The Choice Environment and Diminishing Returns of Choice on the Health Insurance Exchanges

Diana Christine Poehler

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Edwin Wong, Chair

Paul Fishman

Kwun C.G. Chan

Program Authorized to Offer Degree:

Health Systems and Population Health.

University of Washington

Abstract

The Choice Environment and Diminishing Returns of Choice on the Health Insurance Exchanges

Diana Poehler

Chair of the Supervisory Committee:
Edwin Wong, PhD
Department of Health Systems and Population Health

The Affordable Care Act (ACA) Health Insurance Exchanges (HIXs) were designed to encourage price competition in the individual and small group market by providing a platform where enrollees could purchase and compare across numerous health plan options. However, the average number of plan options available to an HIX enrollee quadrupled from 26 to 108 plans between 2019 and 2022. While the extant literature suggests that having an assortment of plan options can benefit enrollees, too many options may be detrimental to enrollee well-being and have cascading effects on market competition. Little is known about the maximum number of choices to offer on the HIXs. In this dissertation, we explore the changing dynamics of health plan menus on the HIXs and identify menu sizes where additional plan options cease to benefit enrollees. Because plan standardization was implemented for a subset of health plans on the

HIXs in plan year 2023, we explore these menu sizes both with and without standardization. We found significant plan churn in the HIXs as few plans were offered for more than two consecutive years and the overall number and characteristics of HIX plans varied across time. We also found that enrollees stopped valuing new plans after 80 plans under no standardization and after 20 plans under some - or complete - standardization. These estimated maximum menu sizes are smaller than the average menu size offered on HIXs in 2022 suggesting that HIXs may be able to reduce the number of health plan options available to enrollees without negatively impacting enrollee well-being. Continuing to expand this knowledge base may guide policymakers as they work to improve competition and enrollee choice on the HIXs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my committee for their guidance and support as I developed this dissertation, including my chair, Edwin Wong, who provided me with support as I developed this dissertation from a seedling of an idea and facilitated the purchase and acquisition of my dissertation data, along with Paul Fishman and Gary Chan, who provided me with invaluable feedback on each manuscript. I also sincerely appreciate the support of Ryan Hansen, who served as my Graduate School Representative. I really appreciate your time and mentorship.

Matthew Poehler, completing this doctoral degree has been a whirlwind of a journey and I am so thankful to have experienced this journey with you. Your endless support through late nights and stressful times grounded me and kept me sane. You were the rock that held our family together and words cannot fully describe how much I appreciate you.

To my friends and family: I could not have completed this dissertation without your love, kindness, patience, and support. Mom, Dad, Karen, Mickey, and Doug – thank you so much for all your help; your support meant the world to me. To late nights, endless giggles, and sanity-saving work dates, Diana Naranjo, you helped me survive this journey and I am so lucky to have met you and be your friend. Beth Devine, thank you so much for your mentorship and friendship.

Thank you to Ben Allaire, Amanda Honeycutt, and my colleagues at RTI International, for helping foster and guide my love of research after my undergraduate degree and providing me with the support to begin – and end – my doctoral journey with RTI.

And finally, thank you to my HSERV cohort and to the UW family housing community for the incredible camaraderie and experiences.

DEDICATION

To Sophia and Abigail.

You are my whole heart, and I am so lucky to be your mother.

Chapter 1. INTRODUCTION

Insurance is a central component of medical care financing in the United States health care system and protects individuals and their families from unexpected, and often catastrophic, medical care costs.[1] Over 90% of individuals in the United States are covered by some type of health insurance, including 49% in employer-sponsored insurance, 37% in government insurance programs like Medicare or Medicaid, and 6% in non-group health insurance.[2] However, individuals who shop for health insurance face a complex decision with large health and financial ramifications.[3-6] The extant literature suggests that individuals may not always be choosing the plan that best meets their needs.[4, 7, 8] Improving each individuals' health plan choice may have cascading impacts for both their family's health and out-of-pocket medical spending as well as have market-wide impacts such as increasing insurer competition, increasing the generosity of plan cost-sharing arrangements, and decreasing plan premiums.[9]

1.1 Problems of Choice

1.1.1 *At a Glance*

Many classical economic theories are built on the concept of rational choice – in market settings, individuals are self-interested and make choices that optimize their well-being.[10] Individuals are thought to have heterogenous tastes and therefore benefit from having an assortment of options.[10, 11] A larger assortment size increases the likelihood that individuals will find an option that matches their preferences.[11]

However, these benefits may be offset as assortment sizes become increasingly large because individuals must process additional information when making their choice.[10, 12, 13] With too many options – referred to as choice overload - individuals may use rules of thumb to make their choice, such as only comparing across a few attributes of each option or defaulting to an option they selected in the past.[14-18] This in turn can be detrimental with individuals having lower satisfaction or even regretting their choice. It can also lead individuals to defer making a choice altogether.[17] There are many factors that can contribute to choice overload beyond just the number of options, including whether the options have many attributes that the individual must consider; if the decision problem is complex with large ramifications; or if the individual is uncertain about their preferences.[10] The problems of choice are pervasive across consumer settings, including jams, electronics, and mutual funds as well as across health insurance markets.[5, 6, 9, 16, 19-21]

1.1.2 *Within Health Insurance*

Public and private health insurance markets in the United States often allow enrollees to choose from a variety of plan options – often referred to as a health plan menu. While most individuals enrolled in employer-sponsored insurance are offered a limited number of plan options, the opposite is true for other health insurance markets.[17, 22] On average, Medicare enrollees choose from anywhere between 20 to 40 plan options and enrollees on the Affordable Care Act's Individual Health Insurance Exchanges (HIXs) choose from around 100 plan options.[17, 22] Even some enrollees under employer-sponsored insurance have close to 50 plan options.[6] Current evidence suggests that enrollees have a difficult time making health insurance choices and often experience anxiety, confusion, and lack of confidence in their choices during open enrollment,[23, 24] and this can be exasperated by a large or complex menu of health insurance options.[17]

The extant literature on health plan menu sizes and choice overload is sparse, but prior work has found that more than 30 health plan options decreases enrollment in Medicare advantage.[25] Further work in employer-sponsored insurance found that, with 48 plan options that differed solely in cost-sharing arrangements, more than half of employees selected a plan that was financially dominated – in other words, more expensive with equal care.[6]

Though few studies explore the optimal number of plans to offer, several studies focus on behavioral outcomes of choice overload, such as persistently selecting a default health plan, referred to as inertia, or deferring the choice of insurance. Six studies evaluated the impact of inertia in Medicare Part D and found that it leads to enrollees "leaving money on the table", but also that insurers knowingly use enrollee inertia to their advantage when pricing Part D plans.[4, 26-30] Further, an excessive number of plan options can reduce the likelihood that enrollees choose to enter the insurance market, leaving the enrollee at risk for high medical costs. Medicaid and private health insurance have been less studied in the literature. Only one study evaluated enrollee inertia in Medicaid and found substantial inertia.[31] Three studies have evaluated enrollee inertia in employer-sponsored insurance and found that enrollees may be overpaying for their insurance needs and preferences by up to $2,000 in out-of-pocket costs. [32-34] In contrast, two studies of Medicare Part D refute enrollee inertia and suggest enrollees learn from their health plan choices overtime in order to make better future choices. [20, 35] In addition, other studies have suggested other causes of inertia in health insurance beyond just choice overload, such as hassle costs from switching plans and inattention to changes in plan benefits overtime.[30, 36]

1.2 Introduction to Health Insurance Exchanges

1.2.1 *Background*

Established by the 2010 Patient Protection and Affordable Care Act (ACA), HIXs provide an online platform where individuals and families could shop for non-group health insurance in each US state.[14] Enrollees are able to purchase health insurance from numerous plan options, fostering competition that was expected to help achieve the ACA's aim of increased access to affordable health insurance.[14,37] An assumption in the HIX's development was that enrollees would shop around for the best priced health plan option for their medical care needs each open enrollment period, which would then send appropriate price signals to insurers.[14]

HIXs are designed to be a one-stop shopping experience where individuals can check their eligibility for Medicaid and ACA subsidies and shop across all plans available to them.[38] The ACA subsidies help low- and middle-income individuals purchase insurance on the HIXs by mitigating out-of-pocket medical care costs. Similar to employer-sponsored insurance and Medicare, HIX enrollees are usually permitted to change plans during an annual open enrollment period.[38] Plans are organized within four metal tier levels – Platinum, Gold, Silver, and Bronze. These tier levels are based on the plan's actuarial value and are designed to simplify the shopping experience for enrollees.[38] Online and in-person assistance is also available to enrollees who need additional help enrolling in insurance.[38]

Though HIXs are implemented at the state-level, states can choose to operate a state-based exchange (SBE), where the state operates most of the core functions of the HIX – such as designing the website and determining enrollee eligibility – or choose for the federal government to operate all core functions of the HIX on the state's behalf (referred to as a federally-facilitated marketplace; FFM).[39] In 2023, 21 states operate a SBE while 30 states have a FFM.[40] Individuals and families are generally eligible to enroll on their state's HIX if they are a non-incarcerated United States citizen or national and do not qualify for Medicare.[39]

1.2.2 *Choice and Impacts on HIXs*

The HIXs provide a unique opportunity to explore health plan choice as the average number of health plans available quadrupled between 2019 to 2022, from an average of 25.9 plans offered in 2019 to an average of 107.7 plans in 2022; this compared to an average of 39 Medicare Advantage plans available to Medicare enrollees in 2022.[22, 41] A large and tumultuous health plan menu on HIXs may lead enrollees to use rules of thumb – often referred to as heuristics - by either comparing across only a select few plan characteristics, such as a health plan's premium and deductible, or by defaulting to reenroll in their current health plan without comparison shopping.[4, 7, 32, 33, 37, 42] These heuristics may have cascading effects across the market. Enrollees may overpay in premiums or choose health plans with benefit structures that either do not meet their current health insurance needs or do not give them the best benefit possible from the insurance choices available to them.[7, 33] Systematic and persistent use of heuristics or reenrolling in a default health plan may send inappropriate market signals to insurers which can result in higher premiums on the HIXs.[15] High premiums on the HIXs reduce the affordability of health insurance for HIX enrollees.

This has implications for the United States government, which ends up spending more on premium subsidies, which are benchmarked against the second lowest cost silver plan in the enrollee's county. As premiums increase on the HIXs, the amount of the premium subsidies also increases.[43]

Though literature has explored the relationships between health plan menu size on enrollee well-being and behavioral outcomes of choice overload, this work has focused on Medicare and employer-sponsored insurance. These results may not generalize to HIX enrollees because HIX enrollees represent different age ranges than Medicare enrollees and have more health needs than individuals enrolled under employer-sponsored insurance.[44] Therefore HIX enrollees may have different thresholds for choice overload than the previously studied populations. No study to date has explored the optimal number of options to offer on the HIXs. In the following chapters, we explore plan choice and menu sizes on the HIXs.

1.3 DISSERTATION OBJECTIVES

The objective of this dissertation is to inform these current discussions by exploring the changing dynamics of health plan menus on the HIXs and identifying menu sizes where additional plan options ceases to increase enrollee well-being. We use structural utility models to measure the benefits of insurance, including both financial and non-financial aspects such as protection from financial risk, health benefits (e.g., drug benefits), and availability of desired providers.

Specific Aim 1: Explore the stability in the choice environment faced by HIX enrollees. What percentage of plans on HIXs are offered for two consecutive years? What is the average number of years that a health plan remains on HIXs? Has a change in plan actuarial value, plan premiums, plan tier, or plan type (PPO, EPO, HMO) occurred over time? We used the Health Plan Finder Data (HPF) for years 2014-2019 to provide descriptive statistics on changes in HIXs at the state-level.

Specific Aim 2: Develop a prediction model estimating expected medical expenditures for each Washington Health Benefits Exchange (WAHBE) enrollee. When comparing across various machine learning and parametric models, which type of model best predicts the expected medical expenditures of WAHBE enrollees in a future plan year? Are there differences in performance when considering only high-expenditure enrollees? We developed a novel medical cost prediction model to estimate future medical expenditures for each enrollee. We compared seven machine learning and five parametric prediction models to determine the best prediction model for WAHBE enrollees.

Specific Aim 3: Identify the point where the value of having a larger health plan menu begins to taper off. At which number of plan options does enrollee willingness-to-pay for each additional option stop significantly increasing? Does this vary by penetration of standardized plans in the market? We used a nested logistic model to assess enrollees' plan choices and infer the benefit each plan provided to an enrollee, operationalized by estimating enrollee willingness-to-pay for each plan. We then simulated changes in willingness-to-pay for sets of plans ranging from one plan option to 100 plan options and identified when willingness-to-pay for each option stopped significantly increasing.

1.4 SETTING AND DATA

We used data from Washington State's HIX from 2014 to 2018 to address these research gaps. Washington State operates an SBE, referred to as the Washington Health Benefits Exchange (WAHBE), that became operational in 2014.[45] WAHBE is a public-private partnership created by Washington state legislature and operates distinctly from the Washington state government with a self-sustaining revenue stream stemming from health carrier charges, enrollees premiums, and premium taxes.[45] WAHBE conducts all core functions required of SBEs, including developing and operating its own enrollment website, the WA HealthPlanFinder, and further adheres to all rules and regulations set forth for SBEs by the annual Payment and Parameters notices from the United States Department of Health and Human Services as well as any Washington-specific legislation.[46]

WAHBE enrolls an average of 200,000 enrollees each year across Washington State's 39 counties.[47-53] Mirroring trends on the FFM, the number of health plans and insurance issuers participating in the HIX has fluctuated since WAHBE's inception, ranging from 40 plans in 2019 to 120 plans in 2022.[54, 55] WAHBE began implementing novel policies ahead of the federal requirements, such as plan standardization, capping the number of non-standardized health plan options that insurance issuers could offer, and a public plan options. However, these policies took effect in 2021 and are not captured in our data period and thereby do not impact the generalizability of our results to other SBEs and FFMs.[45]

This dissertation created a novel dataset by linking three, typically disparate, datasets to conduct the analyses. First, we obtained WAHBE administrative data through a public use request.

WAHBE administrative data contains detailed demographic and enrollment information for each individual enrolled on the HIX, including dates of enrollment; residential county and zip code; income and eligibility for premium and cost-sharing reduction (CSR) subsidies; family identifiers to group family members; age, sex, and detailed race that is complete for all enrollees; and choice of health plan for each enrollment period, including the health plan's unique Health Insurance Oversight Identifier (HIOS ID).

