purchasing insurance, and only 36% of 25-34 year olds. To detect differences in acuteness of selection across age groups, I interact the % premium increase with age. However, when breaking down the sample by age group, the estimation becomes much less precise, and thus the preferred specification I use to identify heterogeneity in selection is a median regression which is more robust to outliers. I allow the coefficient of interest to interact with a polynomial of the individual's age, the regression results of which are shown in Table VI with a quadratic term. The age margins for the effect of a % increase in premiums are plotted in Figure V. Adding a cubic term did not affect the results much. The results suggest that individuals in the middle age groups are the ones with the most acute selection, particularly individuals in the 35-44 age range. Young individuals ages 25-34 have less, and this may be due to the fact that they tend to use little healthcare, thus leaving less scope for selection. The 55-64 age group is the one with the least selection, and at the older end of that age group, there appears to be almost no adverse selection.

3.3 Chronic Conditions

In order to provide further evidence that adverse selection is at work, I investigate whether, in addition to increased spending in the insured pool, the underlying risk of the insured pool also increases with premiums. A higher prevalence of chronic conditions in the insured population in locations with higher premiums would suggest a higher-risk pool, which would coincide with an adversely selected market, and help to assuage concerns about the effects being only driven by pent up demand.

To do this, I use the Healthcare Cost and Utilization Project's (HCUP) Chronic Condition Indicator (CCI) tool, which categorizes ICD-9 diagnosis codes as indicators of chronic or non-chronic conditions. The tool also categorizes the ICD-9 diagnosis codes into one of 18 body system indicators, which are listed in Appendix D. Table A6 shows the prevalence of these conditions in the non-group market. I then re-run the exact same specifications as before, but rather than using an individual's annual medical expenditures as the dependent variable, I use an indicator of whether the individual at any point during the year generated a diagnosis code that indicated a chronic condition.

The results of these regressions suggest that higher premiums are associated with a higher prevalence of chronic conditions in the non-group market. In particular, the coefficient on the premium of the average silver of 0.105 shown in Table VII implies that for each 10% increase in the monthly premium, the probability that an insured individual has a chronic condition increases by 1.05 percentage points. Statewide, the probability that an individual in the non-group market in 2014 had at least one chronic condition was .385; thus, a 10 % increase in monthly premiums would increase this probability to about 0.395. Analogous placebo regressions on the 2013 market show no effect, displayed in Table VII. These results provide further evidence of an adversely selected insurance market, rather than something else such as a temporary effect due to pent up demand.

As a robustness check, I estimate the same regressions but rather than simply using an indicator for a chronic condition as the dependent variable, I use an ordered logit model with the outcome given as a count of the number of chronic conditions. These results are presented in the Appendix Tables, and show similar patterns to the previous analysis when using counts of 0, 1, or 2+ chronic conditions (Table A7) or creating an additional bin of 3+ chronic conditions (Table A8).

4 Demand Estimation and Welfare Implications

The cost curve alone can indicate the presence of adverse selection, but it is only when combined with the demand curve that the estimates can be used to evaluate efficiency loss due to selection. After estimating demand, I use such an economic framework to evaluate welfare consequences in the market. The framework can also be used to evaluate policy interventions. I find that additional premium subsidies, and especially age-targeted premium subsidies, can be welfare improving in the sense that the benefit to consumers outweighs the costs of providing the subsidies. Finally, I extend the model to additionally take into account a social value of coverage.

4.1 Demand Estimation

Because the premium variation available through this research design comes from differences in premiums for all plans, I focus on the binary decision of whether to enroll in a plan or remain uninsured (extensive margin) as the entire choice set becomes more expensive. I use the same premium variation that was used for the cost estimation to estimate demand to be insured as health insurance premiums change. To do so, I use zip code-level survey data from the Colorado Health Institute to approximate the share of uninsured individuals in each zip code. I then estimate the zip code-level regression:

$$\log q_k = \alpha + \beta \cdot \left(\frac{prem_k - prem_k^L}{prem_k^L} \right) + \phi_{g(k)} + \mu_k \tag{2}$$

where q_k is the share of the non-group market that is insured. A downward sloping demand curve would yield a negative coefficient for β, which can be interpreted as the extensive margin elasticity - how does enrollment change as all plans become more expensive.

Although the survey data directly provides an estimate of the uninsured rate at the zip code level, this number is the share of the entire population that is uninsured, while the quantity of interest is the share of individuals eligible for non-group coverage who obtain it. Since the relevant market is only the non-group market segment, the overall uninsured rate needs to be translated into the share of only the non-group market that is uninsured. To do so, I calculate a factor τ which is equal to the inverse of the share of the entire statewide population that is part of the non-group market. Denote by r^k the zip code-level share of the entire population that is uninsured. I then calculate s^k, as $s^k = r^k \cdot \tau$, the non-group uninsured rate in zip code.[18] I then can use estimated zip code level non-group insured rate, $q_k = (1 - s^k)$, as the dependent variable in the demand regressions.

The results from this regression are shown in Table VIII, and the estimated elasticities do indicate the expected sign. The estimated demand elasticities are larger than what has been estimated for the employer-sponsored insurance market, but previous literature does indicate that the non-group market is more demand elastic. Column (1) shows the population-wide elasticity of -2.5. Gruber and Poterba (1994) use quasi-experimental variation to estimate the demand elasticity for health insurance among self-employed, and find estimates within the range of -1 to -2, with the preferred specification showing -1.8. Thus although my estimated elasticity is large for an insurance demand elasticity, it is only moderately larger than previous estimates for the non-group context.

Finally, I test for heterogeneity across age groups in demand. I do this estimating the same regression equation with an interaction for age, the results of which are shown in Column (2) of Table VIII. The estimated margins across age groups are shown in the bottom panel of Table VIII. Because the survey data only includes age data available in age bins, I present the results for four age bins (25-34, 35-44, 45-54, 55-64). The coefficients indicate that young consumers are most price

[18] In other words, $s^k \equiv \frac{\#\text{uninsured}_k}{\#\text{uninsured}_k + \#\text{insured in non-group market}_k} \approx \frac{\#\text{uninsured}_k}{\text{Total pop of zip code k}} \cdot \frac{\text{Total pop of state}}{\#\text{uninsured} + \#\text{insured in non-group market}} = r^k \cdot \tau$

sensitive to health insurance premiums, with the 25-34 age group having an elasticity of -4.01, while those above age 55 having more inelastic demand for insurance of -1.13. This pattern of decreasing elasticity across age groups is also found in Tebaldi (2016) looking at California's ACA Exchange. Moreover, the estimates are of similar magnitudes. Tebaldi (2016) estimates semi-elasticities, finding that among subsidized populations, when all prices increase by $100, 20-29 year old enrollment falls by 12.8% while 45-64 enrollment falls by 3.8%. Given that the average annual premium in California is about $3,000, Tebaldi (2016)'s esimates imply elasticities of -4.3 for the 20-29 year olds, and -1.3 for 45-64 year olds, very much in line with my estimates.

It is important to note, however, that my estimated demand elasticities are imprecise. Although the 25-34 group has a high elasticity of -4.01, and it is significantly different from zero, the 95% confidence interval includes values from [-6.2, -1.7]. Thus rather than focus only on the point estimate, an important takeaway is the pattern that younger people are more price-sensitive than older individuals. Given the imprecise elasticity estimates, I use a range of elasticities for the welfare exercises, including those from the literature, to test for sensitivity of the policy conclusions to these parameters.

4.2 Welfare Estimates and Policy Evaluation

The demand estimates can now be combined with the estimates of selection to evaluate welfare consequences in the market. I then evaluate a policy of additional premium subsidies aimed to decrease the welfare loss due to selection. I compare age-targeted subsidies to a blanket subsidy where subsidies are not conditioned on age in any way.

4.2.1 Welfare Estimates

To calculate the welfare effects of selection, I combine the estimates from the cost and demand regressions. I first impose that there exists a competitive equilibrium point $(p_{\text{eq}}, q_{\text{eq}})$ at the observed share of the non-group population that is insured, such that insurers break-even by earning enough premiums to offset their incurred costs. This implies that the market demand and average cost curves will intersect at this point. I then assume linear demand and cost curves (later I also relax this), and derive the equations for the curves using my estimates from the demand and cost regressions and linearizing around the equilibrium point. With linear forms for the demand and cost equations, it then also becomes straightforward to derive a simple form of the marginal cost curve as well:

$$D = \alpha + \beta \cdot q$$

$$AC = \gamma + \delta \cdot q$$

$$MC = \frac{\partial TC(q)}{\partial q} = \frac{\partial (AC(q) \cdot q)}{\partial q} = \frac{\partial}{\partial q}(\gamma q + \delta q^2) = \gamma + 2\delta q$$

These three curves based on the estimates are plotted in Figure VI. The competitive equilibrium, where the average cost and demand curves intersect, is the break-even pricing for insurers. The efficient allocation occurs where the marginal cost and demand curves intersect. At this point all individuals whose valuation for insurance is higher than their costs are insured. With the linear form for the curves, the efficient allocation give given by $q_{\text{eff}} = \frac{\alpha-\gamma}{2\delta-\beta}$ and $p_{\text{eff}} = \alpha + \beta q_{\text{eff}}$. Given my estimates, however, the demand and marginal cost curves do not intersect over the range of the share insured $\in [0, 1]$, and because demand is always higher than marginal costs, this suggests that the efficient allocation is full insurance.

With this framework it is simple to calculate various welfare quantities of interest. The amount of welfare loss to the marginal consumer is $D(q_{\text{eq}}) - MC(q_{\text{eq}})$. The welfare loss due to selection is the area above the marginal cost curve but below the demand curve, between q_{eq} and q_{eff}. Selection raises average premiums for silver plans from p_{eff} to p_{eq}. The estimated quantities when assuming linear curves are shown on Figure VI. The average welfare loss due to selection is \$44 per person, which can be more easily interpreted as adverse selection raising monthly premiums by \$78 (from \$120 to \$198). This corresponds to an estimate of welfare loss of \$44 per person per month. The linear specification is an upper bound on the welfare loss, particularly given the divergence between the demand and marginal cost curves, which leads to large welfare implications as the share insured approaches 1. For this reason, my preferred specification assumes a non-linear functional form. Specifically, because the parameter estimates from the research design can be interpreted as elasticities, I repeat the analysis assuming constant elasticity functional forms, and directly plug in my parameter estimates for both the average cost and demand curves. The results are shown in Figure VII, and show smaller magnitudes for the welfare loss due to selection compared to the linear case. Monthly per person welfare loss is \$26, or \$312 annually, and selection increases monthly premiums by \$47. Thus, I take the constant elasticity numbers to be a lower bound, and also the preferred specification.

While this welfare estimate is over twice as large a magnitude as found in studies in the employer market, this is only moderately larger than the estimate of \$241 annual per person welfare loss based on Massachusetts 2006 healthcare reform investigated by Hackmann et al. (2015). My preferred estimates indicate that welfare loss from selection to the marginal consumer is \$64 (EFC find \$138 in their context).

To obtain standard errors on the welfare estimates, I use the bounds on the 95% confidence interval for the cost estimates and compute the welfare quantities. Because the demand estimates are imprecise, relying on a zip-code level analysis, I use demand elasticities from the literature as the "true" values, and use only the estimated standard errors from the cost regressions. A sensitivity analysis suggests that the welfare estimates do not vary much given "true" demand estimates from a reasonable range. In the preferred CES specification, the estimated monthly welfare loss of \$25.70 has a 95% confidence interval from \$0.70 to \$51.30. For the linear case, the estimated welfare loss was \$43.80, with the 95% confidence interval ranging from \$0.80 to \$93.80.