Second, we obtained medical and pharmaceutical claims data for all WAHBE enrollees from the Washington All-Payers Claims Database (WA-APCD). The WA-APCD data was established by Washington State Legislature and began data collection in 2014. All private and public health insurance issuers who operate in Washington State must submit claims data for all their Washington enrollees. The Washington Health Care Authority (HCA) oversees implementation of the WA-APCD and any data use requests. The claims data includes medical diagnoses, medical procedures, prescription information, and payment information for inpatient medical care, outpatient or office-based medical care, and pharmaceutical care. The WA-APCD was linked to WAHBE administrative data through cooperation between the Washington HCA and WAHBE. This linked data spans 2014 to 2018.

Third, we obtained detailed information on each health plan offered on HIXs nationwide from the Health Plan Finder (HPF) data. The Centers for Medicaid and Medicare (CMS) produces the HPF data files, which contain detailed information on cost-sharing arrangements, benefit design, service exclusions, provider networks, and cost-sharing variations for every health plan offered on the HIXs across all states.[56] The HPF data are publicly available on the

CMS website and contain the unique HIOS ID for every plan in the data.[57, 58] We joined the HPF data to each enrollee's health plan choice in the linked WAHBE-WA-APCD data using health plan HIOS ID. Because WAHBE administrative data contained zip codes and county of residence, we were also able to identify characteristics of all plans available to each WAHBE enrollee in a given year. To our knowledge, no other studies have combined these three types of data sources.

Though we will only be evaluating one state in our analyses, the results from this dissertation will generalize beyond Washington state. This dissertation is based on choice theory and our results will assess how individuals respond to large menu sizes. Therefore, the results from these analyses can generalize beyond the WAHBE, to any HIX and to any setting where large choice sizes persist.

1.5 DISSERTATION ROADMAP

In chapters two through four, we provide the background, analytic approach, and findings for Aims 1, 2, and 3. Each dissertation aim has associated figures, tables, and appendices. In chapter five, we summarize our results as well as implications for policy and future research from our findings.

Chapter 2. EXPLORING PLAN CHURN AND CHANGING PLAN FEATURES ON US HEALTH INSURANCE MARKETPLACES

2.1 INTRODUCTION

Health Insurance Exchanges (HIXs) were established by the Affordable Care Act (ACA) and allow consumers to shop for non-group health insurance coverage offered within a county. Each state is allowed to choose the manner in which the HIX operates and can choose between establishing and operating a state-based exchange (SBE), allowing the Department of Health and Human Services (HHS) to operate a federally-facilitated exchange (FFE) on behalf of the state, or partnering with HHS to operate a state-based exchange on the federal platform (SBE-FP). HIXs are intended to provide access to affordable health insurance by fostering competition among insurers within each individual market. Competition among insurers was expected to generate lower premiums and more generous cost-sharing arrangements for enrollees.[43]

However, since HIXs became operational in 2014, they have lacked stability in plan pricing and issuer participation.[59, 60] This instability within HIXs may negatively impact HIX enrollees as participants in these markets often struggle to make informed choices when selecting a health insurance plan due to the overall complexity and ongoing change in plan options.[23, 24] In addition, since more than a quarter of all HIX enrollees are automatically reenrolled each year, extensive plan discontinuation on HIXs may result in many enrollees being auto-reenrolled in plans with a different tier, plan type, and even insurer[59, 61]

Previous research on health plans available through HIXs has primarily focused on changes in premiums and insurance issuer participation.[59, 62-66] Changes in plan participation and cost-sharing arrangements have received less attention and existing studies have either focused on a single year or explored changes in only the lowest cost plans.[43, 67] Our research expands the extant literature by exploring trends among all plans offered on the HIXs to better understand program stability.

In this study, we use publicly available data on all health plans offered on the HIXs to track health plan participation and health plan characteristics over time. Our research addresses gaps in the literature by exploring continuity in health plan participation between 2015 and 2019. In addition, we seek to understand whether HIXs have shown stability in the number, type, and characteristics of health plans offered. Understanding plan continuity on HIXs may provide valuable insight for developing future policies, including auto re-enrollment policies and revisions to health plan mapping algorithms.

2.2 METHODS

We conducted an observational study to understand health plan participation and characteristics over time. We obtained information on HIX health plans from the Health Plan Finder (HPF) data for years 2015 to 2019. This dataset is publicly available from the Centers for Medicare and Medicaid Services (CMS) website.[56] The HPF data contains detailed information on plan availability, rates, cost-sharing arrangements, rating areas, and disease management program for plans offered on HIXs. This information is available for health plans offered on-exchange or off-exchange in both the small business health options program (SHOP) and the individual market.

These data are also available for state-based and federally-facilitated HIXs. We limited our analysis to health plans offered on-exchange in the individual market and excluded child-only health plans. All analyses are conducted on plans with standard cost-sharing arrangements.

We measured continuity of plan participation as the percentage of plans in a given year that were also offered in the prior year. We tracked plan participation using the health insurance oversight identifier (HIOS ID). The HIOS ID identifies unique plans offered on HIXs. Per the Centers for Medicare and Medicaid (CMS) guidelines, plans are required to keep the same HIOS ID across open enrollment periods.[68] A plan only receives a new HIOS ID if the changes in cost-sharing arrangements are substantial enough that the plan is classified under a new metal tier.

Our analysis included 3,025 counties across 47 states. We excluded counties in Alaska, South Carolina, Vermont, and Wyoming because the data were missing for at least one year in these states. Since state-based HIXs may differ in rules and regulations for health plans, we calculated plan continuity separately for state-based HIXs, federally-facilitated HIXs, and state-based federal-platform HIXs. In addition, we tracked the number of years that a health plan was consecutively offered in each county and estimated average continuity at the county-level and state-level.

Large plan discontinuity may also be accompanied by changes in the composition of HIX plans. To examine this issue, we conducted descriptive analyses of plan characteristics across years. This included changes in metal tier, plan actuarial value (AV), maximum out-of-pocket costs (MOOP), and deductibles. Each health plan on the HIXs is also assigned a metal tier. The metal tier is a categorical representation of the percentage of health care costs that the plan covers for an average enrollee, referred to as actuarial value.[43] We tracked plan prevalence in each tier across time. We also calculated summary statistics for AVs for each metal tier. For MOOP and deductibles, we calculated a weighted average, accounting for metal tier prevalence since metal tiers affect the cost-sharing arrangements of a plan.

We additionally explored changes in plan type. HIXs offer four main types of health plans including preferred provider organizations (PPO), health maintenance organizations (HMO), exclusive provider organizations (EPO), and point-of-service (POS) plans. These plan types differ in payment structure where HMO and EPO plans are types of managed care plans and PPO and POS plans are types of network-based plans. We tracked the prevalence of each plan type across years.

This research did not include human subjects and was exempt from review by the University of Washington Institutional Review Board.

2.3 Results

The continuity of health plan participation varied by plan year and by marketplace structure (Figure 2.1). Across the US, the percent of plans offered for two or more consecutive years increased over time with 60% of 2015 health plans also offered in the prior year, 63% of 2016 health plans; 67% of 2017 health plans; 64% of 2018 health plans, and 68% of 2019 health plans. These increases were largely driven by fewer plans entering the market each year.

State-based HIXs averaged a higher health plan continuity than both federally-facilitated HIXs and state-based federal-platform HIXs (Figure 2.1). Health plan continuity on state-based HIXS ranged from 66% in 2016 to 87% in 2018 with an average continuity of 75%, while continuity on federally-facilitated HIXs ranged from 57% in 2015 to 63% in 2017 with an average of 60%. State-based federal-platform HIXs largely had the lowest continuity with an average of 57%, ranging from 44% in 2018 to 78% in 2019.

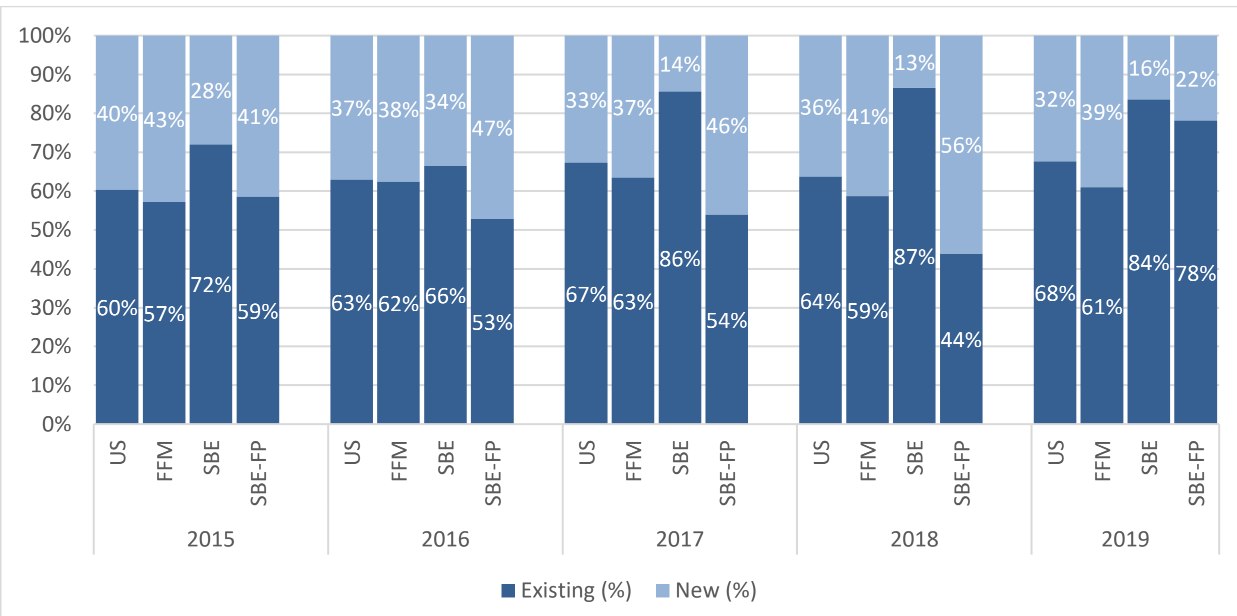

Figure 2.1. Health Plan Continuity by Marketplace Type, Over Time.

NOTES The unit of observation is health plan within a county. Unique plans are identified by HIOS Plan ID. SBE = State-based Exchange, FFM = Federally-facilitated Marketplace, SBE-FP = State-based Exchange, Federal Platform. SOURCE Authors' analysis of Health Plan Finder (HPF) data CMS Center for Consumer Information and Oversight) for 2015-2019 and population density for 2020 from Census Bureau.

Health plan continuity also varied by states and across counties within a state. The average number of years that a health plan was consecutively offered in a state ranged from 1.08 years in Arizona to 2.23 years in Rhode Island (Figure 2). Across all states, the average number of years a health plan was consecutively offered was 1.44 years. Average health plan continuity within state also varied by county. New York had the largest differential between counties, ranging from a continuity of 1.81 years in Cayuga County, NY to 3.82 years in Otsego County, NY.

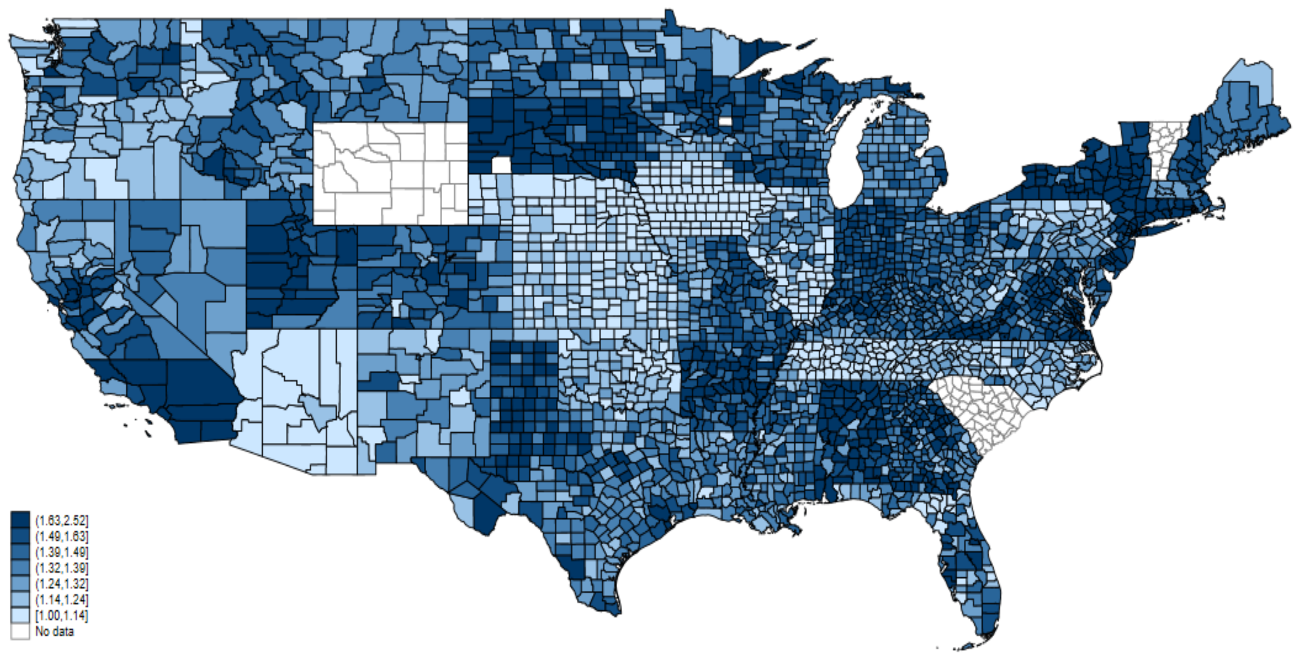

Figure 2.2. Average Number of Consecutive Years a Health Plan is Offered on an HIX.

NOTES The unit of observation is plan-county. Unique plans are identified by HIOS Plan ID within a county. Per Centers for Medicaid and Medicare (CMS) guidelines, plans are required to keep the same HIOS id for multiple years and only receive a new HIOS id if the changes in cost-sharing arrangements are substantial enough that the plan is classified under a new metal tier. Data were not available for South Carolina, Vermont, and Wyoming. SOURCE Authors' analysis of Health Plan Finder (HPF) data (CMS Center for Consumer Information and Oversight) for 2015-2019.

HMOs plans were most common health plan payment structure, accounting for nearly 50% of all health plans on the HIXs, followed by PPO plans, accounting for approximately 25% of plans (Table 2.1). PPO plan offerings spiked in 2015 wherein PPO plans accounted for nearly 40% of all HIX plans, but prevalence decreased by nearly 20 percentage points between 2015 and 2019. PPO plans accounted for less than 19% of plans by 2019. Between 2015 and 2018, POS and EPOs plan offerings remained relatively steady, but EPOs began to trend slightly upwards in 2018 while POS plans began to trend slightly downwards. AVs stayed steady within a tier, with a slight increase in Bronze plan AVs over time. In 2015, the average AV for a Bronze plan was 60.9% while in 2019 the AV increased to 62.1%. MOOP increased steadily over the time frame, ranging from $5,793 in 2015 to $7,160 in 2019. Average deductible also increased over the time frame from $3,461 in 2015 to $4,669 in 2019.

Silver tiered plans were the most prevalent metal tier, followed by Bronze tiered plans, and Gold tiered plans (Table 2.1). Silver tiered plans became more common over time, accounting for 30-40% of all health plans on the HIXs, while Gold tiered plans became less common over time, accounting for between 15-25% of plans. Catastrophic and platinum tiered plans were the least common plan tiers, jointly accounting for less than 15% of all health plans on the HIXs.

Table 2.1. Average Number of Health Plans, Health Plan Continuity, and Health Plan Characteristics by Plan Year Across All States

	2015	2016	2017	2018	2019
Number of Plans					
Average Number of Plans Offered in a County (SD)	37.5 (25.4)	32.9 (22.6)	22.9 (20.4)	13.4 (10.3)	13.7 (11.4)
Median Number of Plans Offered in a County	31.0	29.0	18.0	11.0	10.0
Metal Tier (%)					
Catastrophic	6.6	6.3	7.6	7.1	7.4
Bronze	28.9	30.5	30.9	30.0	31.2
Silver	35.6	37.9	43.2	43.3	40.4
Gold	22.0	21.3	15.4	16.8	18.1
Platinum	6.9	3.9	2.9	2.8	2.9
Actuarial Value in Metal Tier					
Bronze (SD)	60.9 (1.1)	61.5 (0.6)	61.2 (0.8)	61.8 (1.7)	62.1 (1.8)
Silver (SD)	69.9	70.1	70.0	70.0	70.0

	(1.3)	(1.3)	(1.5)	(2.0)	(2.0)
Gold (SD)	79.6 (1.3)	79.7 (1.5)	79.5 (1.4)	79.1 (1.8)	78.8 (1.8)
Platinum (SD)	89.1 (0.9)	89.5 (1.1)	88.8 (1.6)	88.3 (1.2)	88.0 (0.9)
Plan Type (%)					
Preferred Provider Organization Plans (PPO)	37.5	25.5	22.3	18.4	18.8
Health Maintenance Organization Plans (HMO)	46.4	53.4	54.8	57.8	52.9
Exclusive Provider Organization Plans (EPO)	8.6	12.3	13.2	14.3	20.1
Point of Service Plans (POS)	7.5	8.8	9.7	9.5	8.2
Cost-Sharing Arrangements ($)					
Average Deductible (SD)	3,461 (1,905)	3,849 (1,957)	4,510 (1,845)	4,456 (1,986)	4,669 (2,148)
Average MOOP (SD)	5,793 (1,278)	6,203 (1,164)	6,558 (910)	6,973 (910)	7,160 (933)

SOURCE Authors' analysis of Health Plan Finder (HPF) data (CMS Center for Consumer Information and Oversight) for 2015-2019. NOTES SD = Standard Deviation. The unit of analysis for continuity is plan-county. The AV and cost-sharing arrangement estimates are weighted average, using the number of plans in each tier. Catastrophic plans are not required to meet AV targets or submit AV values and are therefore excluded from the AV estimation.

2.4 Discussion

Our study explored the continuity of health plans options across time and found churn in both plan participation and plan characteristics. The majority of health plans averaged less than two years on HIXs. State-based HIXs had higher continuity of health plan participation than federally-facilitated HIXs. Silver tiered plans and plans with an EPO payment structure became more common over time.

Navigating health insurance options can cause enrollees to experience anxiety, confusion, and lack of confidence in their choices.[23, 24] Constant churn in health plan options may reduce enrollees' willingness to shop around for alternative health plan options on HIXs, especially since nearly 25% of HIX enrollees are auto re-enrolled each year.[69]

Because most HIX plans remain on the market for less than two years and plan characteristics are constantly changing, automatic reenrollment policies may result in enrollees re-enrolled in plans whose provider networks, metal tiers, and plan type vary from their original health plan selection. Deductibles increased 35% between 2015 and 2019 while MOOPs increased by 24%. In addition, the prevalence of preferred provider organization plans decreased by nearly half. These changes may lead to surprise medical bills from enrollees whose network arrangements have changed and add additional burden to enrollees who may need to change their care team due to new provider networks. Restructuring how enrollees are auto-reenrolled may improve consumer experiences with HIXs and improve their continuity of care. Future research should explore how enrollees respond to health plan mapping and to changes to auto-reenrollment policies on the HIX.

Our study had several limitations. First, we were unable to access complete health plan data for 2014 and our analysis therefore excludes the first year of the HIXs. Second, our analysis explores changes in plan offerings and does not evaluate enrollment decisions. As a result, we do not observe how changes in the insurance choices affect enrollment decisions nor how enrollment decisions affect insurance options. Changes in continuity and plan characteristics may be driven by low enrollment in certain insurance issuers, plan products, or plan types. We are not able to observe this correlation in our analysis. Third, data for Wyoming, Alaska, Vermont, and South Carolina are missing for at least one year in our data. Our analyses on plan continuity exclude these states. Fourth, we limited our analysis of changes in cost-sharing arrangements to changes in AV, deductibles, tier, and MOOP because these are the most consistently defined cost-sharing variables in the dataset. However, provider network adequacy and other network, service provision, cost-sharing variables may provide valuable insights to changes in health plan generosity. In addition, changes in the risk pool may drive changes in the cost-sharing arrangements of the plan. Because our data do not contain enrollee health characteristics, we are unable to observe correlation between risk pool and changes in cost-sharing arrangements.

Chapter 3. IMPROVING HIX MEDICAL EXPENDITURE PREDICTIONS USING ALL-PAYERS CLAIMS DATABASE

3.1 INTRODUCTION

Health care cost prediction models are used throughout the US health care system to help payers, health systems, and governments efficiently allocate care resources across patient populations.[70-72] These models are used across a variety of settings, including to risk adjust federal programs, which compensate health plans that enroll individuals with expensive medical needs to prevent plans from designing benefits that attract only healthier enrollees, to setting insurance premiums and, under value-based payments, provider reimbursement rates. [71-75] [76] Further, provider groups factor their patient populations' expected medical care costs in planning future care coordination and ensuring that adequate resources are available, particularly for high needs patients.[77, 78] Improving accuracy in health care cost prediction models may have widespread implications for the US healthcare system by improving patient health as health systems better meet the needs of their high risk high needs patients and helping lower medical care costs by ensuring prospective provider and plan payments align with patient risk.

Many existing medical prediction models use linear regression, including the federal risk adjustment models for Medicare and the Health Insurance Exchanges (HIEs).[71, 74, 79, 80] Though linear regression is an unbiased estimator, prediction models using linear regression are easily skewed by high cost cases and may produce inaccurate predictions for subsets of the populations, such as high-cost enrollees, that may distort the allocation of care resources or plan payments.[81, 82] Machine learning relies on fewer modeling assumptions, such as homoscedasticity, and may yield better cost predictions than linear regression in highly skewed or noisy data.[73, 83-87] While several studies have used machine learning algorithms to predict health care costs, these studies are often limited to a specific population, disease, or use only medical and pharmaceutical claims to predict costs and may yield high bias in other settings.[88]

Further, few studies have explored machine learning's predictive power among individuals enrolling in the HIXs. The HIXs provide a platform where individuals can purchase non-group health insurance and provides coverage for nearly 10% of adults in the US.[14,37, 89] Previous research has compared the predictive accuracy of machine learning algorithms with the federal HIX risk adjustment model and, as such, uses a narrow set of predictors that are carefully chosen to prevent insurers from gaming the risk adjustment models.[80] This cautious variable selection process excludes potentially powerful predictors from the machine learning models. A more expansive set of predictors may increase the accuracy of HIX health care cost prediction models. Further, this work uses employer-sponsored claims data that has been sampled to resemble an HIX population which may not accurately reflect HIX medical utilization.[73, 86]

To address these gaps, we compared the performance of seven machine learning and five parametric models in predicting future health care costs among HIX enrollees in Washington State; we explored model prediction error across the full sample of HIX enrollees and how that prediction accuracy varied when considering only high-cost enrollees. Our study is the first to explore the predictive power of machine learning using HIX claims data. Because claims data often lacks detailed information on enrollee and health plan characteristics, we combined an all-payers claims database (APCD) with HIX administrative data and an inventory of HIX health plans to provide a robust profile of enrollee health care utilization and diagnoses merged with demographic and health plan information. Because HIX enrollees have a higher prevalence of chronic conditions and more medical service utilization than other privately insured individuals,[44] accurate medical cost predictions may improve care management and risk adjustment for this understudied population.

3.2 METHODS

3.2.1 *Data Sources*

We used the Washington State All-Payers Claims Database (WA-APCD) from 2014 to 2018 which contains claims data for Medicaid, Medicare, and 30+ commercial health care payers, including all payers who offer plans on the Washington Health Benefits Exchange (WAHBE). WAHBE is an independent organization that was established by Washington State statute and operates Washington State's HIX.[90] The WA-APCD data contains paid amounts for inpatient, outpatient, and prescription drug claims and was created by the Washington State legislature in 2014 and began collecting data the same year.[91]

We linked the WA-APCD to the WAHBE enrollment data for more detailed information on WAHBE enrollees' demographics, dates of enrollment, county of enrollment, selected health plan's Health Insurance Oversight identifier (HIOS Id), and income. WAHBE administrative data is collected for all enrollees purchasing insurance with the HIX. Finally, we linked individuals' HIOS Ids to the Health Plan Finder data, which is available from the Centers for Medicare and Medicaid website and contains information on every health plan offered on the HIXs, including premiums, cost-sharing arrangements, provider networks, and available disease management programs. [56]

3.2.2 *Study Sample*

We used prospective prediction models where the predictor years varied based on an individual's years of continuous enrollment and spanned from 2014 to 2017 while the forecast years spanned from 2015 to 2018 (Table 3.1). To generate a robust prediction model, we identified individuals enrolled on WAHBE for at least two consecutive years between 2014 to 2017. Data from year one provided risk information to predict costs in year two. Enrollees were allowed to be in our data at multiple timepoints if they were enrolled for more than two consecutive years.

Table 3.1. Predictor and Outcome Measurement for Example Patients

Patient Example	**2015**	**2016**	**2017**	**2018**
Enrolled continuously for 2 years, between 2015-2016	Predictor	Outcome		
Enrolled continuously for 3 years, between 2015-2017	Predictor	Outcome		
		Predictor	Outcome	
Enrolled continuously for 4 years, between 2015-2018	Predictor	Outcome		
		Predictor	Outcome	
			Predictor	Outcome

3.2.3 *Outcome Variable*

Our outcome was annual health care costs, defined as the sum of outpatient, inpatient, professional, and pharmacy expenditures between January and December in the forecast year. We used both insurer and enrollee paid amounts to calculate total costs.

3.2.4 *Predictor Variables*

We created a diagnostic profile of each enrollee using the 2018 Department of Health and Human Services Hierarchical Clinical Categories (HHS-HCC), which included 128 clinical conditions and 12 prescription drug categories. HHS-HCCs are similar to the Centers for Medicare and Medicaid Hierarchical Clinical Categories (CMS-HCC) but include medical conditions that are more salient to a commercially insured population, including pregnancy and neonatal complications.[80] We used diagnoses from inpatient, outpatient, and professional claims to define the HHS-HCC clinical categories.

We also measured annual health care utilization, including the number of inpatient, outpatient, professional, emergency room, and operating room visits along with the number of unique brand and generic prescription. We determined the type of medical care visit using the type of setting and place of setting codes listed in the claim header.[92] We counted the number of unique national drug code (NDC) prescriptions in the pharmacy claims data for each enrollee. Generic and branded prescriptions were distinguished using a generic drug indicator in the WA-APCD data.

We also included prior year health care costs, enrollee demographics, and health plan characteristics as predictors in our models. Enrollee demographics included age, race, income, eligibility for health insurance subsidies, county of residence and family size. Income is reported in the WAHBE administrative data and include values validated through federal agencies.[93] We included all health plan characteristics that were available in our data, including health plan deductible, out-of-pocket maximum, actuarial value, rating area, and prescription drug formulary. We identified prescription drug formularies using the unique formulary identifier submitted by insurers to the Health Plan Finder data.

We preprocessed our data with a frequency and correlation cutoff to eliminate highly correlated or near zero variance predictor variables that may generate unstable solutions.[94] Our dataset included 312 predictors, however, only 90 predictors met our preprocessing requirements and were included in our analysis. A full list of our proposed predictor variables and final inclusion status is listed in Appendix A.

3.2.5 *Prediction Models*

We built a prospective model to predict expenditures in a given year t, EXP_t, using predictors in the prior year, X_{t-1}, using the following structure: $EXP_t = f(X_{t-1})$. We prespecified a tuning grid of hyperparameters for each model (e.g., number of binary splits for tree models, penalty size for regularized regression) and iteratively tuned the grid to maximize model performance.(Appendix B) We trained each model using 10-fold cross-validation, repeated three times to avoid overfitting the data. We compared the models based on their group-level prediction error using root mean-squared error (RMSE), R-squared, mean average error (MAE), and prediction ratio. Lower MAE, lower RMSE, higher R-squared, and prediction ratios closer to one indicate reduced error and better model performance. We evaluated performance across the full sample and, using the models that were trained on the full sample, the performance among enrollees whose health care costs were in every 5th percentile ranging from the 0th percentile (all enrollees) to the 95th percentile. We used the CARET package in R to conduct our analyses.

3.2.6 *Parametric Models*

We assessed the performance of five parametric models including linear regression, which is traditionally used in risk adjustment models because it provides interpretable results that are on a scale of interest.[73] We also consider three regularized linear regression models including least absolute shrinkage and selection operation (LASSO), ridge regression, and elastic net regression. These models use the same functional form as standard linear regression but minimize the residual sum of squares plus an additional penalty term; the penalty term shrinks the coefficients on less important variables down towards zero to reduce collinearity.[95] Because health care expenditures often have a high density of zero expenditures, we estimated a two-part model where the first equation was a logistic regression, modeling whether an expenditure was zero or not, and the second equation was a generalized linear model using a gamma distribution and a log-link function, predicting the expenditures among observations with non-zero expenditures.