The efficient allocation, in this case because it equals full coverage, could be achieved by an individual mandate or premium subsidies that were high enough to induce full coverage. My estimates suggest that such a policy would reduce average monthly premiums by about \$47 in the CES case, because the new enrollees would be lower cost, and thus average costs of the insured pool would fall, supporting lower premiums.

4.2.2 Premium Subsidy

One tool that can be used to increase coverage and address adverse selection is to subsidize consumer premiums. Funding the premium subsidies is not costless, however. Implementing subsidies to achieve the efficient level of coverage should only be undertaken if the welfare gains outweigh the costs of funding the subsidies. Indeed, EFC find that in their context of the employer-sponsored insurance market, the cost of a subsidy was too high to justify implementing subsidies. In this context, with more acute adverse selection, the subsidies do appear to lead to a net welfare improvement. To illustrate the evaluation of a subsidy policy, suppose the linear functional form for the cost and demand curves:

$$p_D = \alpha + \beta \cdot q$$

$$AC = \gamma + \delta \cdot q$$

The equilibrium in a competitive market is given by the break-even condition:

$$\begin{aligned} p_D &= AC \\ \alpha + \beta \cdot q &= \gamma + \delta \cdot q \\ q_{eq} &= \frac{\alpha - \gamma}{\delta - \beta} \\ p_{eq} &= \alpha + \beta \cdot q_{eq} \end{aligned}$$

Adverse selection causes under-insurance relative to an efficient level ($q_{\text{eq}} < q_{\text{eff}}$).

To evaluate the welfare effects of the policy, now suppose a subsidy of \$s per person is provided to consumers for the purchase of health insurance. Thus, when insurers post some price p_S, consumers effectively face the price $p_D \equiv p_S - s$. The equilibrium with a subsidy in a competitive market will be a quantity q_S such that:

$$p_D(q_S) + s = AC(q_S)$$

This is the point at which the total revenue to insurers will be equal to the total costs incurred.

Solving for q_S:

$$\begin{aligned} \alpha + \beta \cdot q + s &= \gamma + \delta \cdot q \\ \alpha + s\ \gamma &= (\delta - \beta) \cdot q \\ q_S &= \frac{\alpha + s - \gamma}{\delta - \beta} \\ p_S &= \alpha + \delta \cdot q_S \end{aligned}$$

Thus for any given subsidy amount s, these equations yield new equilibrium prices and quantities (q_S, p_S). The change in welfare resulting from the subsidy policy is then the increase in consumer surplus minus the cost of the policy:

$$\begin{aligned}\Delta W &= CB - C \\ &= \int_{q_{\text{eq}}}^{q_{\text{s}}} (D(q) - MC(q))dq - q_s \cdot s\end{aligned}$$

Benefits from subsidies exist if the value of coverage exceeds marginal cost for consumers between q_{eq} and q_{s}. The evidence of adverse selection suggests that this is the indeed the case in Colorado. However, this benefit needs to be weighed against the cost of the subsidy, $q_s \cdot s$, which is provided to all consumers (including infra-marginal consumers). In the next section, I use this approach to evaluate both an age-targeted premium subsidy, and a blanket subsidy provided to everyone without conditioning on age in any way.

4.3 Age-targeted Premium Subsidies

Heterogeneity across age groups raises the possibility that additional age-targeted subsidies could achieve welfare improvements more effectively than simply additional blanket subsidies. I use the framework from the previous section to compare a blanket subsidy for everyone to a subsidy that is targeted to age groups in order to maximize the effectiveness of the subsidy policy. This would be particularly germane when the public funds available would be limited, as the government could target dollars to their highest return.

The heterogeneity across age groups in both cost and demand factors leads to large estimated implications for age-targeted premium subsidies. On the cost side, age groups differ in how acutely they are selected; that is, as new individuals enter the market, there are differences in how quickly the average costs of that age group fall. On the demand side, different age groups have different levels of price responsiveness, which will affect by how much individuals respond to additional subsidies by entering the market. In addition, a large amount of the cost of a subsidy is the payment to infra-marginal individuals who would have purchased insurance anyway, and the take-up rates vary widely across age groups. By targeting the subsidies to age groups that had lower take-up rates to begin with and are more price responsive, as well as the most acutely selected, policymakers can minimize costs while achieving larger benefits.

To take into account the age heterogeneity, I proceed as before but deriving separate curves by age:

$$D = \alpha_{\text{AGE}} + \beta_{\text{AGE}} \cdot q$$

$$AC = \gamma_{\text{AGE}} + \delta_{\text{AGE}} \cdot q$$

$$MC = \gamma_{\text{AGE}} + 2\delta_{\text{AGE}} \cdot q$$

The same welfare quantities can then be computed as before, and each weighted by the share of the non-group market in each age bin. This approach allows for costs to reflect a changing age distribution of the insured pool, as allowed for in Tebaldi (2016), as well as to reflect selection within any age group.

One important consideration when looking at age heterogeneity is the restriction on pricing dictated by the Affordable Care Act, that any plan can charge a 64 year old at most three times the price charged to a 21 year old for the same plan in the same location. This 3:1 pricing ratio essential creates a subsidy from younger to older individuals.[19] I consider two cases for the policy evaluation regarding the age restrictions. First, I suppose that once the age-targeted premium subsidies are implemented, insurers can price to reflect the new cost structure regardless of the 3:1 restriction. For example, a subsidy targeted at the 25-34 age group would bring in new healthy enrollees to that age group, and I suppose insurers price to break even within each age group. In the second case, I require that insurers maintain the 3:1 ratio after the premium subsidies, which creates spillovers across the age groups. For example, now subsidizing 25-34 year-olds would still lower costs, but insurers would not be able to pass the lower costs on only to those in that age group. By keeping the distribution of premiums across the age groups fixed in the 3:1 ratio, premiums for all age groups would be lowered to some extent because the market-wide average costs have decreased. This lowering of all premiums would increase coverage for all age groups, even those that did not directly receive a subsidy, which is the source of the spillover of the policy. In this second case, insurers break even overall, but this may reflect profiting from some age groups while incurring losses in others.

The welfare effects of additional subsidies can be computed using the age-specific demand, average cost, and marginal cost curves, averaging across age bins weighted by the population share for each bin. Specifically, suppose there is an $\$s$ premium subsidy targeted at the 25-34 age group. In a competitive market, the equilibrium price will move to a point where insurers are breaking even taking into account the subsidy. This implies that insurers could lose up to $\$s$ per person, and still break even. In the first case I consider where insurers break even by each age group, the new equilibrium quantity supported is given by the quantity q_{s} that solves: $AC_{\text{AGE}}(q_s) - D_{\text{AGE}}(q_s) = s$ and must hold for all age groups. With linear curves that intersect at most once, this subsidy equilibrium quantity q_s will be unique for any given s. The equilibrium price that achieves q_s will be $p_s = D_{\text{AGE}}(q_s) + s$. The cost of such a subsidy is $q_s \times s$, and the benefit will be the area between q_{eq} and q_s that lies above the marginal cost curve but below the demand curve.

The results from this exercise are shown in the top panel of Table IX, and show that it is optimal to provide additional subsidies to every group except the 55-64 age group, with both the 25-34 and 35-44 groups receiving high enough subsidies to attain essentially full coverage within those two age groups. The first column of the table shows the optimal additional monthly subsidy amount across age groups, and it indicates that those below age 55 optimally would receive an additional \$13-15 per month. The second column shows the share of each age group that is covered after the subsidy, and only the 25-34 and 35-44 age groups reach nearly full coverage under the optimal subsidy. The table also reports the per person welfare benefits and costs from the subsidy. For the 25-34 age group, a benefit of \$5.17 indicates that the optimal subsidy to this age groups recovers over \$5 per

[19] During the debates around the ACA legislation, insurers argued for a ratio of 5:1 or even 7:1.

person in the market of the welfare loss due to adverse selection, while the cost of the policy is \$3.33 per person in the market. Overall, the subsidy to this age group yields a net welfare gain of \$1.83 per person of the welfare that was lost due to selection. The optimal policies together yield a net benefit of over \$5.50 per person in the market.

The final column in the table shows the ratio of the benefits over costs. Certainly a policy should only be undertaken if this ratio is at least 1, indicating that the benefits of the policy outweigh the costs. However, evaluating the policy should also take into account the cost of raising public funds. A value of 1.3 has been used in the literature, indicating a cost of about \$1.30 for every dollar of revenue that is raised through taxation.[20] Using this number, additional subsidies should only be implemented if the benefit/cost ratio is above 1.3. Given my estimates, this is the case for all age groups except the 55-64 group. The overall benefit/cost ratio of almost 1.6 suggests that the benefits of additional premium subsidies outweigh the costs of providing the subsidies, even taking into account the cost of public funds.

The age-targeted policy can be compared to a policy of premium subsidies that spends the same amount of public funds, but does not use any age targeting. As shown in the bottom panel of Table IX, this policy will lead to a lower net welfare gain and lower levels of coverage compared to spending the same amount of money on age-targeted subsidies. The overall benefit/cost ratio of 1.4 suggests, however, that such a policy would still have benefits that outweigh the costs, and so additional premium subsidies would be welfare enhancing even without age targeting. The best use of public funds, however, would entail age targeting.

When pricing on age is not flexible, but restricted as required by the ACA, the gains from using age-targeted premiums will be mitigated, but still present and will create spillovers as discussed above, so that even age groups that do not receive subsidies can benefit. Rather than decreasing the premium for only the subsidized age group as the average cost of that age group fall, insurers would need to decrease the premiums for all age groups to maintain the 3:1 pricing distribution. This induces spillovers to non-targeted age groups. The results that take this into account are shown in Table A9, and the spillovers show up clearly in the non-targeted groups. For example, the 55-64 age group receives essentially no direct subsidy in the optimal case, while the 35-44 and 25-34 age groups are heavily subsidized. However, the unsubsidized 55-64 age group still experiences a net benefit. This is because although they receive no subsidy, as healthy individuals in the subsidized age groups enter the market, the marketwide average costs fall and thus insurers lower premiums. Keeping the ACA age pricing ratio means that the premiums will fall for all groups, and thus even those that are not subsidized receive a benefit. Overall, the benefit/cost ratio for the age-targeted subsidies with the ACA pricing restrictions are 1.1, suggesting that the benefits are greater than the costs. But when taking into account a cost of public funds of 1.3, this policy is no longer clearly cost effective. Under a pricing ratio of 5:1 or 7:1, however, the subsidies would have effects closer to the first case, and perhaps be cost effective.

Alternatively, the equilibrium can be derived using demand and average cost curves that are

[20] See Ballard et al. (1985) and Poterba (1996).

non-linear. In particular, since I estimate parameters that can be interpreted as elasticities, I can assume that the demand and cost curves exhibit an isoelastic functional form. Specifically the curves have the form:

$$D(p) = \alpha_{AGE} \cdot p^{\beta_{AGE}}$$

$$AC(p) = \gamma_{AGE} \cdot p^{\delta_{AGE}}$$

Here, β_{AGE} and δ_{AGE} are quantities estimated directly from regressions using the research design, with the demand coefficient β_{AGE} coming from a $\log - \log$ specification regressing share insured on premiums. α_{AGE} and γ_{AGE} are, as before, pinned down by imposing a break-even equilibrium. The same logic as the linear case applies to calculate the welfare quantities of interest, the only difference being that numerical integration is used in some calculations (for example, finding the area below the demand curve and above the marginal cost curve). Otherwise the same welfare exercises can be carried out using these new functional forms.

I find very similar patterns to the linear case when using the constant elasticity forms, but with smaller magnitudes of optimal subsidy amounts. The patterns are similar in that the optimal age-targeted subsidy gives no subsidy to the 55-64 age bin, and there are welfare gains made by the policy. The blanket subsidy now suggests a slightly lower $7.88 per person per month, and again is near the cost of public funds threshold with a benefit/cost ratio of 1.36. The full set of results are shown in Tables A10 and A11. Interestingly, when maintaining the ACA price restriction, the CES functional form specifications show a higher benefit/cost ratio of 1.25, suggesting that such a policy would be close to conferring benefits to consumers equal to the cost of the policy. These results are shown in Table A12. If there is a social value to expanding coverage, as I discuss in the next section, this policy could very well have benefits that outweigh the costs even when maintaining a 3:1 pricing ratio.