3.2.7 *Machine Learning Models*

We explored the performance of seven machine learning models, including multivariate adaptive regression splines (MARS), tree models, ensemble models, and a deep learning model.

MARS is a piecewise linear regression model that combines several linear regression models together in order to model highly non-linear relationships.[96] Decision tree models split the data into homogenous groups based on predictor values and estimate the outcome within each group.[97] Ensemble methods seek to build more powerful models by combining many simple models into a single predictive model. Averaging predictions across multiple models often reduces the variance of the predictions. We considered four ensemble models, including cubist, random forest, gradient boosting machines (GBM), and extreme GBM. Each model applies different techniques to iteratively learn (Table 3.2). We additionally considered an artificial neural network (ANN), which is a type of deep learning model. ANNs use a series of interconnected nodes to identify patterns in the data. The flexibility of ANNs allows the models to learn from highly non-linear and complex relationships.

Table 3.2. Description of Machine Learning Models

Algorithm	**Description**	**Key Tuning Parameters**
Parametric		
Linear Regression	Assumes a linear relationship between outcome and predictors and estimates coefficients by minimizing the RSS.	None
Least Absolute Shrinkage Operator	A type of *regularized linear regression* that seeks to minimize RSS and a penalty term of the sum of absolute values of the coefficients. Allows coefficients of less important predictors to shrink to zero, creating a sparse model.	• L1 penalty term
Ridge Regression	A type of *regularized linear regression* that seeks to minimize RSS and a penalty term of the sum of squares the coefficients. Coefficients of less important predictors to shrink towards zero but cannot equal zero.	• L2 penalty term
Elastic Net	A type of *regularized linear regression* that seeks to minimize RSS and a penalty term of the weighted average between an L1 and L2 penalty.	• L1 penalty term • L2 penalty term

Two-Part Models	Uses two separate equations to model an outcome. The first equation models whether an outcome meets a predetermined criteria and the second equation models the outcome contingent on meeting that criteria. Two-part models are often used to model heavily skewed data, like healthcare expenditures.	None
Machine Learning Models		
Multivariate Adaptive Regression Splines	A piecewise linear model that develops cut points or knots in the data and fits a separate regress between a given two knots. Uses a hinge function to combine the multiple linear models.	• Number of terms in the pruned model • Degree of interaction in predictors
Classification and Regression Tress (CART)	Basic decision tree that uses recursive binary splitting to determine homogenous groups. The predicted value of an observation is the mean medical expenditures within the observation's terminal node.	• Tree depth, as determined by cost complexity parameter
Cubist	Based on recursive binary splitting but predict a multivariate linear regression model in the terminal node rather than estimate the mean outcome. Also incorporates committees and k-nearest neighbors.	• Number of committees • Number of K-Nearest Neighbors
Random Forest	Ensemble method that fits multiple decision trees to resampled data, but only considers a random subset of predictors for each split. Estimates final predictions by taking the average of predictions across trees.	• Number of trees • Number of predictors sampled for each split • Minimum node size
Gradient Boosting Machines	Ensemble method that fits multiple small decisions trees to resamples of the data, where the residuals from the prior tree inform how to improve the next tree.	• Number of trees • Interaction depth • Learning rate • Number of minimum observations in a node

Extreme Gradient Boosting	Ensemble method that is similar to gradient boosting machines but introduces regularization based on penalty terms.	• Number of boosting iterations • L2 regularization • L1 regularization • Learning rate
Neural Network	Deep learning model comprised of a series of nodes that conduct data transformations where the outputs from one node inform the inputs to the next node. Repeats through multiple layers of nodes.	• Number of units in the hidden layer • Weight decay • Learning rate

3.3 RESULTS

We identified 61,881 enrollees across 84,390 enrollment periods that met our inclusion criteria. Our sample averaged about 2.9 (SD=5.1) unique brand NDCs and 1.3 (SD=3.7) unique generic NDCs annually (Table 3.3). The number of professional office visits averaged 5.3 visits (SD=9.8) in a year while the number of emergency room visits (0.3; SD=1.2), operating room visits (0.1; SD=0.4), and outpatient hospital visits (0.7; SD=2.3) averaged less than one visit. Only three HHS-HCCs had sufficient prevalence in our sample to meet our preprocessing requirements; diabetes without complications (3.6%), diabetes with complications (2.3%) and asthma (2.5%).[80] Health care costs were skewed in the prediction year with a mean and median cost of $3,977 (SD=$16,109) and $497, respectively. Health care costs were also skewed in the forecast year with a mean cost of $4,323 (SD=$18,342) and median cost of $524. The correlation between forecast year and predictor year health care costs was 0.5.

Table 3.3. Description of Medical Expenditures and Selected Predictors in the Analysis Sample

Forecast year	
Medical Expenditures ($)	
Mean	4,323 (SD = 18,341)
Median	523
75th Percentile	2,186
90th Percentile	7,571
95th Percentile	18,745
Predictor Year	
Medical Expenditures ($)	
Mean	3,977 (16,110)
Median	497
Utilization (number of visits, mean(sd))	
Emergency Room Visits	0.3 (1.2)
Operating Room Visits	0.1 (0.4)
Outpatient Hospital Visits	0.7 (2.3)
Professional Visits	5.2 (9.9)
Prescription (number of NDCs, mean(sd))	
Generic	1.3 (3.7)
Brand	2.9 (5.1)
Clinical Conditions (%)	
Diabetes with Chronic Complications	2.3
Diabetes without Complications	3.6
Asthma	2.5
Demographics	
Age (years) (mean/sd)	47.9 (14.3)
Male (%)	45.5
Family Size (mean/sd)	2.2 (1.3)
Health Plan Subscriber (%)	
Race (%)	
White	60.7
Vietnamese	2.2
Chinese	3.3
Other	2.4
CSR Eligibility (%)	
No Subsidy	46.5
73% Actuarial Value Subsidy	16.6
87% Actuarial Value Subsidy	22.1
94% Actuarial Value Subsidy	10.1
Health Plan Characteristics	
Premium Amount ($)	789 (433)
Deductible ($)	4,184 (1,950)
Out-of-Pocket Maximum ($)	6,331 (939)

Actuarial Value	64.7 (14.8)

NOTES Only predictors that met cutoff criteria are shown. Most races (eg Black, American Indian or Alaskan Native, etc) had low prevalence in our data and were excluded due to near-zero variance in the data; these races are not presented in the table.

Because we scaled our outcome to health care costs per $10,000, the root mean squared error (RMSE) and mean average error (MAE) represents error per $10,000. Linear regression and regularized linear regression models performed best with R-squared of nearly 29.9% and RMSE of 1.5 per $10,000 (Table 3.4). However, the performance was similar across parametric and machine learning models with R-squared varying between 19.5% in two-part models to about 29.9% in linear models. MARS and ANN models were the next best performing models and had an R-squared of 29.7% and 29.2%, respectively. The error ranges in predictions were also similarly sized across models (Figure 3.1).

Table 3.4. Comparison of Average Model Performance in Predicting Health Care Costs, by Model

Model	RMSE	R-squared	MAE	Prediction Ratio
LASSO	1.526	0.299	0.420	0.978
Elastic Net	1.526	0.299	0.420	0.999
Linear	1.526	0.299	0.421	0.999
Ridge Regression	1.527	0.299	0.422	0.999
MARS	1.529	0.297	0.418	0.998
Neural Network	1.534	0.292	0.432	1.001
Gradient Boosting Machine	1.539	0.290	0.426	1.001
Extreme Gradient Boosting Machine	1.559	0.272	0.423	0.895
Random Forest	1.566	0.281	0.438	1.091
Cubist	1.572	0.258	0.417	0.886
CART	1.621	0.224	0.444	0.677
Two-Part Model	1.944	0.195	1.049	1.371

NOTES CART = Classification and Regression Trees; GBM = Gradient Boosting Machine; LASSO = Least Absolute Shrinkage and Selection Operator; MARS = Multivariate Adaptive Regression Splines; MAE = Mean Average Error, RMSE = Root Mean Squared Error. Lower MAE, lower RMSE, higher R-squared, and prediction ratios closer to 1 indicate reduced error and better model performance. Health care costs were scaled to cost per $10,000.

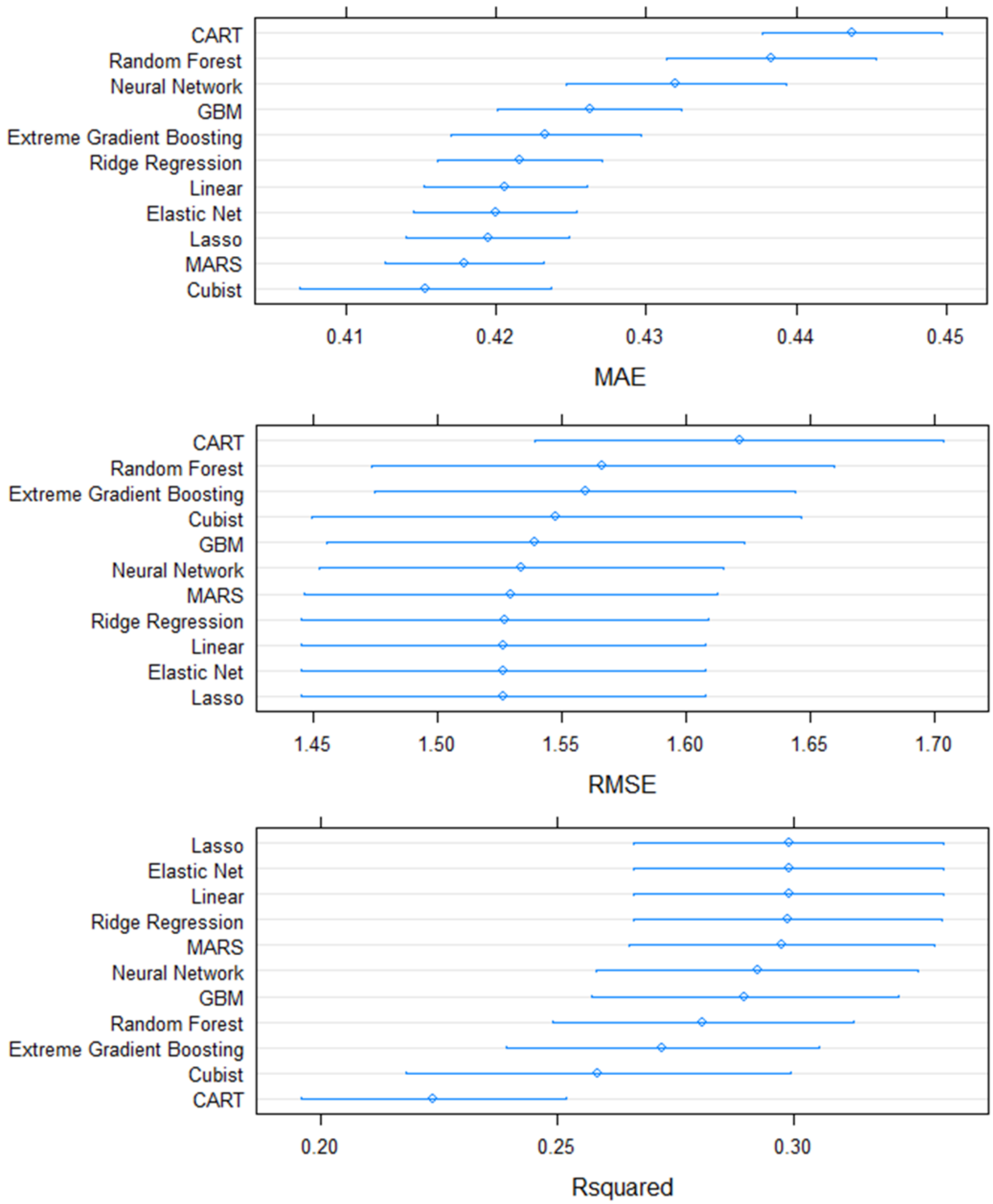

Figure 3.1. Comparison of the Distribution of Prediction Accuracy, by Model.

NOTES CART = Classification and Regression Trees; GBM = Gradient Boosting Machine; LASSO = Least Absolute Shrinkage and Selection Operator; MARS = Multivariate Adaptive Regression Splines; MAE = Mean Average Error, RMSE = Root Mean Squared Error. The points in the whisker's plots indicate the mean error and the capped bars indicate the maximum and minimum errors. Lower MAE, lower RMSE, and higher R-squared values indicate reduced error and better model performance. Health care costs were scaled to cost per $10,000.

The R-squared performance of most models decreased at the tail of the distribution (Figure 3.2). However, the R-squared of random forests increased across higher percentiles ranging from 21.4% to 36.8% in the 60th and 95th percentiles, respectively. The R-squared of GBMs also remained relatively consistent as compared to other models, ranging between 27.7% to 26.7% in the 5th and 95th percentiles, respectively. This compared to LASSO, where R-squared decreased from 29.9% to 17.6% in the 5th and 95th percentiles. The RMSE performance across the distribution tail had similar trends as R-squared (Figure 3.3).

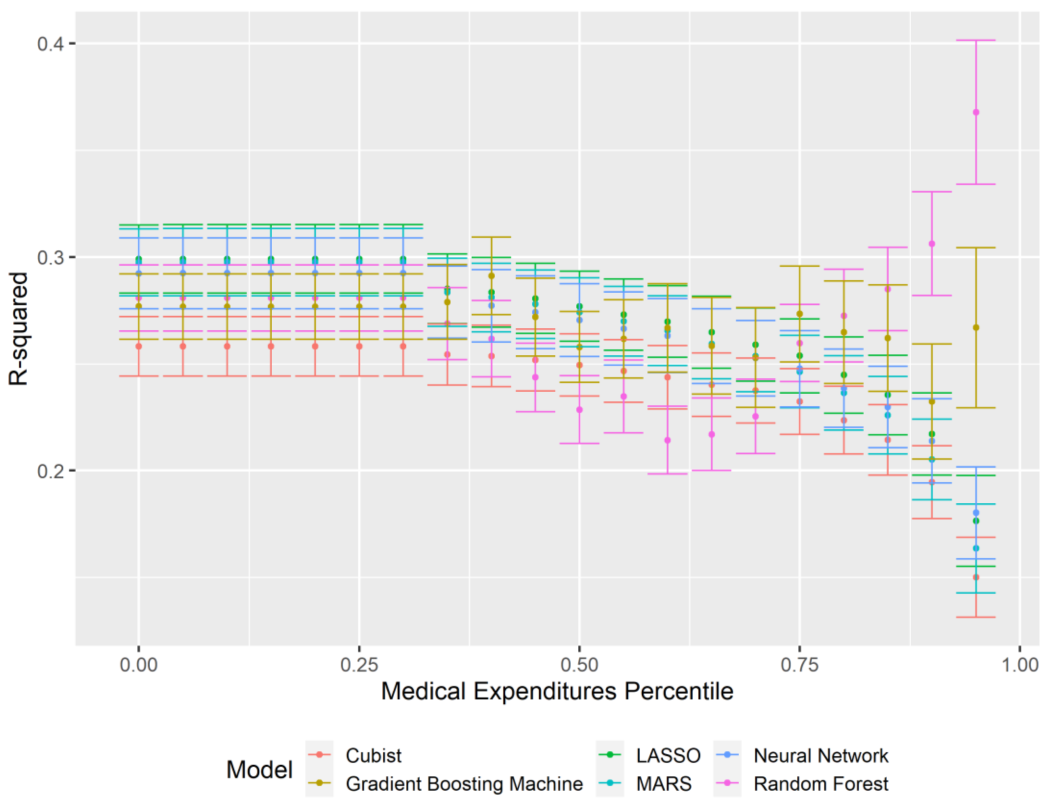

Figure 3.2. R-squared Performance of Select, High Performing Models at the Tail Distribution.

NOTES LASSO = Least Absolute Shrinkage and Selection Operator; MARS = Multivariate Adaptive Regression Splines; RMSE = Root Mean Squared Error. The lines indicate average prediction error at each percentile. Percentiles are percentiles within the predicted distribution of each model. Higher R-squared values indicate reduced error and better model performance. Health care costs were scaled to cost per $10,000. The bars represent the standard errors around the R-squared estimate.

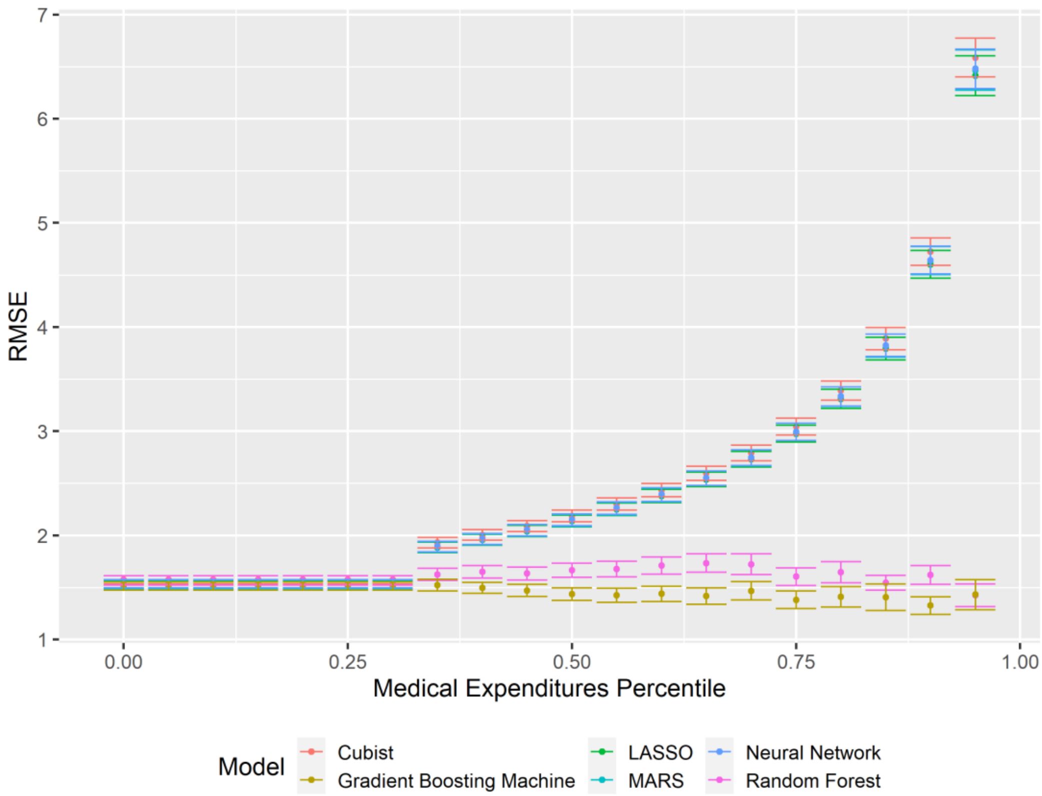

Figure 3.3. RMSE Performance of Select, High Performing Models at the Tail Distribution.

NOTES LASSO = Least Absolute Shrinkage and Selection Operator; MARS = Multivariate Adaptive Regression Splines; RMSE = Root Mean Squared Error. The lines indicate average prediction error at each percentile. Percentiles are percentiles within the predicted distribution of each model. Lower RMSE values indicate reduced error and better model performance. Health care costs were scaled to cost per $10,000. The bars represent the standard errors around the R-squared estimate.

We overlayed the observed distribution of health care costs with the predicted distribution in select models (Figure 3.4). The observed distribution was right skewed with most expenditures under $1,000 (61%). This compared with 49% of predicted observations in cubist model, 36% in random forests, 30% in linear regression, 24% in ANNs, 13% in extreme GBM, and 0% in GBM. The disparity lessened at the tail, where 14% of observations in the observed distribution had expenditures exceeding $5000. The GBM was the next best performing model at the tail with 16%

of observations exceeding $5000 in medical expenditures, followed closely by cubist (17%) and extreme gradient boosting (18%).

Figure 3.4. Medical Expenditures: Observed Versus Predicted Distributions for Select, High Performing Models.

3.4 Discussion

We compared seven machine learning and five parametric models using a large selection of predictors including medical care utilization, diagnoses, health plan characteristics, and detailed demographics. Our study found minimal performance differences between linear models, deep learning and ensemble methods in the full sample. However, linear regression performed increasingly worse when we considered only observations in the tail of the distribution while the performance of ensemble methods, like random forest and GBM, varied minimally across the full distribution as compared to the tail distributions. Ensemble methods could be considered a strong alternative, particularly when identifying high-cost patients is a key priority or when data are heavily right skewed.

Stakeholders should carefully consider their priorities when selecting a model. If model interpretability is a key goal, stakeholders should continue using linear regression models as these models provided the best performance across the full distribution. However, linear regression models poorly predicted very low expenditures and tail observations. Machine learning models may better identify these subpopulations, but at the cost of reduced model interpretability and increased computation time as more complex machine learning models often take longer to train than linear regression. This could be undesirable for a model that needs to be updated frequently. Further, machine learning models provide less transparency than linear regression as many deep learning models incorporate randomization into the estimation process, yielding results that may not be reproducible.[98] Our models were trained using predictors that were available in claims data, HIX administrative data, and health plan characteristics data. Using another set of predictors – either more expansive or more restrictive – may yield less comparable prediction performances across models.

The accuracy of our models falls within the range of other work that prospectively predicts medical expenditures in privately insured individuals, where R-squared values range from 0.15 to 0.48, and performs better than work in the HIX population, where R-squared values range from 0.15 to 0.25.[73, 86, 99] Though machine learning models typically outperform linear regression in the general prediction literature, linear regression yields a higher prediction accuracy in the full distribution than machine learning models in our study. However, this finding aligns with existing literature in health care cost prediction where linear regression typically performs as well – and sometimes better – than machine learning models.[73, 100, 101] The parametric assumptions of linear models lead linear regression to predict a straight line that minimizes the average error term under an assumption that the data are normally distributed.[73] Since medical predictions are often right skewed rather than normally distributed, linear models yield the best average prediction error but do so by overpredicting expenditures at the lower end of the distribution to average out underpredictions at the tail of the distribution.[73, 75, 80]

We had three key innovations in our study. First, to our knowledge, our study is the first study to explore the performance of different machine learning models in the HIX population, while using observed data from HIX enrollees. Existing work uses Marketscan data that has been sampled to resemble HIX enrollees.[73, 86] We compared across a variety of models, including linear regression models, decision trees, ensemble models, and deep learning models. Second, we used a robust dataset that included claims data linked to administrative enrollment information and detailed information about each health plan. This gave us a comprehensive profile of each enrollee during the predictor year. Third, unlike the prior prediction work in the HIX population, we used a comprehensive approach to our covariates, which allowed us to achieve better model accuracy than other studies that use machine learning to predict medical expenditures in an HIX population.

Our study had several limitations. First, only 90 predictors (30%) met our preprocessing requirements, including only three HHS-HCC clinical categories. Second, our analysis was limited to WAHBE enrollees who were continuously enrolled for at least two years and may not generalize to enrollees with shorter enrollment stretches or to other HIXs. Third, our analysis did not include some potentially powerful predictors, including access to care, usual sources of care, employment status, industry codes for employment, the count of claims incurred by an enrollee, or disease interactions.[102, 103] However, these variables may add little predictive power to models when prior year expenditures are also included as a predictor.[102] Fourth, unlike inferential statistics, our prediction framework does not yield standard errors and therefore cannot test hypotheses on a specific relationship between predictors and the medical expenditures.

3.5 CONCLUSIONS

Linear regression only marginally outperformed machine learning models in our study. We further found that, while prediction error in linear regression increased at the tail, prediction error of ensemble methods varied minimally across the full distribution of cost. As such, ensemble methods could be considered a strong alternative to linear regression when predicting medical expenditures in an HIX population. However, the choice between parametric and machine learning models should account for stakeholder priorities as these models face tradeoffs in interpretability, predictive accuracy in subgroups, and computation time. Future studies should consider exploring the predictive power of machine learning models, especially ensemble methods, across a nationwide sample of HIX enrollees.

Chapter 4. DIMINISHING RETURNS: CHOICE AND HEALTH PLAN MENU SIZE ON THE HEALTH INSURANCE EXCHANGES.

4.1 INTRODUCTION

Though most health insurance markets in the United States allow individuals to choose from an array of plans, many individuals struggle to choose a health insurance plan that meets their medical and financial needs, particularly in situations with either a large menu of plan options or complex designed insurance plans.[4, 5, 15, 42, 104, 105] This presents a problem, given the critical role insurance plays in protecting individuals and families from catastrophic medical care costs.[1, 3-6] Underinsuring – or even, failing to enroll - in health insurance coverage can lead to high out-of-pocket expenditures or may result in individuals forgoing necessary medical care, while over insuring may result in enrollees paying too much in health insurance premiums relative to their likely health care needs.[106] Either extreme of insurance enrollment may be mitigated by reducing the burden that enrollees face during open enrollment via enrollees making health plan choices that are more comparable with their medical needs and preferences.

Between 2019 and 2022, the number of health plans offered on the HIXs quadrupled from an average of 26 plans to an average of over 100 plans; this compares to an average of 39 Medicare Advantage plans offered in 2022.[22, 41] HIXs provide individuals and families with access to affordable and competitive non-group health insurance.[38] The rapidly increasing number of plan options on the HIXs has led the Centers for Medicaid and Medicare (CMS) to propose limiting the number of plans that insurance issuers can offer in a service area in plan year (PY) 2024 to prevent HIX enrollees from being overwhelmed by their choices.[107] This proposal is in addition to existing

CMS policies that attempt to simplify open enrollment for HIX enrollees, such as assigning each plan a metal tier label (bronze, silver, gold) that reflects the generosity of the plan and requiring participating insurers to offer at least one plan with a standardized plan design to facilitate easier comparison across a multitude of options.[43, 107] Plans using a standardized design have the same cost-sharing arrangements (eg deductible, out-of-pocket maximum, copays across all benefit categories, etc.) but can differ in non-financial aspects of the plans, such as provider networks, and premiums.[22] Limiting the number of plan options on HIXs may not only simplify open enrollment for HIX enrollees, but may also increase comparison shopping, which in turn may improve price competition on the HIXs, reduce premiums for enrollees, and decrease the burden to taxpayers due to federal subsidies on premiums and cost-sharing arrangements.[22, 105]

Generally, having a larger number of plan options – often referred to as a health plan menu - is thought to be good as individuals may have heterogenous preferences for health plans and having an assortment of plan options increases the likelihood that all individuals find options that fits their needs.[10, 11] However, following the law of diminishing returns, the benefits from each additional option are expected to reduce as the number of plan options increases.[10, 11] With larger menu sizes, each plan will become less distinctive from other, already available options and having the additional options provides less satisfaction to enrollees. Eventually, with a large enough menu size, the assortment of health plans is expected to be comprehensive and new health plan options provide little to no extra value to enrollees.

Further, existing research suggests that individuals making decisions may be overwhelmed when presented with an overabundance of options.[10] Given that individuals have limited mental resources, decisions with too many options can lead to worse outcomes due to an overwhelming amount of information that results in individuals being unable to assess and compare all available options. This phenomenon is known as the paradox of choice or choice overload.[10]

Individuals may face a host of negative consequences from choice overload as these individuals often rely on rough rules of thumb – or, heuristics – to simplify their choices.[10, 108, 109] These heuristics may result in individuals having lower satisfaction with their plan, regretting their choice, or avoid making a plan choice at all.[4, 7, 25] Too much choice may also lead the individual to dread health plan open enrollment periods.[17] Previous research within employer-sponsored insurance and Medicare Advantage has found that offering more than 50 plan options may overload enrollees, however, this work may not generalize to HIXs because HIX enrollees are generally younger than Medicare enrollees while having more chronic conditions and lower health insurance literacy than those in employer-sponsored insurance.[6, 22, 25, 44, 110]

No studies to date have explored overload on the HIXs and this research gap may, in part, be attributable to the challenges in assessing choice overload.[6, 17, 22] The researcher must be able to separate heterogenous preferences for health plan characteristics from heuristically chosen plans that do not align with the enrollees' preferences.[111] The existing literature on choice overload in health insurance is limited to using stated preference experiments, which may not accurately reflect enrollee experiences during open enrollment,[105, 112] or natural experiments where at least one plan option was dominated – in other words, was more expensive with worse benefits – as the choice of a dominated plan is considered irrational.[14, 105, 113] These natural experiments are evaluated only at the observed menu sizes and therefore can only identify if the observed menu sizes caused overload.[14, 105, 113] In lieu of these situations, researchers are left with few tools to assess the point of choice overload and may explore other outcomes that relate to overload, such as the likelihood of switching plans.[35]

4.2 New Contributions

This study takes a first step at addressing choice overload on the HIXs by identifying the point where new health plan options provide little to no extra value to enrollees. While this point of diminished benefits will not necessarily equate to the point of choice overload, it can help inform CMS's policies around plan limits. Assuming the HIXs offer more plans than this point of diminished returns, reducing the number of plans offered on the HIXs to the point where new health plan options provide little to no extra value to enrollees would help reduce the number of options that enrollees must compare without negatively impacting enrollee satisfaction. We measure value as enrollee willingness-to-pay for a given number of health plans; willingness-to-pay is a commonly used metric in health economics to evaluate satisfaction on a dollar scale. Our study seeks to answer the following research questions:

1. Is there a health plan menu size after which willingness-to-pay for additional plans diminishes?
2. Does standardizing health plans affect this menu size?