These estimates and policy recommendations coincide with survey results from the Commonwealth Fund, which cite affordability as one of the primary reasons that individuals remain uninsured, even under the existing subsidy schedule.[21] However, this analysis has been focused entirely on the loss to (private) consumer welfare due to selection. If there is a social value of increasing insurance coverage, either because of an instrinsic value to society or the externality costs to society in the form of charity care when uninsured individuals use emergency department services, then the current estimates would under-state the value of premium subsidies. This social value of increasing coverage can be embedded into this framework and may better reflect the total value of additional premium subsidies.

4.4 The Social Value of Coverage

This framework can be extended to include a social value to increasing insurance coverage in the welfare evaluation. Indeed, the first goal of the Affordable Care Act was to decreased the nationwide

[21] "Of uninsured adults who are aware of the marketplaces or who have tried to enroll for coverage, the majority point to affordability concerns as a reason for not signing up" (Collins et al., 2016).

uninsured rate. There are a variety of potential sources of this social value. Society may place an intrinsic value on insuring those who may not be able to afford insurance, as has been part of the motivation for other publicly funded insurance programs in the U.S. Alternatively, society may value increasing coverage because of the externality from uninsured individuals forgoing care until they utilize more expensive emergency department services. The Emergency Medical Treatment and Active Labor Act (EMTALA) requires hospital emergency departments that accept Medicare payments to provide medical care to anyone seeking treatment, regardless of their ability to pay. This care provided to uninsured individuals is typically uncompensated care, imposing an externality on the insured population which can result in a social value of coverage above the individual's willingness to pay for insurance.

Now, suppose that society has a social value of \$$r$ per person to having them insured. One can simply add this social benefit to the change in welfare equation from before. The change in welfare resulting from a subsidy of \$$s$ per person is now the consumer benefit minus the cost of the policy as before, but with the additional social benefit of \$$r$ for each additional person covered by the policy:

$$
\begin{aligned}
\Delta W &= CB - C + SB \\
&= \int_{q_{\text{eq}}}^{q_{\text{s}}} (D(q) - MC(q))dq - q_s \cdot s + (q_s - q_{\text{eq}}) \cdot r
\end{aligned}
$$

A social value of insurance of $r = 0$, the lower bound, results in the cases investigated above. For $r > 0$, this would make the subsidies more beneficial relatively to the previous results. For the constant elasticity case, my estimates suggest that any value of $r \geq \$78$, i.e. a value to society of at least \$78 per month of insuring an individual, would lead to an efficient outcome where subsidies are used to obtain full coverage for the entire population. When only a blanket subsidy is available (i.e. no age-targeted subsidies available), the threshold is $r \geq \$205$ to obtain full coverage as the efficient policy target. Assuming the linear functional forms, these thresholds for r are \$121 and \$354, respectively.

Although I do not estimate r in this paper, previous literature and policy discussions can give a sense of what values might be plausible. For example, the ACA individual mandate penalty is one example of a number that aimed to reflect the cost to society of remaining uninsured. For 2016, the penalty for failing to obtain insurance coverage was \$695 per adult, or about \$58 per month. Although this number is still below the \$78 per month required to justify subsidies that attain full coverage for the entire population given my estimates, it suggests that the subsidies in my baseline cases that exclude the social value of insurance do offer only a lower bound on the socially optimal subsidy.

5 Robustness and Extensions

This section checks for robustness across different definitions of medical market (using Hospital Service Areas rather than Hospital Referral Regions), and then investigates what drives the results through extensions. Breaking the medical spending into whether it is ER or non-ER spending reveals that the selection is primarily driven by changes in non-ER spending across the boundaries. Had the results been driving by differences in ER spending, one might worry that moral hazard could have been at play, in that the insured population began to engage in more risky activities. The fact that the differences driven entirely by the non-ER spending providers further support that individuals have private (or at least non-priced) information about their health status, and use this information to make their insurance decisions.

5.1 Robustness

One difficulty in using zip codes in this context is that the rating areas are defined by groups of counties, and zip codes can cross county boundaries. In particular, there are several zip codes in Colorado that cross county boundaries of two counties that are in different rating areas. In the preceding analysis, I assigned each of these zip codes to the county which contained the largest proportion of the zip code's population. For robustness, here I present results from analysis that completely drops all of these "ambiguous" zip codes.[22] The results for the OLS are shown in Table A13, and the results for the IV are shown in Table A14. For these specifications, the sample size has decreased due to there being fewer boundary zip codes available, but the results are quite similar and thus fairly robust to dropping these zip codes.

I also estimate the same regressions, but when constructing the pairs of zip codes, use HSA as the definition of the medical market, rather than the HRR. Because HSAs are much smaller than HRRs, there are fewer candidate boundary zip codes in this case, and thus fewer pairs of zip codes and smaller sample sizes. In particular, the number of zip code pairs falls from 32 to 19. The OLS results from using the HSA criteria are shown in Table A15, and though there is more variation across premium measures, and less statistical significance, the point estimates are not very different from before. Using the leave-out cost instrumenting strategy yields more statistical significance and less variation across premium measures, as shown in Table A16. Using the instrument thus matches quite closely to the previous results from using HRRs.

Finally, rather than restricting zip codes to be matched into pairs, I allow for groups of more than two zip codes to be matched together. This yields a larger sample size. However, here the coefficients are smaller than before, and are much less statistically significant, as shown in Table A17. Using the leave-out cost instrument, no premium measure remains statistically significant, and the coefficients have fallen to around 0.4.

I have additionally considered quantile regressions for the boundary equation. Because many individuals have no healthcare spending, there are many zeros in the data for the dependent variable, annual medical spending. In some zip codes, almost half of the insured individuals in the boundary sample have no healthcare spending in 2014. I therefore run quantile regressions, and at the 75th percentile, the measured adverse selection is similar to the mean regressions. Using low percentiles

[22] I drop these zip codes before running the matching algorithm, so compared to the previous analysis, it is possible now for zip codes to be arranged in different pairs.

(median or lower) does not yield meaningful results because in those ranges, many individuals have little or no medical spending and so there is no variation to be explained.

5.2 Extensions

In order to look at what types of spending might be driving the selection, I run all of the previous specifications, breaking down the medical spending paid by the plan into whether the expenses were incurred as part of emergency room care or not.[23] These results are shown in Table X. Table X uses the % increase for BlueCross BlueShield's Silver plan, and columns (1)-(3) show that with the zip code group fixed effect, emergency room spending does not change much as one moves across the rating area boundary. Columns (4)-(6) show clearly that the adverse selection detected previously is driven by changes in spending in non-emergency room spending. The zip code group fixed effect causes the premium coefficient for ER spending to drop to 0.07, while for non-ER spending, the coefficient increases to 0.89, similar to the magnitude found in the main results. This suggests that the selection is due to difference in non-emergency spending as one moves across the rating area boundary. The results are very similar when the % increase in premiums is measured using products from other insurers, or whether it is measured as the change in the premium of the average silver plan available.

5.3 Sorting Across Metal Levels and Moral Hazard

The preceding analysis has focused on the extensive margin and the overall costs of the entire insured pool, due to the nature of the premium variation offered by the rating areas. However, individuals are able to respond to higher premiums not only by leaving the insurance market, but also by sorting into a perhaps less generous plan (e.g. Bronze rather than Silver plan). Thus, there may be adverse selection across metal levels as well. However, if an individual's behavior depends on the metal level in which they are enrolled, i.e. if there is moral hazard, this may lead one to overstate the estimated selection effect.

To make concepts precise, I will use the term "extensive moral hazard" to refer to an individual changing their behavior with respect to healthcare utilization when they are insured vs. uninsured, and I will refer to "intensive moral hazard" as an individual changing behavior depending on the metal level in which he or she is enrolled. Because I have estimated the changing utilization based on only the insured sample, extensive moral hazard will not effect the estimates presented here. The estimates can be thought of as, given a certain level of extensive moral hazard, representing the selection effect.

Intensive moral hazard, however, may bias the selection effect. However, in the typical case, if higher premiums lead an individual to sort into a less generous plan (e.g. Silver to Bronze), and then moral hazard leads individuals in less generous plans to use less healthcare, that would leave the previous estimates to understate the true selection effect. If on the other hand, one

[23] In the sample, emergency spending is about 11% of the total medical spending

could assume that healthcare utilization decisions did not depend on the metal level, then the previous estimates have identified exactly the selection effect. In other words, if we assume away intensive moral hazard, the estimates presented thus far are the selection effect. With intensive moral hazard, theory suggests these estimates would be an underestimate of the selection effect. All of these concepts are precisely defined, and the reasoning why moral hazard would cause the estimated effect to be an underestimate of selection, in an explicit framework in Appendix E.

In order to try to estimate the intensive moral hazard itself, I investigate sorting across plan metal levels. There are, however, two complications. First, the rating areas do not offer exogenous premium variation across metal levels. That is, for each insurer, the prices of are all adjusted depending on the rating area by multiplying the premium by the same rating area factor. Secondly, this is not a particularly large market segment, and there is variation across insurers on how well the plan metal level of each individual can be identified in the data. Moreover, the sample of individuals on the boundary is small, and therefore for some metal levels, very few individuals could be identified in boundary zip codes. For example, I am only able to identify a small number of individuals from the boundary sample who are in Gold or Platinum plans.

I therefore focus here on Bronze and Silver plans, and present results analogous to before, but for individuals only in each of these plan metal levels. Table A19 shows the same regression, but now using only the subsample of individuals which have been identified to have enrolled in a Bronze plan. The results show strong, positive coefficients, indicating that the average costs of the Bronze enrollee pool increases with premiums. This can be driven by two effects: (1) the relatively healthy individuals from Bronze plans drop out of the market as premiums increase, and (2) some less healthy individuals who were previously in more generous plans may sort into Bronze plans. The results for Silver plans, shown in Table A20, are also positive, but not statistically significant.

The intensive moral hazard, or how individuals respond in healthcare utilization when moving from Bronze to Silver, can then be estimated. This is done by deriving, for each metal level, the marginal cost curve from the demand and average cost curves, as was done for the entire sample in Figure VI. Then, the point by point difference between the estimated MC curves for Silver vs. Bronze plans give an estimate of moral hazard. This is because this difference would indicate the differences in healthcare expenditures for the marginal individuals as they move from Silver to Bronze. If the two marginal cost curves are identical, that would indicate that individuals do not change the amount of healthcare services they use during the year because of being moved from a Bronze to a Silver plan. In practice, however, this is imprecise because the boundary research design requires eliminating much of the available data, leaving only a small sample of boundary residents. Moreover because each insurer's rating factor is applied to all plans offered, there is not directly exogenous premium variation across metal levels. Thus, in this context, it has proven difficult to estimate directly and confidently the magnitude of the intensive margin moral hazard.

6 Discussion and Conclusion

This study has presented robust evidence of adverse selection in Colorado's non-group health insurance market, including the state's ACA Health Insurance Exchange. Welfare analysis suggests that selection leads to fairly large welfare losses in this market, leaving room for policy interventions that increase coverage to potentially improve market efficiency. The heterogeneity across age groups in both the estimated selection effect and demand side price responsiveness implies that age-targeted policies can be particularly effective. Indeed, I find that while an additional blanket subsidy would be a cost effective improvement to welfare, age-targeted subsidies would be a more effective use of public funds.

There are several limitations to this study. First, because the identification of selection relies on differences across the rating area boundaries, the analysis must be restricted to a small subset of all individuals in the dataset. Moreover, these individuals tend to reside in more rural areas of the state, as the rating areas are designated so as to not cut through metro areas.