4.3 METHODS

4.3.1 *Study Setting*

Our analyses use data from a single, state-based HIX, the Washington Health Benefits Exchange (WAHBE). WAHBE began enrollment in 2014 and averages 200,000 enrollees per year across 39 of Washington State's counties.[114] WAHBE is an independent organization, set up by Washington State statute, that oversees the core operations of the exchange, including maintaining the WA HIX website and facilitating enrollment.

4.3.2 *Data Sources*

We combined three data sources for this analysis, WAHBE enrollment data, claims data for WAHBE enrollees, and health plan characteristics for WAHBE plans. WAHBE enrollment data contains information on each enrollees' family size, income, eligibility for federal health insurance subsidies, and choice of health plan. We also identified the default health plan for each WAHBE enrollee, which is designed to prevent accidental disenrollment if an enrollee does not reenroll on WAHBE's website during open enrollment.[115]

We linked WAHBE enrollment data to publicly available health plan characteristics from the Health Plan Finder (HPF) data, and to WAHBE enrollee claims data from the Washington All-Payers Claims Database (WA-APCD). The HPF data contain a comprehensive profile of every plan offered on the HIXs, including premiums, deductibles, out-of-pocket maximums, provider network, actuarial value, and disease management programs. Health insurance issuers who offer health plans on the HIXs must submit this data to the Department of Health and Human Services each year.[57] HPF data are publicly available on CMS's website.[56] We used the 2017 HPF data to identify the sets of plans offered on WAHBE in each county in Washington State.

We obtained WA-APCD data for 2014 to 2018, which includes claims data across all Washington insurers. The WA-APCD data contains diagnostic and health care procedure information as well as paid amounts for inpatient, outpatient, and prescription drug claims. WA-APCD began collecting data in 2014.

4.3.3 *Study Sample*

We included unique subscribers who were continuously enrolled on WAHBE for the 2016 plan year and either had an observed plan choice for the 2017 plan year or voluntary disenrolled from WAHBE in 2017.[116] We excluded enrollees who involuntarily disenrolled for reasons such as non-payment. We included subscribers from a representative set of five Washington counties in our analysis, including King, Spokane, Skagit, Benton, and Okanogan County. We selected these counties based on population size, rurality, physical location, and racial and ethnic diversity (Table 4.1 and Figure 4.1).

Table 4.1. Counties Selected for Analysis

County	Location	Population Size (Percentile)	Ethnic Diversity (Percentile)	Rurality
King	Western WA	75-100th	50-100th	Urban
Spokane	Eastern WA	75-100th	50-100th	Urban
Benton	Eastern WA	50-75th	0-50th	Urban
Okanogan	Eastern WA	25-50th	0-100th	Rural
Skagit	Western WA	50-75th	0-50th	Rural

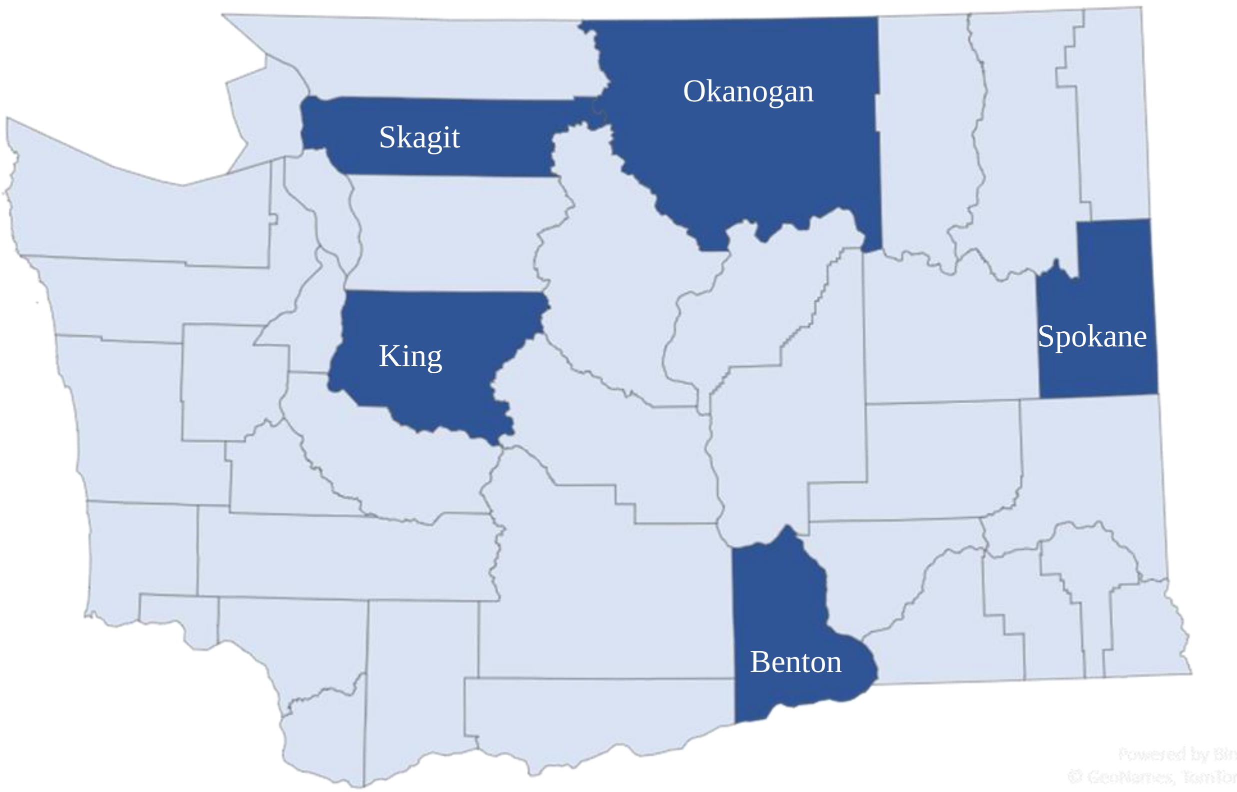

Figure 4.1. Counties Selected for Analysis

4.3.4 *Conceptual Model*

Structural choice models are frequently used to identify utility and WTP across a variety of settings, including health insurance.[4, 7, 15, 32] Utility measures all benefits that an enrollee recieves from a health plan while WTP standardizes utility across a variety of individuals by transforming utility to a dollar scale. Intrinsically, WTP measures the maximum amount that an enrollee would pay for a each health plan menu, measuring enrollee preferences for both the types of plans and the number of plans offered in the menu.[117] WTP is expected to monotonically increase as enrollees are expected to value larger menu sizes more than smaller menus because larger menus provide more opportunities to find a plan that meets their medical needs and preferences. Utility and WTP are often approximated using limited dependent data models, such as multinomial logistic regression, mixed logistic regression, or nested logistic regression.[118, 119]

4.3.5 *Econometric Model*

We used a nested logistic regression to model health plan choice on WAHBE. Nested logistic models can account for decisions with multiple decision levels. We nested our model assuming that enrollees made two decisions when choosing a plan: 1) choice of health plan type, such as health maintenance organization [HMO], preferred provider organization [PPO], or exclusive provider network [EPO] plans and 2) choice of health plan within a given health plan type (4.1). Our nested logistic model estimated the probability, *Pr*, of subscriber, *s*, choosing plan, *p*, within a given health plan type, *k*, as compared to plan type, *l*, when choosing across all health plan types, *N*. This probability was estimated using observed plan and subscriber attributes, *O*, and assumes a correlation of λ.

$$Pr_{sp} = \frac{e^{O_{s,p}/\lambda_k}\left(\sum_{p \epsilon N_k} e^{Os,p/\lambda_k}\right)^{\lambda_k - 1}}{\sum_l^K \left(\sum_{p \epsilon N_l} e^{Os,p/\lambda_l}\right)^{\lambda_l}} \tag{4.1}$$

We nested on plan type because individuals may have strong preferences for health plan types that lead them to consistently choose an HMO plan over a PPO plan or vice versa. A strength of nested logistic regression for our research question is that, unlike many discrete choice models, nested logistic regression models do not assume that the relative probability of two alternatives is held constant even when a third alternative is introduced. This assumption of proportional substitution may be too restrictive in health insurance because individuals can have strong preferences for health plan types.[118]

We additionally modeled a degenerate branch representing disenrollment from WAHBE in order to model all options available to the enrollee.[116] We modeled choice at the subscriber-level as families typically make health insurance choices jointly while considering the medical and financial needs of all enrollees under the same subscriber.[120] We modeled each county separately to allow for different preferences for plan-specific characteristics by geographic location. We modeled choice as a function of both plan-specific characteristics and subscriber-specific characteristics.

4.3.6 *Plan Attributes*

To account for varying preferences for financial characteristics of each plan option, we included the plan's deductible, out-of-pocket maximum, and actuarial value.[4, 42] To account for varying preferences in non-financial aspects of the plan such as geographic proximity to providers and brand perceptions, we included fixed-effects for insurance issuer brand and for the plan's provider network.[121] Each health care provider network offered on the HIX is given a unique identifier in the Health Plan Finder data; we used this identifier to develop our provider network fixed effects.[121] Because the disenrollment option would not have any health plan characteristics, we randomly imputed health plan brand and provider network for the disenrollment branch using observed brands and networks in the county and imputed out-of-pocket maximum, deductible and actuarial value using the mean values in the county.[122] We assumed a zero-dollar premium for the disenrollment nest.

4.3.7 *Subscriber Attributes*

We also included an indicator for whether a health plan was the subscriber's default health plan as WAHBE enrollees are reenrolled into their default plan unless the subscriber actively chooses a new plan. We calculated the family-level expected premium for each plan available in a subscriber's menu using premium rate information from the HPF data. To ensure accuracy, we compared our calculated premiums to each family's observed premium in the WAHBE administrative data.[123] We additionally included subscriber invariant characteristics that may influence choice of health plan, including subscriber sex, family eligibility for health insurance subsidies, family size, and future family-level medical expenditures.

Because a family's expected medical expenditures are integral to their choice of insurance plan, we predicted each family's expected medical expenditures in 2017 based on their 2016 medical care usage and diagnoses. We developed a linear regression prediction model for next-year medical expenditures. We applied the prediction model to the 2016 claims data in our sample to obtain each individual's expected medical expenditures for the 2017 plan year. We then summed expected medical expenditures to the family level.

4.3.8 *Willingness-to-Pay*

WTP measures the utility each subscriber assigns to a set of health plans. We first estimated the utility for each health plan, p, and each subscriber, s, as follows (4.2):

$$\lambda_{s,p} = e^{\beta_{PREM} PREM_{sp} + \beta X_p + \gamma Z_p} \qquad (4.2)$$

Where *PREM* represents the subscriber-specific plan premiums, X represents the linear combination of financial aspects of each plan (eg deductible, out-of-pocket maximum), and Z represents the linear combination of non-financial aspects of each plan (eg issuer brand, provider network). WTP is the ratio of preference for plan characteristics as compared to preferences for premiums (4.3).

$$WTP = \frac{(-1)}{\beta_{PREM}} \left(\sum_{sp} \lambda_{sp}\right) \qquad (4.3)$$

Where β_{PREM} is the coefficient on premium from our nested logistic regression models, λ represents preferences for financial and non-financial aspects of each plan, p, as estimated by coefficients from our nested logistic regression. To estimate total preferences for a given menu size, we summed the WTP across all plans, p, available to a subscriber, s, and across all subscribers.[4, 7]

4.3.9 *Simulation*

We simulated the WTP for each menu size from 1 to 100 health plans. We chose 100 plans as the upper limit because, nationally, HIX enrollees had an average of around 100 plans available to them and because 100 plans exceeded the point of overload in Medicare and employer-sponsored insurance.[6, 22, 25] We first developed a repository of all WAHBE plans that were offered in any Washington county between 2014-2019. We based our simulations on this repository of observed plan options in Washington in order to mimic realistic plan options for WAHBE enrollees. We then used the following simulation strategy:

1. For each given menu size, we sampled that given number of health plans from our plan repository. We sampled without replacement.
2. Following equations 2 and 3 above and using the coefficient from our nested logistic regression models, we estimated the WTP across all plan options and across all subscribers to estimate total WTP for a given menu size.
3. We then repeated steps 1 and 2 above, 100 times.

To simulate a number of plan options that were smaller than the set of plans offered in the county in 2017, we randomly selected a subset of health plans from the observed health plan menu in 2017 in that county. To simulate a number of plan options that were larger than the observed menu, we first used all observed plans in that county in 2017 and then randomly added health plans from our repository to reach the desired menu size. For example, Okanogan offered 13 plans in 2017; to simulate 20 plans for that county, we used all 13 observed plans and then sampled an additional 7 plans from the WAHBE plan repository. When increasing the menu size, we limited the repository to plans that were not offered in that county in 2017 so that we could better simulate a realistic situation as we would not expect an insurance issuer to offer two of the exact same plans within a county. Because our nested logistic models only yielded preferences for networks and issuers that were available in the observed menu, we randomly imputed a network and issuer for sampled plans when the plans 'observed network or issuer were not available in the analysis county.

We conducted secondary simulations that sampled health plans by metal tier. Health plans offered on HIXs are classified into metal tiers (ex: Bronze, Silver, Gold) that categorize the health plan's actuarial value. There are three main metal tiers offered on WAHBE: Bronze and expanded Bronze tiered plans, where the plan's actuarial value is between 56% to 65%; Silver tiered plans, where actuarial values are between 66% to 72%; and Gold tiered plans, where actuarial values are between 76% to 82%.[124] We did not conduct a simulation for Platinum tiered plans (AV ranging from 86% to 92%) because only 5 platinum tiered plans were offered on WAHBE between 2014 and 2019.