A second limitation is the threat to identification stemming from the fact that individuals purchasing insurance through the state-run insurance exchange are eligible for premium subsidies, and in 2014 about 60% of individuals purchasing through the exchange in Colorado received some financial assistance. Data on individual incomes and premium subsidies are not available in the APCD.[24] In terms of the estimates presented in this paper, I argue that the premium subsidies will lead to an underestimation of the selection effect, and therefore they should be interpreted as a lower bound on adverse selection.

This study has provided the first quasi-experimental evidence of adverse selection using cost data directly from the ACA Exchange markets. However, it will be important to take into account further research on these markets, including evidence from other states, datasets, and research designs. Because only the first year of data was available for this study, there remain many important questions relating to the dynamics of the markets and the outcomes once the markets have begun to stabilize. As data becomes available for more years and contexts, there will be an increasingly greater scope to understand the effects of selection and competition in healthcare and insurance markets.

DUKE UNIVERSITY

[24]The premium subsidies are available to individuals making below 400% of the Federal Poverty Line (FPL), and they work by capping the % of income that an individual pays for health insurance. For example, someone at 200% FPL would have to pay no more than 6.3% of their income in order to purchase the second-lowest cost silver plan in their area.

References

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

Bajari, P., C. Dalton, H. Hong, and A. Khwaja (2014, December). Moral hazard, adverse selection, and health expenditures: A semiparametric analysis. *The RAND Journal of Economics 45*(4), 747–763.

Ballard, C. L., J. B. Shoven, and J. Whalley (1985). General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States. *The American Economic Review 75*(1), 128–138.

Bundorf, M. K., J. Levin, and N. Mahoney (2012). Pricing and Welfare in Health Plan Choice. *The American Economic Review 102*(7), 3214–3248.

Cabral, M. and N. Mahoney (2014, November). Externalities and Taxation of Supplemental Insurance: A Study of Medicare and Medigap. SSRN Scholarly Paper ID 2372163, Social Science Research Network, Rochester, NY.

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

Chiappori, P. and B. Salanie (2000). Testing for Asymmetric Information in Insurance Markets. *Journal of Political Economy 108*(1), 56–78.

Cohen, A. (2005). Asymmetric Information and Learning: Evidence from the Automobile Insurance Market. *The Review of Economics and Statistics 87*(2), 197–207.

Collins, S. R., M. Z. Gunja, M. M. Doty, and S. Beutel (2016, August). Who Are the Remaining Uninsured and Why Haven't They Signed Up for Coverage? Technical report, Commonwealth Fund.

Cutler, D. M. and S. J. Reber (1998, May). Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection. *The Quarterly Journal of Economics 113*(2), 433–466.

Cutler, D. M. and R. J. Zeckhauser (2000). Chapter 11 The anatomy of health insurance. Volume 1, Part A, pp. 563–643. Elsevier.

Einav, L., A. Finkelstein, and M. R. Cullen (2010, August). Estimating Welfare in Insurance Markets Using Variation in Prices. *The Quarterly Journal of Economics 125*(3), 877–921.

Einav, L., A. Finkelstein, and J. Levin (2010). Beyond Testing: Empirical Models of Insurance Markets. *Annual Review of Economics 2*(1), 311–336.

Enthoven, A. C. (1993, January). The History And Principles Of Managed Competition. *Health Affairs 12*(suppl 1), 24–48.

Ericson, K. M. M. and A. Starc (2015, January). Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange. *Review of Economics and Statistics 97*(3), 667–682.

Farley, P. J. and A. C. Monheit (1985). Selectivity in the demand for health insurance and health care. *Advances in Health Economics and Health Services Research 6*, 231–252.

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

Finkelstein, A. and J. Poterba (2004). Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market. *Journal of Political Economy 112*(1), 183–208.

Geruso, M. (2013). Selection in Employer Health Plans: Homogeneous Prices and Heterogeneous Preferences. Technical report.

Gruber, J. and J. Poterba (1994). Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed. *The Quarterly Journal of Economics 109*(3), 701–733.

Hackmann, M. B., J. T. Kolstad, and A. E. Kowalski (2015, March). Adverse Selection and an Individual Mandate: When Theory Meets Practice. *American Economic Review 105*(3), 1030–1066.

Handel, B., I. Hendel, and M. D. Whinston (2015, July). Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk. *Econometrica 83*(4), 1261–1313.

Handel, B. R. (2013). Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review 103*(7), 2643–82.

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

Hendren, N. (2013, September). Private Information and Insurance Rejections. *Econometrica 81*(5), 1713–1762.

Jacobs, P. D., N. Duchovny, and B. J. Lipton (2016, July). Changes In Health Status And Care Use After ACA Expansions Among The Insured And Uninsured. *Health Affairs 35*(7), 1184–1188.

Keane, M. and O. Stavrunova (2016, January). Adverse selection, moral hazard and the demand for Medigap insurance. *Journal of Econometrics 190*(1), 62–78.

Knepper, M. and M. Niedzwiecki (2014). Public Expansions and Private Spillovers: What the 2014 Medicaid Expansions Mean for Private Insurance Markets. Technical report.

Kolstad, J. T. and A. E. Kowalski (2010, May). The Impact of Health Care Reform On Hospital and Preventive Care: Evidence from Massachusetts. Working Paper 16012, National Bureau of Economic Research.

Kolstad, J. T. and A. E. Kowalski (2012, March). Mandate-Based Health Reform and the Labor Market: Evidence from the Massachusetts Reform. Working Paper 17933, National Bureau of Economic Research.

Long, S. H., R. F. Settle, and C. W. Wrightson (1988, October). Employee premiums, availability of alternative plans, and HMO disenrollment. *Medical Care 26*(10), 927–938.

Moriya, A. S. and K. Simon (2016, May). Impact of Premium Subsidies on the Take-up of Health Insurance: Evidence from the 2009 American Recovery and Reinvestment Act (ARRA). *American Journal of Health Economics 2*(3), 318–343.

Orsini, J. and P. Tebaldi (2016, March). Regulated Age-Based Pricing in Health Insurance Exchanges. SSRN Scholarly Paper ID 2464725, Social Science Research Network, Rochester, NY.

Poterba, J. M. (1996, January). Government Intervention in the Markets for Education and Health Care: How and Why? In *Individual and Social Responsibility: Child Care, Education, Medical Care, and Long-Term Care in America*, pp. 277–308.

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

Scheuer, F. and K. Smetters (2014, May). Could a Website Really Have Doomed the Health Exchanges? Multiple Equilibria, Initial Conditions and the Construction of the Fine.

Shepard, M. (2014, November). Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange. *Harvard, Job Market Paper*.

Shi, J. (2016, January). Income Responses to Health Insurance Subsidies: Evidence from Massachusetts. *American Journal of Health Economics 2*(1), 96–124.

Tebaldi, P. (2016). Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA. Technical report, JMP.

Wrightson, W., J. Genuardi, and S. Stephens (1987). Demographic and utilization characteristics of HMO disenrollees. *GHAA journal 8*(1), 23–42.

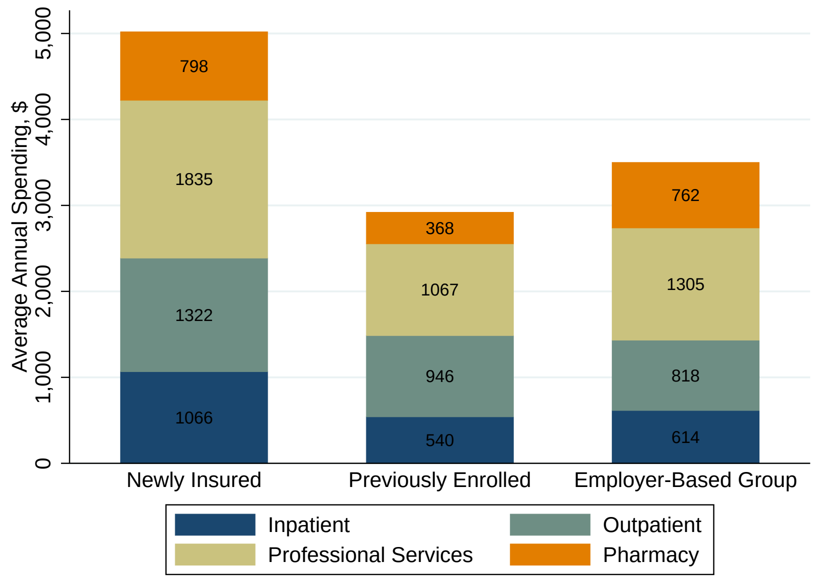

Figure I: 2014 Annual Medical Spending

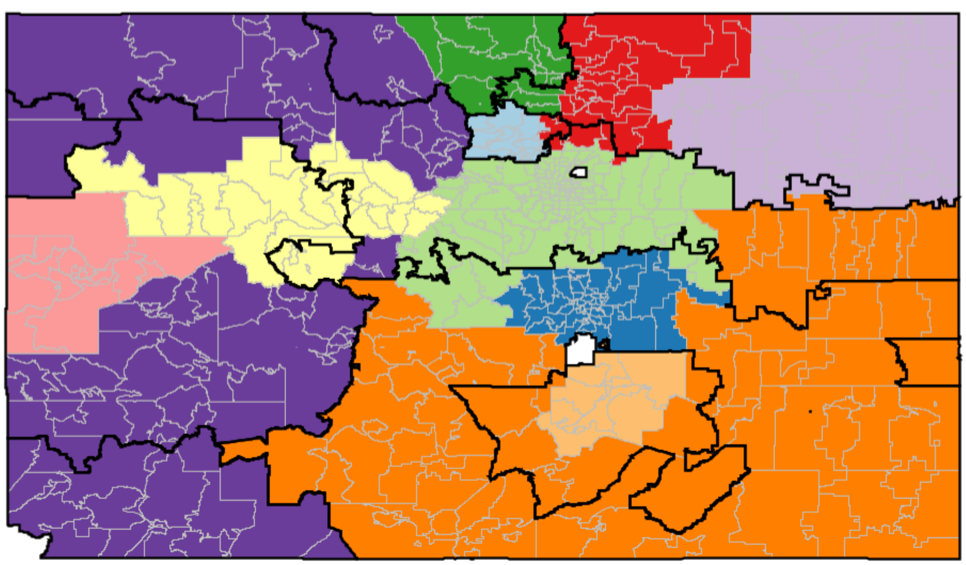

Figure II: 2014 Rating Areas in Colorado. 5-digit Zip codes are shown grouped into Rating Areas based on color. The outlines designate the grouping of Zip codes into medical markets, here defined as the Hospital Referral Region (HRR).

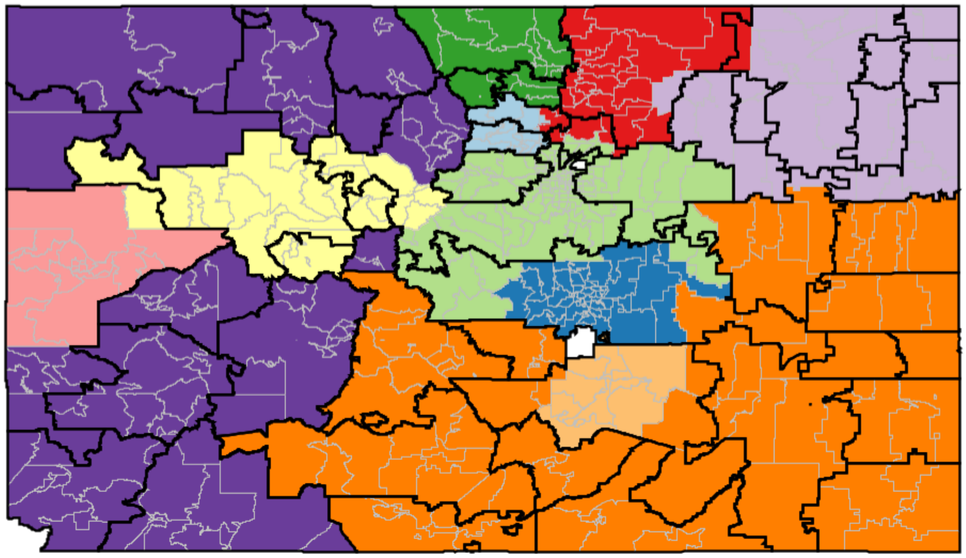

Figure III: 2014 Rating Areas in Colorado. 5-digit Zip codes are shown grouped into Rating Areas based on color. The outlines designate the grouping of Zip codes into medical markets, here defined as the Hospital Service Areas (HSA).