We additionally simulated the impact of standardizing plans, including their deductibles, out-of-pocket maximums, and actuarial values, by repeating the primary analysis but using a pool of plans whose cost sharing arrangements were similar to the 2017 HIX standardization policies (Table 4.2).[22] Because there were no plans in our repository that exactly matched the 2017 standardization recommendations, we set a band of ±1 percentage point in actuarial value and ± $500 in the deductible and out-of-pocket maximum to identify a pool of plans whose cost-sharing arrangements were similar to the standardization recommendations. Our pool of standardized plans contained only 59 plans, so we simulated WTP under standardization across menu sizes ranging from one plan to 55 plans. We simulated WTP assuming 25%, 50%, 75% and 100% of plans in a given menu size were standardized, here forth referred to as standardization penetration levels. With each menu size and standardization penetration, we sampled the standardization penetration level from our pool of standardized plans and the remaining plans from a pool of non-standardized plans. For example, to model a menu of ten plans with 50% penetration, we sampled five plans from our standardized plan repository and five plans from our non-standardized repository.

Table 4.2. 2017 Standardization Recommendations and Characteristics of Health Plans in the Standardized and Full Health Plan Simulation Repositories

		Bronze	Silver	Gold
Sample Size	**Standardized Repository**	22	23	14
	Full Repository	114	173	92
Actuarial Value	**2017 Standardized Recommendations**	61.88	70.63	79.98
	Standardized Repository (Mean [Min, Max])	61.5 [60.8; 62.1]	71.1 [70.4; 71.8]	79.0 [78.1, 81.2]
	Full Repository	61.2 [58.5; 62.1]	69.6 [66.0; 72.2]	79.1 [76.2; 82]
Deductible	**2017 Standardized Recommendations**	$6,650	3,500	1,250
	Our Comparable Range	6,741 [6,350; 7,150]	3,674 [3,000; 4000]	1,150 [750; 1,500]
	Full Repository	5,725 [3,750; 7,150]	3,549 [1,250; 7,150]	1,189 [0; 3,800]
OOP	**2017 Standardized Recommendations**	$7,150	7,150	4,750
	Our Comparable Range	7,036 [6,800; 7,650]	7,063 [6,750; 7,350]	4,528 [4,450; 5,000]
	Full Repository	6,662 [5,250; 7,850]	5,810 [3,300; 7,900]	6,415 [4,000; 7,900]

4.3.10 *Change Point Analysis*

We used a multiple change point analysis to identify menu sizes where the slope of WTP changes. We then used these change points as knots in a regression spline analysis to identify the smallest menu size where the marginal change in slope between two consecutive knots was not significant; this menu size represents the point where WTP stopped significantly increasing. Our change point analysis and regression splines were run separately for each county for each standardization penetration level (0%, 25%, 50%, 75%, and 100%).

4.4 RESULTS

Our sample included a total of 24,816 subscribers and ranged from 458 subscribers in Okanogan County to 17,738 subscribers in King County (Table 4.3). The average family size in our sample was less than two family members. Advanced payment tax credit subsidies averaged $214 per month and ranged from an average of $193 in King County to an average of $380 in Okanogan. Family-level medical expenditures were right-skewed, with a mean of $5,141 and median of $2,156. Among subscribers who qualified for an ACA cost-sharing reduction subsidy, the 87% Actuarial Value cost sharing reduction subsidy (CSRs) was the most common CSR (17.4%). This variant is available to enrollees whose income is between 150% and 200% of the federal poverty line. However, the majority of subscribers did not qualify for any cost-sharing reduction subsidies (62.3%).

Table 4.3. Characteristics of 2017 WAHBE enrollees, by analysis county

Variable	Total	King	Spokane	Skagit	Benton	Okanogan
Number of Subscribers	24,816	17,738	3,718	1,114	1,080	458
Male (%)	45.03 (0.32)	46.58 (0.37)	41.23 (0.81)	38.87 (1.46)	41.30 (1.50)	46.07 (2.33)
Subscriber Age (Years) (mean/sd)	46.53 (0.09)	45.55 (0.11)	48.13 (0.23)	49.51 (0.43)	47.88 (0.45)	52.05 (0.59)
Family Size (mean/sd)	1.93 (0.01)	1.87 (0.01)	2.09 (0.02)	2.05 (0.04)	2.16 (0.04)	1.98 (0.05)
Advanced Payment Tax Credit ($) (mean/sd)	214.10 (1.60)	193.36 (1.80)	204.50 (3.59)	351.87 (9.38)	233.06 (7.59)	379.59 (15.25)
Monthly Family Income ($) (mean/sd)	3579.77 (27.24)	3636.85 (35.40)	3422.38 (35.48)	3357.04 (61.83)	3607.58 (162.07)	3267.21 (77.27)
Cost-Sharing Reduction Subsidy (%)						
Tier 1: None	62.34 (0.31)	64.03 (0.36)	56.70 (0.81)	61.58 (1.46)	57.59 (1.50)	55.90 (2.32)
Tier 2: Zero Cost-Sharing	0.31 (0.04)	0.19 (0.03)	0.59 (0.13)	0.81 (0.27)	0.28 (0.16)	0.22 (0.22)
Tier 3: Limited Cost-sharing	0.13 (0.02)	0.10 (0.02)	0.24 (0.08)	0.18 (0.13)	0.09 (0.09)	0.22 (0.22)

Tier 4: 73% AV Variant	9.85 (0.19)	9.06 (0.22)	12.10 (0.53)	10.14 (0.90)	12.50 (1.01)	13.10 (1.58)
Tier 5: 87% AV Variant	17.36 (0.24)	16.08 (0.28)	21.14 (0.67)	19.84 (1.20)	20.28 (1.22)	23.36 (1.98)
Tier 6: 94% AV Variant	10.01 (0.19)	10.54 (0.23)	9.23 (0.47)	7.45 (0.79)	9.26 (0.88)	7.21 (1.21)
Family-Level Medical Predictions ($)						
Mean (SE)	5,141 (72.14)	5,130 (85.13)	4,990 (181.94)	5,223 (363.08)	5,806 (380.79)	4,766 (427.31)
Median	2,156	2,169	2,105	2,157	2,279	2,215
75th	4,796	4,814	4,585	4,739	5,291	4,625
90th	10,882	10,942	10,440	10,145	11,679	9,634
95th	18,467	18,545	17,240	18,162	21,695	16,082

NOTES Families with incomes under 400% of the federal poverty line (FPL) are eligible for advanced payment tax credits; these credits are offered on a sliding income scale. Cost sharing reduction subsidies are also based on income in relation to FPL where: Tier 1 = >250% of FPL; Tier 2 = <300% and Native American or Alaskan Native; tier 3 = >300% and Native American or Alaskan Native; tier 4 = 201-250; tier 5 = 151-200%; and tier 6 = < 150%. Family-level medical expenditures were predicted from 2016 claims data using a linear regression model with 10-fold cross validation, repeated 3 times.

In 2017, health plan menu sizes ranged from 13 plans in Okanogan to 69 plans in King County (Table 4.4). Most plans were silver tiered plans (51%), followed by bronze tiered plans (28%) and gold plans (18%). The number of health insurance issuers ranged from two in Skagit county to seven issuers in King and Spokane counties. While most issuers only offered one or two provider networks across their health plan options, issuers offered an average of four prescription drug formulary tier structures within a county. Okanogan had the largest average monthly premiums ($383) and the largest average annual maximum out-of-pocket costs (MOOP; $6,770), but the smallest annual deductibles ($3,925). This contrasted with Benton, where the average monthly premium was the smallest ($346), but the average annual deductible was the largest ($4,491).

Table 4.4. Characteristics of health plans offered in 2017, by analysis county

	King	Spokane	Skagit	Benton	Okanogan
Number of Plans, 2017					
Total	69	42	15	37	13
Gold	14	7	2	5	3
Silver	35	23	7	21	6
Bronze	19	11	5	10	4
Catastrophic	1	1	1	1	0
Characteristics of Health Plans					

Insurance Issuers	7	7	2	5	3
Provider Networks	10	8	3	6	3
Prescription Drug Formulary Tiers	31	27	13	23	8
Average premium, single 45-year old	355 (79)	352 (86)	359 (64)	346 (79)	383 (80)
MOOP	6,702 (433)	6,635 (714)	6,683 (643)	6,654 (665)	6,770 (976)
Deductible	4,060 (1,928)	4,354 (2,110)	4,407 (2,421)	4,491 (2,101)	3,925 (1,730)
Coinsurance	20.5 (13.4)	17.6 (12.4)	11.3 (9.2)	17.2 (12.6)	31.0 (3.2)
Additional Characteristics of County HIX (universal to all plans)					
Rating Area[2]	Rating Area 1	Rating Area 4	Rating Area 2	Rating Area 5	Rating Area 5
Second Lowest Cost Plan, 2017[1]	Ambetter Balanced Care 2 (2017)	Ambetter Balanced Care 2 (2017)	VisitsPlus Silver HD 17	Ambetter Balanced Care 2 (2017)	LifeWise Essential Silver HSA EPO 3000
Premium for second lowest cost silver plan (single, 45 year old) [1]	268.89	252.76	343.34	267.28	358.89

NOTES MOOP = Maximum out-of-pocket costs. Metal tiers are defined by health plan actuarial value (AV) where: gold tiered AV = 78-82%, silver tiered AV = 70-72%, bronze tiered AV = 58-62%, catastrophic tiered AV < 60%. Rating areas denote geographic areas where an insurance issuer's rates must be uniform across all counties in that area. [1]Source: https://www.wahealthplanfinder.org/content/wahbe/global/en/current-customers/your-1095-a-statement/affordability-exemption/second-lowest-cost-silver.html [2]Source: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Health-Insurance-Market-Reforms/wa-gra

Coefficients from our nested logistic regression models are reported in Table 4.5; the coefficient on premiums were negative and less than -0.00 for all counties. The WTP per person monotonically increased as menu sizes increased, ranging from $1,181 (SE = 55.5) with one plan option to $11,964 (SE = 4.9) with 100 plan options (Figure 1). The slope of WTP changed at seven key menu sizes, ranging from five plan options to 81 plan options. The WTP curve stopped significantly increasing after 81 plan options, with a WTP per person of $11,501. WTP per person also varied by county with the WTP per person for one plan option ranging from $3,843 in Benton to $-1,050 in Okanogan and, for 100 options, ranging from $16,559 in Benton to $4,312 in Okanogan (Figure 4.2). The point where WTP stopped substantially increasing, which was measured as the smallest menu size where the marginal change in slope between two consecutive knots was not significant, varied from 81 plans in King, 62 plans in Spokane, 63 plans in Benton, 35 plans in Skagit, and 81 plans in Okanogan.

Table 4.5. Plan-Specific Coefficients from Nested Logistic Regression Models, by County

Variable	**King**	**Spokane**	**Skagit**	**Benton**	**Okanogan**
Premium	-0.001**	-0.001**	-0.001**	-0.001	-0.001*
Deductible	< .001**	<0.001**	<0.001	<.001**	0.000
OOP	<0.001**	<0.001*	<0.000*	<.001*	0.000
Actuarial Value	0.005**	0.010**	0.014**	0.019**	0.001
Default Plan	1.647**	1.518**	1.151**	1.859**	2.015**
Issuer Brand					
Brand A		*Base*			*Base*
Brand B					
Brand C	*Base*	.165		*Base*	
Brand D	0.120	0.089		0.006	-0.063
Brand E	- 0.183*	0.103	*Base*	0.171	
Brand F	-0.357**	0.127		-0.033	
Brand G	-0.203**		0.296*	-0.052	
Brand H	0.059	0.213			-0.111
Brand I	-0.153**	0.264*			
Provider Networks					
Network 1	*Base*	*Base*	*Base*	*Base*	
Network 2	0.004	0.165			*Base*
Network 3	-0.045				
Network 4	0.073	0.089	0.084	-0.288*	
Network 5		0.103		-0.161	
Network 6		0.127			-0.171
Network 7	0.151*			-0.206	-0.181
Network 8	-0.143*	0.213		-0.260	
Network 9		0.264			
Network 10	0.326**				
Network 11	0.279**				
Network 12	0.092	0.378*		-0.242	
Network 13	0.092*				

NOTES Insurance issuer brand and provider network have been masked; ** p-value < 0.001; * p-value < 0.05.

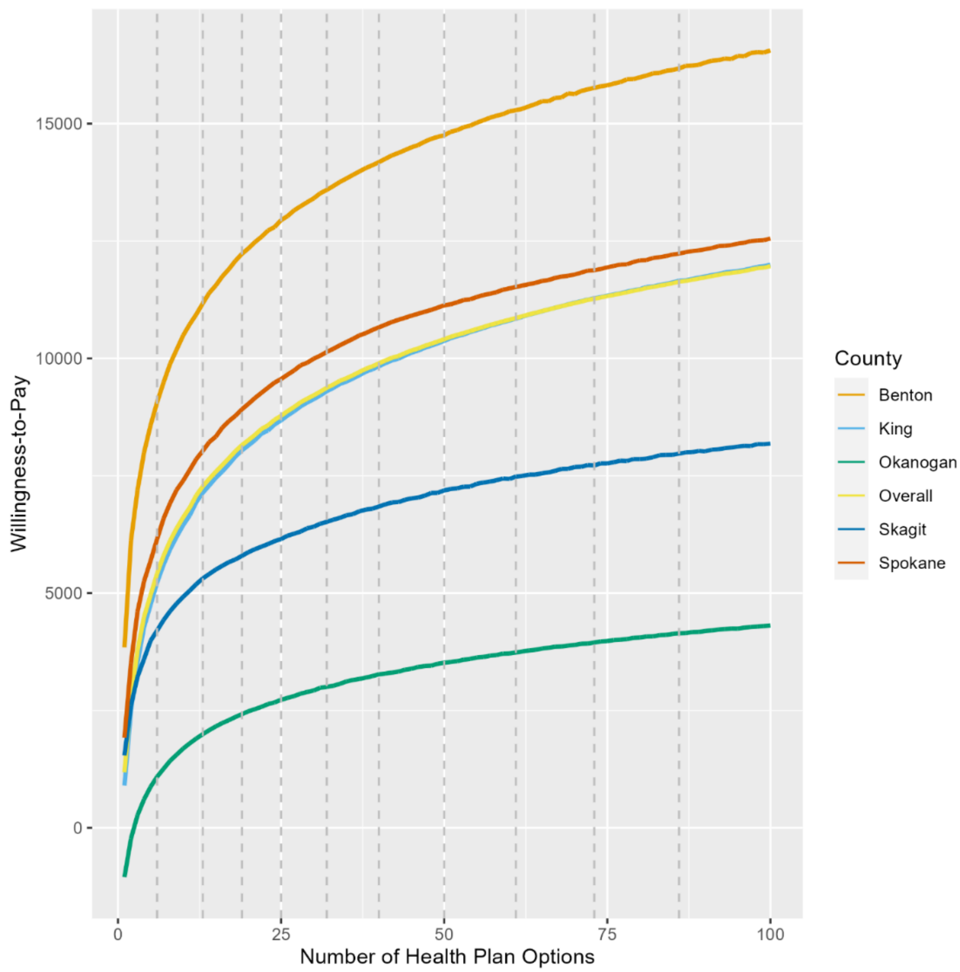

Figure 4.2. Willingness-to-pay for each health plan menu size, by county

NOTES Willingness-to-pay per person for a given health plan menu size. Grey lines indicate change points in the Overall counties curve.