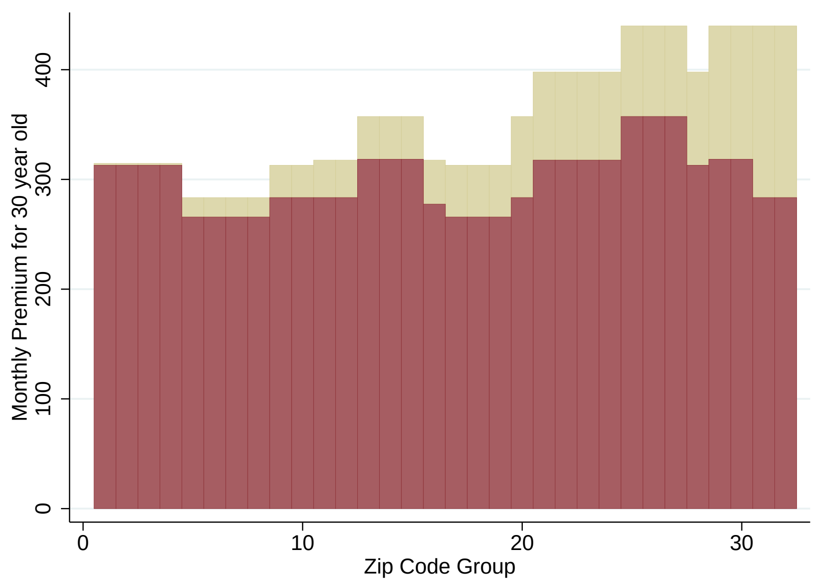

Figure IV: Change in BlueCross BlueShield Silver Premium Across Rating Area Boundary. There are 32 pairs of neighboring zip codes that cross a rating area while remaining in the same HRR. This graph shows the change in monthly premium for BlueCross BlueShield's Silver plan across each of the 32 pairs of zip codes

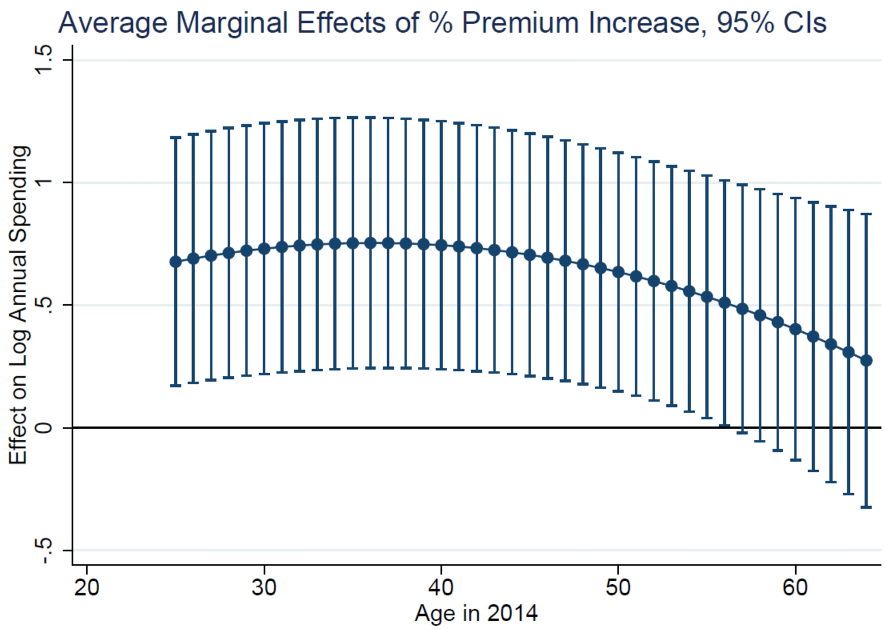

Figure V: Cost Heterogeneity Across Age. This shows graphically the marginal effects of the parameter of interested when interacted with a polynomial of age. Selection is estimated to be most acute for the middle age groups, lower for the 25-34 ages, and nearly zero for individuals above 60.

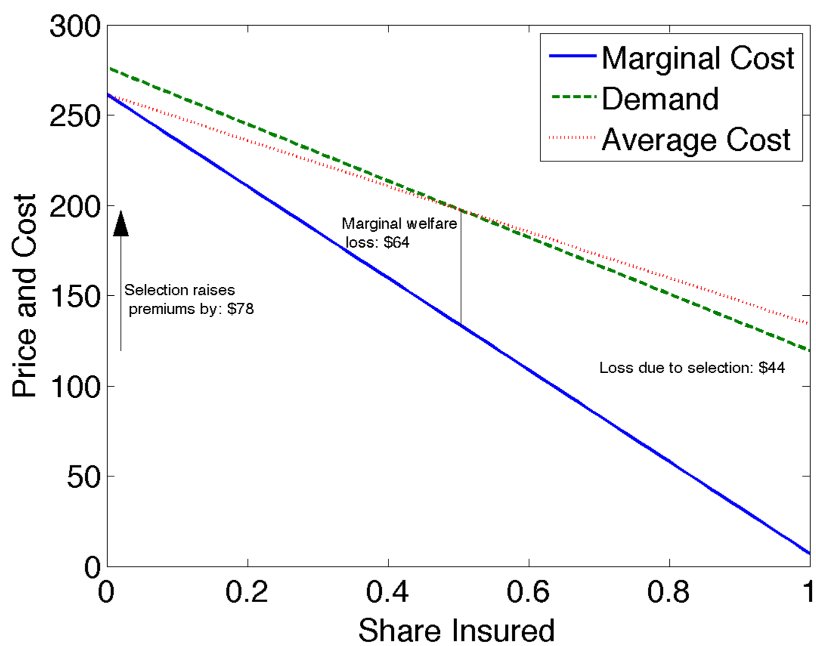

Figure VI: Welfare estimates with linear functions. These numbers come from combining the demand and cost estimates, and imposing a competitive break-even equilibrium at the observed quantities. Because demand is always above marginal costs, this suggests the efficient allocation is full coverage. The calculated welfare loss due to selection is $22 per person per month. I take the linear functional form as an upper bound for welfare estimates.

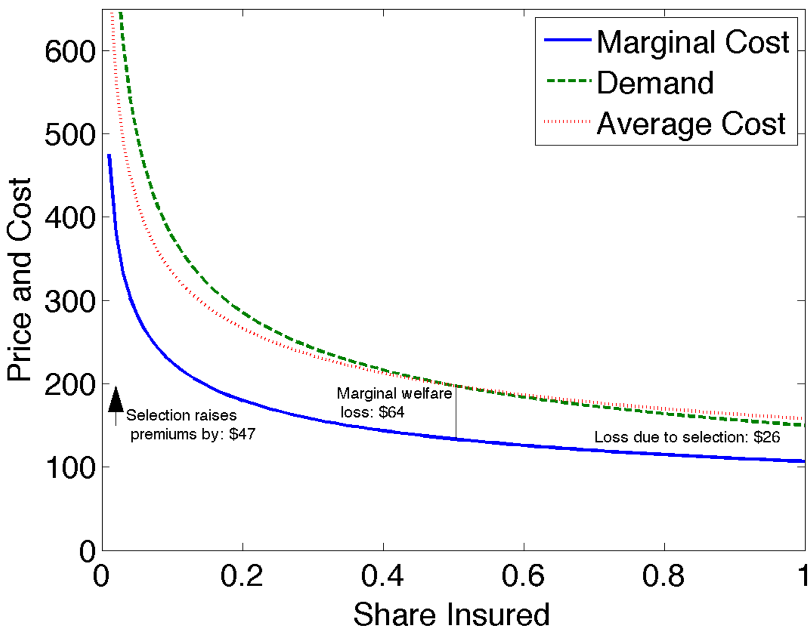

Figure VII: Welfare estimates with constant elasticity functions. These numbers come from combining the demand and cost estimates, and imposing a competitive break-even equilibrium at the observed quantities. Because demand is always above marginal costs, this suggests the efficient allocation is full coverage. The calculated welfare loss due to selection is $25 per person per month. I take this functional form as a lower bound for welfare estimates.

Table I: Non-group Take Up Rates in Colorado, 2014

	Direct Purchase 2014	Uninsured 2014	Take-up Rate, %
All Ages	800	603	57.0
<18	143	105	57.6
18-24	58	113	33.9
25-34	68	120	36.2
35-44	71	94	43.0
45-54	126	83	60.3
55-64	118	78	60.2

Numbers in thousands. *Source:* CPS ASEC Supplement, 2015. Data can be obtained via the CPS Table Creator available at www.census.gov.

Table II: Balance in Demographic Characteristics

	Leaveout Cost	Premium
Age 18-34, %	0.607***	0.611***
	(0.000)	(0.001)
Median Age	-0.427***	-0.500***
	(0.001)	(0.000)
Less than H.S., %	0.337*	0.251
	(0.059)	(0.149)
Bachelor's, %	0.033	0.151
	(0.856)	(0.443)
Labor Force Participation Rate	-0.335	-0.107
	(0.163)	(0.590)
Unemployment, %	0.000793	-0.000582
	(0.489)	(0.529)
Per Capita Income	-164.4	-56.62
	(0.235)	(0.698)
Median Household Income	-124.3	-109.0
	(0.685)	(0.722)
Native Born, %	-0.419***	-0.510***
	(0.000)	(0.000)
Hispanic, %	0.00803***	0.00927***
	(0.001)	(0.000)
2013 % Uninsured 35-64	0.000075	-0.00105
	(0.969)	(0.678)

Notes: Results of regression of leaveout cost instrument on tract-level demographic characteristics, including the zip code pair fixed effect. Each row represents a regression, with the first column listing the dependent variable and the second column the coefficient on leave-out costs. p-values are given in parentheses. Each tract is weighted by the % of the total population of the zip code represented by the tract.

Table III: External Validity of Boundary Sample

	Full Sample	Boundary Sample
Female, %	0.53	0.51
	(0.50)	(0.50)
Avg Age	35.61	36.42
	(18.88)	(19.18)
Population 18-34, %	23.47	20.45
	(8.83)	(5.58)
Population 35-64, %	41.63	43.63
	(5.74)	(5.95)
Median Age	38.25	40.28
	(5.50)	(5.86)
Less than H.S., %	6.77	10.69
	(6.16)	(6.64)
Bachelor's, %	27.81	22.54
	(9.53)	(10.43)
Labor Force Participation	69.89	67.17
	(6.62)	(10.59)
Per Capita Income	36308.30	30785.89
	(10212.73)	(9948.30)
Native Born Pop, %	91.89	91.24
	(5.04)	(5.49)
N	248273	15710

Notes: Comparison of boundary sample to entire population insured in non-group market. The age and gender variables are from the medical claims database. Other variables are Zip code level demographics from the 5-year ACS (2010-2014).

Table IV: Evidence of Selection: IV

	(1)	(2)	(3)	(4)	(5)
Premium Measure:	AvgSilver	BCBS	RMHP	NHV	2LCS
OLS:					
Premium % Increase	0.683*	0.881***	0.695**	0.989**	0.612***
	(0.347)	(0.291)	(0.338)	(0.422)	(0.173)
IV:					
Premium % Increase	0.747**	0.811**	0.732**	1.102**	0.618***
	(0.346)	(0.333)	(0.366)	(0.440)	(0.226)
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	9735	9735	9735	9735	9735

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

Notes: Results from regression of log residualized annual medical expenditures on % increase in premiums. The second row uses leave-out costs as an instrument for premiums. The columns represent different measures of changing premiums when stepping across the boundary. Column (1) is the change in the premium for the average silver plan. Columns (2)-(4) use the change in premium for the exact same silver plan offered by three statewide insurers (BlueCross BlueShield, Rocky Mountain Health Plan, New Health Ventures). Column (5) uses the change in premium for the second lowest cost silver plan (2LCS). The results generally imply that a 1% increase in the insurance premiums in an area increases the annual medical expenditures of the insured population by about 0.8%. Standard errors are clustered at the zip code pair level.

Table V: Placebo IV Regressions

	2013 Non-group		2014 Group		2014 Medicaid	
Premium Measure:	AvgSilver	BCBS	AvgSilver	BCBS	AvgSilver	BCBS
Premium % Increase	0.274	0.289	-0.0301	-0.0347	-0.189	-0.220
	(0.367)	(0.387)	(0.214)	(0.248)	(0.239)	(0.289)
Zip Pair FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10429	10429	38516	38516	68007	68007

Standard errors corrected for clustering at zip-pair level in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Results from placebo test of regression of log annual medical expenditures on % increase in premiums, using leave-out costs as an instrument for premiums. The placebo test runs the same regression using the 2014 Medicaid enrollees, whose behavior should not be affected by rating area boundaries. Standard errors are clustered at the zip code pair level.

Table VI: Breakdown by Age

	(1)
Premium % Increase	-0.0534
	(0.395)
Premium % Increase × Age in 2014	0.0447**
	(0.0213)
Premium % Increase × Age in 2014 × Age in 2014	- 0.000618***
	(0.000296)
Zip Pair FE	Yes
Observations	9730

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Results from median regression of log annual medical expenditures on % increase in Average Silver premiums, broken down by age. Adverse selection appears to be driven primarily by individuals below age 55, and particularly for those ages 35-44. Standard errors are clustered at the zip code pair level.

Table VII: Indicator for Chronic Condition

	2014 Non-group		2013 Non-group (Placebo)	
	AvgSilver	BCBS	AvgSilver	BCBS
Premium % Increase	0.105***	0.0779**	-0.0104	-0.00209
	(0.0315)	(0.0362)	(0.0408)	(0.0455)
Zip Pair FE	Yes	Yes	Yes	Yes
Observations	9736	9736	10430	10430

Standard errors corrected for clustering at zip-pair level in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: This table shows the results using an indicator for the presence of chronic conditions as the dependent variables, as well as placebo regressions. The coefficient of 0.105 in the first column indicates that for each 10% increase in the monthly premium, the probability that an insured individual has a chronic condition increases by 1.05 percentage points.

Table VIII: Demand elasticity estimates

	(1)	(2)
Premium % Increase	-2.572***	-4.973***
	(0.839)	(1.557)
Premium % Increase × Agebin		0.960*
		(0.526)
Group FE	Yes	Yes
Observations	172	172
Margins		
Age 25-34		-4.013***
		[-6.280,-1.745]
Age 35-44		-3.052***
		[-4.778,-1.327]
Age 45-54		-2.092**
		[-3.818,-0.367]
Age 55-64		-1.132
		[-3.400,1.135]

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Results from a zip code-level regression of the log of each zip code's estimated insured rate on the premium measures. The negative coefficient indicates the expected downward sloping demand curve, that is, there is a relatively higher share of uninsured individuals when the premium is higher. The pattern across age groups indicates that the younger consumers are more price sensitive.

Table IX: Optimal subsidy policy by age, no restrictions on age pricing

	Subsidy, $	Share	Benefit	Cost	Net	Ratio (Benefit/Cost)
Optimal Subsidy						
25-34	13.89	0.97	5.17	3.33	1.83	1.55
35-44	15.06	0.99	5.89	3.25	2.64	1.81
45-54	13.16	0.84	4.11	3.04	1.06	1.35
55-64	0.00	0.60	0.00	0.00	0.00	-
Total	9.63	0.84	15.16	9.63	5.53	1.57
Blanket Subsidy						
25-34	11.99	0.88	4.25	2.63	1.62	1.61
35-44	11.99	0.88	4.40	2.29	2.11	1.92
45-54	11.99	0.82	3.72	2.70	1.01	1.38
55-64	11.99	0.65	1.23	2.01	-0.78	0.61
Total	9.63	0.80	13.59	9.63	3.96	1.41

Notes: This table shows the effects of age-targeted premium subsidies. The top panel shows the results of the optimal subsidy for each age group. The bottom panel shows the effects of a policy of spending the same amount of money as the age-targeted subsidy policy, but only using a blanket subsidy. The first column shows the monthly per person subsidy amount. Share indicates the share of the age group that is covered under the optimal subsidy amount. The benefit, cost, and net amounts indicate the per person welfare quantities resulting from the subsidy. A ratio greater than one indicates that the benefits are greater than the costs. These numbers based on allowing flexibility in age pricing. The key takeaway is that spending the same amount of money but without age targeting leads to lower welfare gains and lower coverage levels.

Table X: ER vs. Non-ER spending

	(1) ER	(2) ER	(3) ER	(4) Non-ER	(5) Non-ER	(6) Non-ER
Premium % Increase	-0.405***	0.0535	0.0724	-0.327*	0.765**	0.898***
	(0.140)	(0.276)	(0.267)	(0.166)	(0.295)	(0.305)
Female			-0.0250			0.897***
			(0.0508)			(0.0533)
Age FE	No	No	Yes	No	No	Yes
Zip Pair FE	No	Yes	Yes	No	Yes	Yes
Observations	9736	9736	9735	9736	9736	9735

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Results from regression of log annual medical expenditures on % increase in BCBS premiums, broken down by whether spending occurred due to an emergency room visit. Columns (1)-(3) show that with the zip code group fixed effect, emergency room spending does not change as one moves across the rating area boundary. Columns (4)-(6) show that the adverse selection detected previously is driven by changes in spending in non-emergency room spending. Standard errors are clustered at the zip code pair level.

Online Appendix: Not For Publication

The first two sections of this Appendix provide further details of the ACA, including the Essential Health Benefits (EHB) required for insurance plans, and an explanation of how the premium tax credits work. Section C is the data appendix describing the steps used to determine the sample for estimation. Section D provides the details of the Chronic Condition Indicator tool. Section E includes the theoretical model which guides the empirical specifications and in particular shows why the main specification provides a lower bound of the effect of selection on the extensive margin. Finally, Section F includes the Appendix Tables.

A Essential Health Benefits

The Affordable Care Act's the Essential Health Benefits[25] are:

1. Ambulatory patient services (Outpatient care)
2. Emergency Services (Trips to the emergency room)
3. Hospitalization (Treatment in the hospital for inpatient care)
4. Maternity and newborn care
5. Mental health services and addiction treatment
6. Prescription drugs
7. Rehabilitative services and devices
8. Laboratory services
9. Preventive services, wellness services, and chronic disease treatment
10. Pediatric services

All qualified health plans (QHPs) sold in the individual and small group must cover these ten essential benefits beginning January 1st, 2014. However, the exact scope of services offered can vary. Grandfathered plans are not required to meet these requirements, though they will generally meet some of them.

QHPs must also cover at least 60% of out-of-pocket expenses on average, and must have reasonable annual out-of-pocket maximums. Most common services such as preventative services and wellness visits have no cost sharing. In addition, there are no annual or lifetime limits on Essential Health Benefits.

[25]See "ObamaCare Essential Health Benefits," `http://obamacarefacts.com/essential-health-benefits/`.

B How do premium tax credits work?

The premium tax credits give premium assistance to those earning below 400% of the federal poverty line (FPL) by capping the share of income that an individual or family would have to pay for health insurance. For example, a family of four earning 200% FPL makes $3,925 per month. In the 200% FPL bracket, the ACA requires that the family should spend no more than 6.3% of income on health insurance, or $247 per month. If in this rating area, the second lowest cost silver plan was $400 per month, the family would be eligible for a monthly subsidy of $400-$247 = $153. This $153 could be applied to any plan available in the location.

C Data Appendix

The sample used in estimation includes all individual in Colorado's non-group market in 2014, as well as 2013 for the placebo checks. This data appendix provides a step-by-step description of how the dataset used for estimation is constructed.

I begin by importing all individuals in the insurers' member eligibility files, excluding those on Medicaid, Medicare, or Medigap coverage or those who have have indicated "No" to whether it is the individual's primary insurance coverage. Dental plans are also dropped. In order to isolate individuals in the non-group market in 2013 and 2014, several variables are used because although there is a variable for "market category", there is some inconsistency across payers in how this variable is coded. For some payers, the market category code clearly distinguishes the large group, small group, and non-group markets. For other payers, a combination of the "group policy number" and "coverage type" variables seemed to give a fairly good indication of whether individuals were participating in the non-group market. Having isolated those insured in the non-group market in 2013 and 2014, these individual are then linked to their medical claims for profession, outpatient, inpatient, and pharmacy claims generated during the years 2013 and 2014.

Each individual's location in this sample is identified by a 5-digit zip code of residence. In order to construct pairs of neighboring 5-digit zip codes that could be used as comparison groups, I started with a shapefile of all zip codes in the state, and constructed a matching of each zip with all of its neighboring zip codes. I then kept only matched zip code pairs that were neighbors and also (1) belonged to different rating areas while (2) belonging to the same local medical market. Pairs that shared a border of less than .1 of a mile, or were only neighbors based on a common node (i.e. shared a border length of 0, but intersected at a corner) were excluded. For the primary estimation, zip codes A and B were paired into a "zip code group" if zip A's closest neighbor (in terms of sharing the largest border) was B, and vice versa.

The sample used for estimation includes all individuals determined to be in the non-group market, and who also live in a zip code that was paired with another zip code that belonged to another rating area in 2014. The sample was also restricted to include only individuals below age 65, and only individuals who were insured for at least 9 months of 2014.

For the welfare exercises, I need to calculate an estimate of the distribution of ages in the entire non-group market, including both individuals who are uninsured as well as those who are uninsured. I do this by combining the observed numbers of uninsured individuals in the sample of boundary zip codes with statewide estimates of take-up rates from the Colorado Health Access Survey (CHAS). For example, suppose each individual has been grouped into one of four age bins denoted by i. For each age bin, the observed number of insured individuals is denoted obs_i. The survey data provides an estimate of the take-up rate for bin i, denoted $takeup_i$. Supposing that the total number of individuals in the market from age bin i is $total_i \cdot takeup_i = obs_i$, then the unobserved total number of individuals can be solved for using the two observed quantities by:

$$total_i = \frac{obs_i}{takeup_i}$$

D Chronic Conditions Indictor tool

The chronic condition regressions use the Healthcare Cost and Utilization Project's (HCUP) Chronic Condition Indicator (CCI) tool, which categorizes ICD-9 diagnosis codes as indicators of chronic or non-chronic conditions.

The tool also categorizes the ICD-9 diagnosis codes into one of 18 body system indicators, which are as follows:

1: Infectious and parasitic disease

2: Neoplasms

3: Endocrine, nutritional, and metabolic diseases and immunity disorders

4: Diseases of blood and blood-forming organs

5: Mental disorders

6: Diseases of the nervous system and sense organs

7: Diseases of the circulatory system

8: Diseases of the respiratory system

9: Diseases of the digestive system

10: Diseases of the genitourinary system

11: Complications of pregnancy, childbirth, and the puerperium

12: Diseases of the skin and subcutaneous tissue

13: Diseases of the musculoskeletal system

14: Congenital anomalies

15: Certain conditions originating in the perinatal period

16: Symptoms, signs, and ill-defined conditions

17: Injury and poisoning

18: Factors influencing health status and contact with health services

E Theoretical Underpinning of Empirical Analysis

E.1 Model Setup

This section details the theoretical model underpinning the empirical estimation, an exercise which allows for a precise decomposition of the various margins over which individuals are able to adjust to premiums increasing, which effects are captured in the current analysis, and the direction of the biases caused by effects not explicitly captured in the analysis.

To do so, I first begin by introducing the relevant notation. Each individual receives a draw of private information from a distribution, denoted $\theta_i \sim F(\theta)$. An insurance contract j is a defined by a pair (ϕ_j, p_j), the plan generosity and monthly premium, respectively. For this example, let $j \in \{N, B, S\}$, that is the choice set consists of choosing no insurance, a bronze plan, or a silver plan.

Each i has a valuation v_{ij} for plan j. For example,

$$v_{ij} = \theta_i \phi_j - p_j + \varepsilon_{ij}$$

The heterogeneity in θ_i implies selection. A higher θ_i means an individual has a higher value for plan generosity. If θ_i is an individual's risk, then this model implies adverse selection. However, if θ_i were something like risk preference, such as a measure of risk aversion, then the selection would be advantageous.

Each i chooses how much healthcare to seek, $m_i(\theta_i, \phi_j)$, which is a function of both the individual's private information and their level of coverage. If there is no moral hazard, that means that the healthcare sought does not depend on coverage, and so $m(\theta_i)$. Moral hazard means that the behavior can change with the level of coverage.

The cost to the insurer from offering the policy to i is denoted $c(\phi_j, m(\theta_i, \phi_j))$. In the case without moral hazard, costs are denoted as $c(\phi_j, m(\theta_i))$, which is an increasing function of both arguments. Then the nature of selection depends on $\frac{\partial m(\theta_i, \phi_j)}{\partial \theta_i}$, where if this term is > 0 implying adverse selection, while < 0 implies advantageous selection.

E.2 Use of Premium Variation

To illustrate the usefulness of premium variation, consider a simpler case where there are only two options, N, no insurance, or B, insurance with a bronze plan. We allow for the choice of healthcare expenditures to depend on the coverage level of the plan, so that $m(\theta_i, \phi_N) \neq m(\theta_i, \phi_B)$. Let $I(j)$ denote the population that chose option j.

With medical claims from individuals making both choices, one could compare the average expenditures from each choice, in the spirit of testing for a positive correlation between demand for insurance and expenditures:

$$\mathbb{E}_\theta[m(\theta_i, \phi_B) | i \in I(B)] > \mathbb{E}_\theta[m(\theta_i, \phi_N) | i \in I(N)]$$

However, finding this positive relationship could be due to either moral hazard even with random sorting (i.e. $m(\theta_i, \phi_B) > m(\theta_i, \phi_N)\,\forall i$), or it could be due to selection even if there is no moral hazard (i.e. $m(\theta_i, \phi_B) = m(\theta_i, \phi_N)\,\forall i$, but if $\theta_i > \theta_j \iff v_{iB} > v_{iN}, \implies m(\theta_i, \phi_B) > m(\theta_j, \phi_B)$). It could also be due to a combination of both effects.

Suppose exogenous premium variation is available which provides two populations with an identical distribution of $\theta_i \sim F$, but facing different premiums for choice B. Denote these populations as I^H and I^L for those facing higher and lower premiums, respectively. The difference in the two populations will be how marginal individuals respond to the different premiums. If the first individuals to drop out of the insurance market (switch from B to N) are relatively healthy, and thus the market is adversely selected, that means a low $\theta \implies$ low m. An empirical implication of this pattern is that:

$$\mathbb{E}_\theta[m(\theta_i, \phi_B)|i \in I^H(B)] > \mathbb{E}_\theta[m(\theta_i, \phi_B)|i \in I^L(B)]$$

Thus the average costs of the insured populations can be compared as a test for the existence of adverse selection. This is exactly the logic developed in Einav et al. (2010).

E.3 Market with More Than Two Choices

This logic is slightly complicated by the existence of more than two choices, where there is a menu of plan generosity available. In addition, there is a general problem that medical claims datasets will typically not include information on the uninsured. In my particular context, there is an additional problem that the exact plan details are not always available, so that information such as whether the insured individual is in a bronze or silver plan is known only for a subsample of the observations.

Recall that the premium variation provides two populations, I^H and I^L. To match my empirical context, assume here that the high premium side means that all insurance plans (all metal levels) are more expensive for the I^H population than for the I^L population, relative to remaining uninsured (denoted choice N). If the choice of an individual's metal level is not known, one could start by simply comparing the costs to the insurer in each insured population, which would be equivalent to running the test of whether:

$$\begin{aligned}
&\mathbb{E}_\theta[c(\phi_S, m(\theta_i, \phi_S))|i \in I^H(S)]\cdot\frac{I^H(S)}{I^H(S)+I^H(B)} + \mathbb{E}_\theta[c(\phi_B, m(\theta_i, \phi_B))|i \in I^H(B)]\cdot\frac{I^H(B)}{I^H(S)+I^H(B)} \\
&> \mathbb{E}_\theta[c(\phi_S, m(\theta_i, \phi_S))|i \in I^L(S)]\cdot\frac{I^L(S)}{I^L(S)+I^L(B)} + \mathbb{E}_\theta[c(\phi_B, m(\theta_i, \phi_B))|i \in I^L(B)]\cdot\frac{I^L(B)}{I^L(S)+I^L(B)}
\end{aligned}$$

However, there are three effects that can cause this inequality to hold:

- (i) Selection on Extensive Margin (main effect of interest)
- (iia) Selection on Intensive Margin
- (iib) Moral Hazard on Intensive Margin

Note that by making this comparison only on the insured sample, moral hazard on the extensive margin, that is, individual behavior changes when uninsured compared to being insured, are controlled for. There may exist such an effect, but the estimated effects from the other channels are estimated taking moral hazard on the extensive margin as given.

The effect (iia) occurs because even without moral hazard or adverse selection, plan generosity changes. That is, even for the same level of healthcare utilized m_i, $c(\phi_B, m_i) < c(\phi_S, m_i)$ because the silver plan is more generous than the bronze plan, and this will incur higher claims for the insurer. Even without moral hazard, this (iia) effect will lead to an underestimation of the selection effect when comparing expenditures across the H and L populations.

Rather than estimating the difference in average costs incurred by the plan, using the total medical expenditure $m(\theta_i)$ will address the effect of (iia) by controlling for plan generosity. Indeed, even in the presence of adverse selection, estimating the average costs incurred by the plan can lead to no effect because of the countering effect of changing plan generosity. For example, this can occur if the average costs to the plan of both the silver and bronze plans increase, but as a greater share of individuals are in the bronze plan, the average cost to the plan of the entire population can be flat or even decrease.

Thus, in my empirical estimation, I use the total annual medical expenditure of an individual, which corresponds to $m(\phi_j, \theta_i)$ in this model, because it controls for plan generosity and thus addresses the effect (iia). This corresponds then to testing for the following relationship:

$$\begin{aligned}
&\mathbb{E}_\theta[m(\theta_i, \phi_S)|i \in I^H(S)] \cdot \frac{I^H(S)}{I^H(S) + I^H(B)} + \mathbb{E}_\theta[m(\theta_i, \phi_B)|i \in I^H(B)] \cdot \frac{I^H(B)}{I^H(S) + I^H(B)} \\
&\qquad > \mathbb{E}_\theta[m(\theta_i, \phi_S)|i \in I^L(S)] \cdot \frac{I^L(S)}{I^L(S) + I^L(B)} + \mathbb{E}_\theta[m(\theta_i, \phi_B)|i \in I^L(B)] \cdot \frac{I^L(B)}{I^L(S) + I^L(B)}
\end{aligned}$$

Though this test controls for moral hazard on the extensive margin and selection on the intensive margin, there remains the effect of selection on the extensive margin (the effect of interest), but this could be biased by the effect of moral hazard on the intensive margin (effect (iib)).

However, theory predicts that moral hazard is not symmetric, but rather the logic of moral hazard implies that utilization should not increase as plan generosity decreases. That is, $m(\theta_i, \phi_B) \leq m(\theta_i, \phi_S)\,\forall i$. This allows for a sign to be placed on the bias from moral hazard on the intensive margin, and it can be shown that this will lead to an underestimate of the selection effect (i).

If there were exogenous premium variation available across each plan available, one could estimate the costs of the switchers between each level of plan generosity, and quantify the relative effects of (i) and (iib). However, the premium variation I have available in this context makes all plans more expensive relative to the outside option of remaining uninsured, N. Thus, my analysis focuses on this extensive margin. The point of this section is that although this premium variation from rating area boundaries does not lend itself to quantifying the effects of moral hazard on the intensive margin, to the extent that it exists, it should only lead to an underestimate of the main effect of interest: selection on the extensive margin.

F Appendix Tables

Table A1: Healthcare Utilization in 2014 by Market Segment

	Newly Insured 2014	Previously Enrolled	Employer-Based
Inpatient Admissions (claims per 1,000)	51	25	41
Outpatient Visit Rates (claims per 1,000)	624	545	560
Professional Medical Services (claims per 1,000)	7807	5103	7282
Pharmacy (claims per 1,000)	9352	5349	9036

Table A2: Healthcare Expenditures of New Enrollees

	(1)	(2)	(3)
Previously Enrolled	-1077.4***	-807.6***	-829.3***
	(84.83)	(83.59)	(83.12)
Employer-Based Group	-1241.5***	-919.5***	-870.3***
	(89.80)	(85.13)	(79.32)
Female		64.72***	66.33***
		(24.88)	(24.80)
Constant	3521.4***	6093.7***	5952.6***
	(91.87)	(256.0)	(251.4)
Age FE	No	Yes	Yes
Zip3 FE	No	No	Yes
Observations	1377072	1377034	1377034

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

This table compares the total annual medical expenditures in 2014 of new enrollees in the non-group market (omitted category) to those who were previously insured and those insured in the employer-sponsored group market. The results show that, even when controlling for age and geography, newly insured individuals spent on average $830 more compared to those who were previously insured in the non-group market, and $870 more than individuals in the group market.

Table A3: Average annual spending, by plan generosity

	mean	sd	p25	p50	p75
Platinum	15386.94	68386.93	415.37	1819.54	8588.33
Gold	5998.29	29711.6	350.06	978.13	2970.14
Silver	4852.28	23305.12	306.61	811.64	2410.25
Bronze	3225.78	16245.04	210.63	557.46	1475.1
Total	4716.4	24664.98	273.27	726.05	2164.83

This table shows the average annual medical expenditures for individuals in the non-group market in 2014, by metal level. This positive correlation between plan generosity and spending is in the spirit of Chiappori and Salanié's test for asymmetric information. These patterns, however, cannot disentangle moral hazard from adverse selection.

Table A4: Naive regression of costs on premiums

	(1)	(2)	(3)	(4)
Ln(Premium)	0.245	1.511***	0.589**	0.511***
	(0.195)	(0.0481)	(0.221)	(0.187)
County FE	No	Yes	No	No
HRR FE	No	No	Yes	No
HSA FE	No	No	No	Yes
Age FE	Yes	Yes	Yes	Yes
Female FE	Yes	Yes	Yes	Yes
Metal Level FE	Yes	Yes	Yes	Yes
Observations	36668	36668	36668	36668

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

This table shows the cross sectional regression of log annual spending on premiums, without any research design. Due to the endogenous setting of premiums by insurers, these estimates should not be taken as evidence of selection. The sample includes all individuals insured for at least 9 months in 2014, and for which the plan metal level could be identified.

Table A5: Balance in Choice Set Across Boundary

	Diff. # metal levels available	# Zip pairs	Diff. # insurers available	# Zip pairs
	0	16	0	8
	1	16	1	20
	2	0	2	4
Total		32		32

This table shows the difference in the choice set within the 32 across-boundary pairs of zip codes matched using the HRR criteria for local medical market. The number of metal levels change across some boundaries because although Bronze, Silver, and Gold plans are available statewide, there are some areas in which no Platinum level plan is available. This is unlikely to affect selection on the extensive margin, however. Also, because insurers make county level entry decisions, there are some cases where the number of insurers offering plans changes across the rating area boundary. However, everywhere has at least 4 insurers operating with a fairly large menu of plans.

Table A6: Descriptives of Chronic conditions

VARIABLES	(1) N	(2) mean	(3) sd
Infectious diseases	372,699	0.00273	0.0522
Neoplasms	372,699	0.0122	0.110
Endocrine and immunity disorders	372,699	0.109	0.312
Diseases of blood	372,699	0.00483	0.0693
Mental disorders	372,699	0.0598	0.237
Diseases of the nervous system	372,699	0.0640	0.245
Diseases of the circulatory system	372,699	0.0499	0.218
Diseases of the respiratory system	372,699	0.0559	0.230
Diseases of the digestive system	372,699	0.0222	0.147
Diseases of the genitourinary system	372,699	0.0492	0.216
Complications of pregnancy	372,699	0.000875	0.0296
Diseases of the skin	372,699	0.0113	0.106
Diseases of the musculoskeletal system	372,699	0.0540	0.226
Congenital anomalies	372,699	0.00795	0.0888
Ill-defined conditions	372,699	0.00491	0.0699
Injury and poisoning	372,699	0.000445	0.0211
Factors influencing health	372,699	0.00479	0.0691

Table A7: Dep var: 0, 1, or 2+ Chronic Conditions

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
chron_2more					
Premium % Increase	0.434***	0.337**	0.293*	0.299	0.259***
	(0.162)	(0.169)	(0.167)	(0.246)	(0.0963)
Female	0.271***	0.271***	0.272***	0.271***	0.271***
	(0.0339)	(0.0339)	(0.0340)	(0.0340)	(0.0338)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	9735	9735	9735	9735	9735

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Ordered logit of 0, 1, or 2+ chronic conditions

Table A8: Dep var: 0, 1, 2, or 3+ Chronic Conditions

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
chron_3more					
Premium % Increase	0.403**	0.310*	0.276*	0.252	0.241***
	(0.160)	(0.161)	(0.159)	(0.244)	(0.0929)
Female	0.272***	0.272***	0.273***	0.272***	0.272***
	(0.0339)	(0.0339)	(0.0339)	(0.0340)	(0.0339)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	9735	9735	9735	9735	9735

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Ordered logit of 0, 1, 2, or 3+ chronic conditions

Table A9: Welfare by age, maintaining ACA restrictions on age pricing

	S Amt	**Share**	**Benefit**	**Cost**	**Net**	**Ratio (Benefit/Cost)**
25-34	27.50	0.97	5.16	6.60	-1.44	0.78
35-44	27.82	0.95	5.35	5.77	-0.42	0.93
45-54	21.90	0.99	7.06	5.99	1.08	1.18
55-64	1.80	0.74	3.40	0.34	3.05	9.88
Total	18.69	0.91	20.97	18.69	2.27	1.12

Notes: This table shows the results of the optimal subsidy for each age group, when requiring ACA age pricing. The pricing restriction blunts the effectiveness of age-targeted subsidies, while also creating spillovers across age groups. This is why the older age groups have very high benefits/costs, as they are receiving almost no subsidy and thus have very small costs. However, due to the spillovers from healthy individuals entering the market in other age groups, they still experience a benefit as market-wide costs (and thus premiums) fall.

Table A10: Welfare by age, no restrictions on age pricing, CES form

	S Amt	Share	Benefit	Cost	Net	Ratio (Benefit/Cost)
25-34	13.89	0.97	5.17	3.33	1.83	1.55
35-44	15.06	0.99	5.89	3.25	2.64	1.81
45-54	13.16	0.84	4.11	3.04	1.06	1.35
55-64	0.00	0.60	0.00	0.00	0.00	-
Total	9.63	0.84	15.16	9.63	5.53	1.57

Notes: This table shows the results of the optimal subsidy for each age group, using the CES functional form. The first column shows the optimal monthly per person subsidy amount across age groups. Share indicates the share of the age group that is covered under the optimal subsidy amount. The benefit, cost, and net amounts indicate the per person welfare quantities resulting from the subsidy. A ratio greater than one indicates that the benefits are greater than the costs. These numbers based on allowing flexibility in age pricing.

Table A11: Welfare by age, no restrictions on age pricing, CES form

	S Amt	Share	Benefit	Cost	Net	Ratio (Benefit/Cost)
25-34	11.99	0.88	4.25	2.63	1.62	1.61
35-44	11.99	0.88	4.40	2.29	2.11	1.92
45-54	11.99	0.82	3.72	2.70	1.01	1.38
55-64	11.99	0.65	1.23	2.01	-0.78	0.61
Total	9.63	0.80	13.59	9.63	3.96	1.41

Notes: This table shows the results of spending the same amount of money as the age-targeted subsidy policy, but only using a blanket subsidy, using the CES functional form. The first column shows the monthly per person subsidy amount. Share indicates the share of the age group that is covered under the optimal subsidy amount. The benefit, cost, and net amounts indicate the per person welfare quantities resulting from the subsidy. A ratio greater than one indicates that the benefits are greater than the costs. These numbers based on allowing flexibility in age pricing. The key takeaway is that spending the same amount of money but without age targeting leads to lower welfare gains and lower coverage levels.

Table A12: Welfare by age, maintaining ACA restrictions on age pricing

	S Amt	Share	Benefit	Cost	Net	Ratio (Benefit/Cost)
25-34	9.05	0.80	1.88	1.79	0.09	1.05
35-44	18.66	1.00	3.14	4.06	-0.93	0.77
45-54	10.64	0.87	3.58	2.55	1.03	1.41
55-64	0.00	0.68	1.88	0.00	1.88	-
Total	8.40	0.83	10.48	8.40	2.08	1.25

Notes: This table shows the results of the optimal subsidy for each age group, when requiring ACA age pricing, using the CES functional form. The pricing restriction blunts the effectiveness of age-targeted subsidies, while also creating spillovers across age groups. This is why the older age groups have very high benefits/costs, as they are receiving almost no subsidy and thus have very small costs. However, due to the spillovers from healthy individuals entering the market in other age groups, they still experience a benefit as market-wide costs (and thus premiums) fall.

Table A13: Robustness to Dropping Ambiguous Zip Codes: OLS

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.764**	0.959***	0.811**	1.090**	0.593***
	(0.342)	(0.284)	(0.338)	(0.415)	(0.182)
Female	0.913***	0.913***	0.913***	0.913***	0.912***
	(0.0729)	(0.0730)	(0.0729)	(0.0729)	(0.0729)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	6922	6922	6922	6922	6922

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Robustness to analysis run after dropping all zip codes that span counties that are part of different rating areas. Standard errors are clustered at the zip code pair level.

Table A14: Robustness to Dropping Ambiguous Zip Codes: IV

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.802**	0.880***	0.793**	1.182***	0.591***
	(0.336)	(0.319)	(0.358)	(0.416)	(0.209)
Female	0.913***	0.913***	0.913***	0.913***	0.912***
	(0.0712)	(0.0712)	(0.0712)	(0.0712)	(0.0710)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	6922	6922	6922	6922	6922

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Robustness to analysis run after dropping all zip codes that span counties that are part of different rating areas. Standard errors are clustered at the zip code pair level.

Table A15: Robustness to Using HSA Definition: OLS

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.598	0.762*	0.513	0.990**	0.525**
	(0.508)	(0.379)	(0.567)	(0.443)	(0.240)
Female	0.765***	0.766***	0.766***	0.765***	0.764***
	(0.0643)	(0.0644)	(0.0643)	(0.0645)	(0.0651)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	6342	6342	6342	6342	6342

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Robustness to using Hospital Service Areas as medical market definition. Standard errors are clustered at the zip code pair level.

Table A16: Robustness to Using HSA Definition: IV

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.775*	0.784**	0.817*	0.953**	0.602**
	(0.416)	(0.366)	(0.493)	(0.450)	(0.252)
Female	0.765***	0.766***	0.766***	0.765***	0.763***
	(0.0620)	(0.0622)	(0.0618)	(0.0623)	(0.0628)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	6342	6342	6342	6342	6342

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Robustness to using Hospital Service Areas as medical market definition. Standard errors are clustered at the zip code pair level.

Table A17: Zip Code Groups: OLS

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.480	0.638*	0.640*	0.665	0.470**
	(0.358)	(0.343)	(0.317)	(0.482)	(0.213)
Female	0.786***	0.786***	0.786***	0.786***	0.785***
	(0.0449)	(0.0449)	(0.0448)	(0.0449)	(0.0451)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	16818	16818	16818	16818	16818

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: In these specifications, rather than restricting zip codes to be matched into pairs, I allow for groups of more than two zip codes to be matched together. This yields a larger sample size. However, here the coefficients are smaller than before, and are much less statistically significant. Standard errors are clustered at the zip code pair level.

Table A18: Zip Code Groups: IV

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.424	0.476	0.441	0.673	0.351
	(0.369)	(0.395)	(0.381)	(0.522)	(0.276)
Female	0.786***	0.786***	0.786***	0.786***	0.785***
	(0.0440)	(0.0441)	(0.0440)	(0.0441)	(0.0443)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	16818	16818	16818	16818	16818

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: In these specifications, rather than restricting zip codes to be matched into pairs, I allow for groups of more than two zip codes to be matched together. This yields a larger sample size. However, here the coefficients are smaller than before, and are much less statistically significant. Standard errors are clustered at the zip code pair level.

Table A19: Results for Bronze Plans

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	3.454***	3.453***	3.316***	4.752***	2.254***
	(0.803)	(0.887)	(1.060)	(1.027)	(0.558)
Female	0.476***	0.460**	0.474***	0.468**	0.472**
	(0.172)	(0.172)	(0.171)	(0.172)	(0.175)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	721	721	721	721	721

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Results from regression of log annual medical expenditures on % increase in premiums, but only for individuals in Bronze metal level plans. Standard errors are clustered at the zip code pair level.

Table A20: Results for Silver Plans

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.570	0.830	1.181**	0.754	0.604
	(0.668)	(0.849)	(0.567)	(0.953)	(0.435)
Female	0.524**	0.525**	0.528**	0.522**	0.525**
	(0.207)	(0.208)	(0.209)	(0.208)	(0.209)
Age FE	Yes	Yes	Yes	Yes	Yes
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	732	732	732	732	732

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Results from regression of log annual medical expenditures on % increase in premiums, but only for individuals in Silver metal level plans. Standard errors are clustered at the zip code pair level.

Table A21: Zip Pairs with HRR criteria: OLS

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.606*	0.775***	0.650*	0.893**	0.579***
	(0.328)	(0.282)	(0.320)	(0.402)	(0.156)
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	9736	9736	9736	9736	9736

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Notes: Clustered standard errors.

Table A22: Zip Pairs with HRR criteria: IV

	(1) AvgSilver	(2) BCBS	(3) RMHP	(4) NHV	(5) 2LCS
Premium % Increase	0.634*	0.689**	0.621*	0.936**	0.524**
	(0.333)	(0.327)	(0.345)	(0.436)	(0.220)
Zip Pair FE	Yes	Yes	Yes	Yes	Yes
Observations	9736	9736	9736	9736	9736

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

Notes: Clustered standard errors.