The WTP per person also varied by tier, ranging from $1,416 with one Bronze plan to $12,200 for 100 Bronze plans, from $1,090 with one Silver plan to $11,939 for 100 Silver plans, and from $875 with one Gold plan to $11,740 for 100 Gold plans (Figure 4.3). WTP per person by tier also varied by county, from $-936 for a single Bronze plan in Okanogan to $4,369 for a single Bronze plan in Benton and from $4,438 to $16,851 for 100 bronze plans in Okanogan and Benton counties, respectively. WTP per person for a single Silver plan ranged from $-1,046 in

Okanogan to $4,021 in Benton while WTP for 100 plans ranged from $4,352 in Okanogan to $16,627 in Benton. WTP per person for a single Gold plan ranged from $-1,155 in Okanogan to $3,869 in Benton while MWTP for 100 Gold plans ranged from $4,216 in Okanogan to $16,365 in Benton.

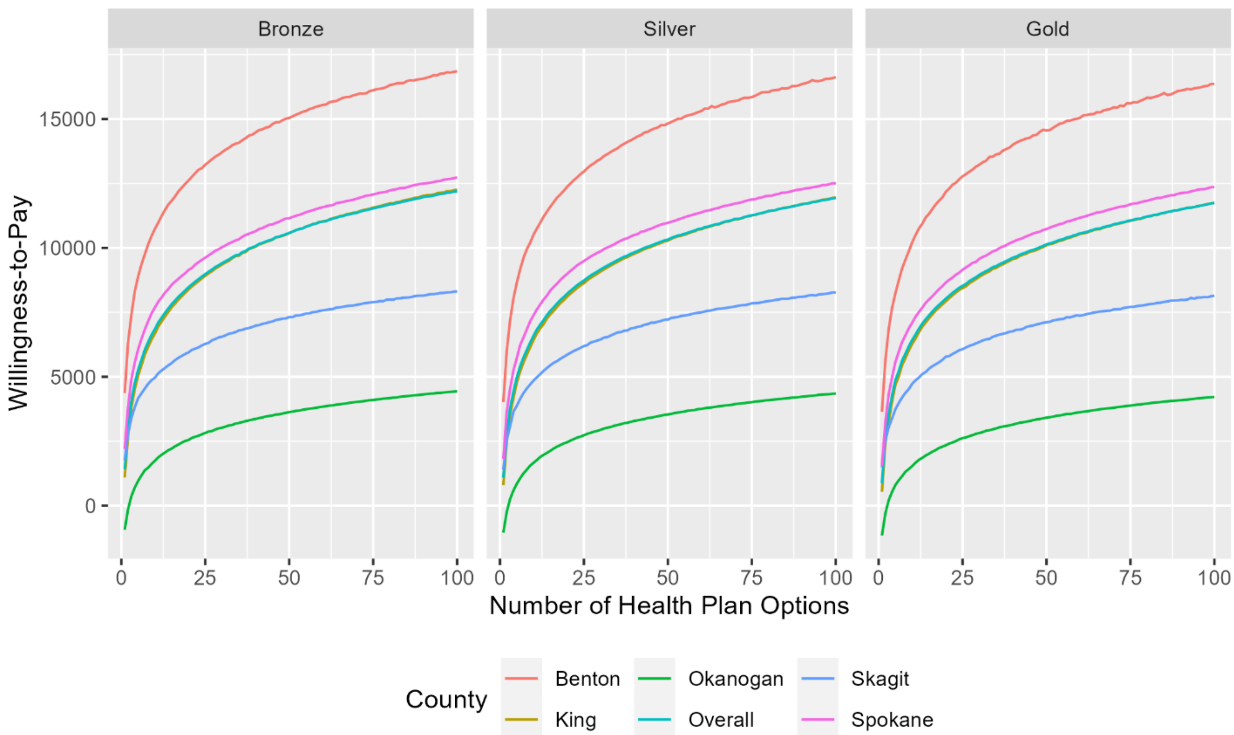

Figure 4.3. Willingness-to-pay for each health plan menu size, by county and metal tier

NOTES Willingness-to-pay per person for a given health plan menu size.

The WTP per person under standardization ranged from $1,032 (SE = $47.2) with one plan option to $10,724 (SE = 7.5) for 55 plan options under 25% standardization penetration and from $1,627 (SE = 23.1) with one plan option to $9,658 (SE = 4.4) for 55 plan options under 100% standardization penetration Figure 4.4). The slope of MWTP changed at up to four different menu sizes, ranging from 6 plans to 36 plans, and these change points were similar across the 25%, 50%, and 75% levels of standardization penetration. WTP stopped significantly increasing after 17 plans with 75% standardization penetration and after 19 plans with 25% penetration. Though increases in WTP were statistically significant across all menu sizes in with 50% and 100% penetration, WTP per person increased by about 10% between highest changepoint (35 plans) and our simulation cap of 55 plans for 50% penetration; this increase was less 5% for 100% penetration. WTP by penetration also differed by county with WTP per person for one plan option ranging from $3,920 (SE = 138.8) to $-1,069 (SE = 20.2) for Benton and Okanogan, respectively, with 25% penetration; from $4,706 (SE = 105.9)to $-969 (SE = 16.9) with 50% penetration; from $5,186 (SE = 35.6) to $-908 (SE = 9.9) with 75% penetration; and from $5,178 (SE = 37.6) to $-926 (SE = 10.5) with 100% penetration.

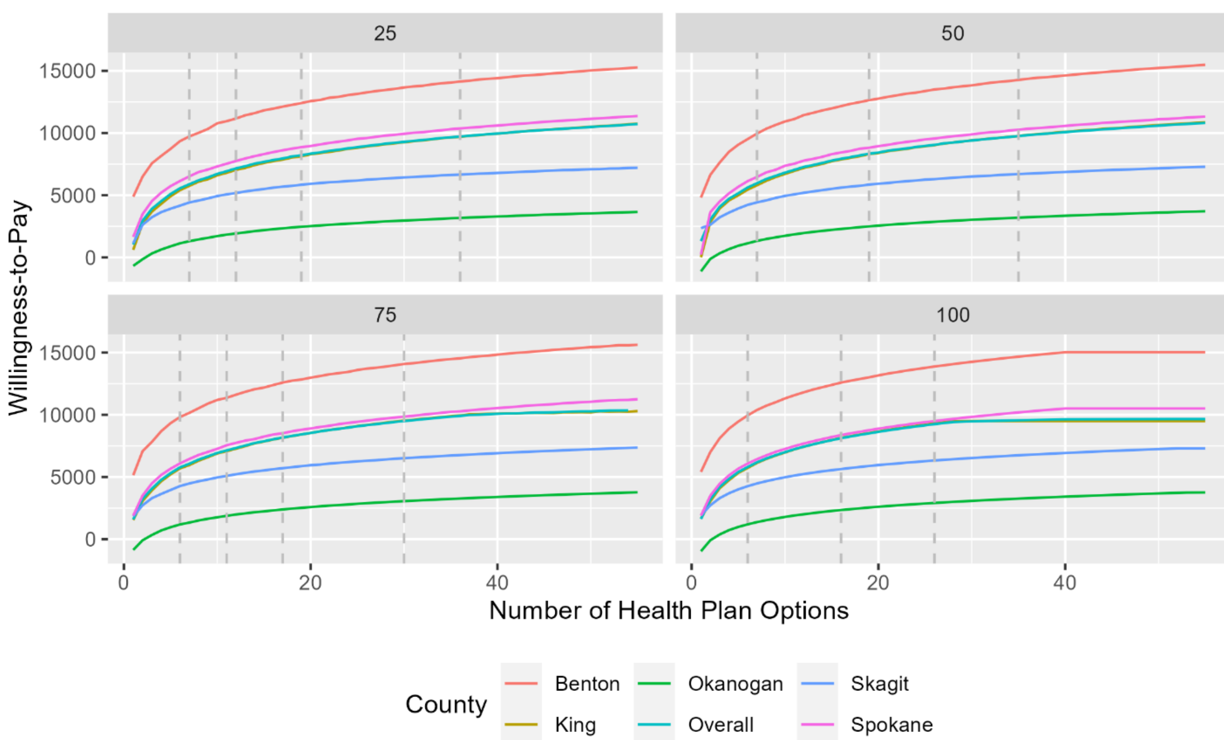

Figure 4.4. Willingness-to-pay for each health plan menu size, by county and standardization penetration

NOTES Willingness-to-pay per person for a given health plan menu size. Grey lines indicate change points in the Overall counties curve.

4.1 DISCUSSION

Health plan menu sizes have rapidly increased on HIXs, leading CMS to consider limiting the number of plans that each issuer is allowed to offer on the HIXs in attempt to reduce health plan menu sizes.[22, 107] Large menu sizes may lead to choice overload and result in enrollees using rules of thumb to choose a plan; have financial and health ramification to enrollees and their families; and result in higher premiums throughout the HIX markets. While the extant literature explores which menu sizes induce choice overload in Medicare and employer-sponsored insurance, these studies may not extrapolate to an HIX population given the differences in health needs and health insurance knowledge for HIX enrollees. Our study is the first to date to explore menu sizes on the HIXs. Our study used a nested logistic regression model to identify individuals' preferences for health plan characteristics and then used these preferences to simulate the value enrollees had for various health plan menu sizes, which was measured as WTP for a given menu size. We found that, across all counties, the maximum amount that enrollees would pay for a single health plan option was $1,759, however, enrollees would be willing to pay up to $10,513 to have 100 health plan options available to them. WTP varied by county, ranging from $16,559 in Benton to $4,312 in Okanogan for 100 plans; these variations are expected as individuals in different geographic locations may have different preferences for health plan characteristics such as cost-sharing arrangements, provider networks, and issuer brands.

We also evaluated the menu sizes where increases in WTP diminished, implying that enrollees gained no additional benefits from having larger plan menus. Our study found that WTP for new plans stopped significantly increasing after about 81 plan options, but this threshold varied by county. Standardizing plan options reduced this threshold to about 20 plan options; this reduction is expected as, by design, plan characteristics under standardization vary minimally so each additional plan is less distinctive from other, already available options and therefore new plan options add little additional benefit for enrollees. Therefore, the benefits from new standardized plans would be expected to diminish at a smaller menu size as compared to a set of non-standardized options.

Our findings are more conservative than prior work in Medicare and employer-sponsored insurance. These comparable studies explored the impact of the number of plan options on enrollee choice of health plan, focusing on poor choices that can be affected by choice overload – such as forgoing enrollment or choosing a dominated plan – and found that more than 30 Medicare Advantage plan options and more than 48 employer-sponsored insurance plans may overload individuals. Though our study does not account for cognitive costs that may induce overload, our goal was to identify a preliminary menu size that could be used to guide future plan limiting policies on the HIXs.

Notably, the average number of plans offered on the HIXs in 2022 was higher than our estimates and this suggests that HIXs could reduce their plan offerings by about 30 plans without negatively impacting enrollees. Limiting health plan menu sizes can be implemented in a couple of ways. First, HIXs can limit menu size using an active purchaser model, where states accept active bids from insurers and manages which insurers can participate in the marketplace.[125] Several state-based HIXs are currently employing an active purchaser model, including California, Rhode

Island, and Massachusetts.[125] Second, HIXs could also limit menu size by requiring insurers to have meaningful differences across the health plans they offer.[126] Meaningful differences requirements have been used in California's HIX and Medicare Advantage to reduce the number of duplicative plans in the market.[126] Using active purchaser agreements and meaningful differences requirements limit menu sizes while allowing variation in plan benefits designs.

We had two key innovations to our study. First, our work is the first to explore the implications of health plan menu size on HIXs, including with full plan standardization, with partial plan standardization, and with no standardization. Second, we combined an all-payers claims dataset, health plan administrative data, and health plan characteristics to get a comprehensive picture of each enrollee's demographics, their medical expenditures, and their health plan options. We also used each enrollee's observed medical expenditures to predict their expected future expenditures, which became a covariate in our nested logistic model to capture each subscriber's expectations of their family's future medical needs.

Our study also had several limitations. First, our analysis was limited to a representative subset of counties in Washington State due to computational feasibility. Washington State is a state-based exchange, and our results may not generalize to all counties in Washington nor to federally-facilitated HIXs. Though our sample is limited to Washington state, our study models economic behavior around insurance choice and our results may inform policymakers and other stakeholders across both state-based and federally-facilitated HIXs. Second, we required a full year of enrollment in 2016 to ensure that we could predict 2017 medical expenditures. Subscribers with discontinuous enrollment may be systematically different than continuously enrolled subscribers. Third, we used a utility-maximizing framework to model health plan choice, which assumes a subscriber made a rational choice. Utility-maximizing frameworks are standard in choice modeling, but may not always hold in individuals' real world insurance choices .[4, 7, 8, 20, 26, 27, 32, 35, 36, 42, 113, 127] Fourth, our simulation approach maintained the existing distribution of metal tiers, however, changing the distribution may yield different WTP. Future work should explore whether the thresholds reported in this study vary when the distribution of metal tiers is changed.

4.2 SUMMARY

Large heath plan menu sizes may yield substantial health and financial ramifications for enrollees and impede competition on HIXs. Though menu sizes have rapidly increased, there is a dearth of literature on choice overload within HIXs. We found that enrollee WTP for new plan stopped significantly increasing after 81 plans with no standardization and after 20 plans with standardization. These estimated menu sizes were smaller than the average menu size in 2022, suggesting that HIXs could reduce plan offerings by about 30 plans without negatively impacting enrollees.

Chapter 5. CONCLUSIONS

5.1 SUMMARY OF FINDINGS

This dissertation explored the choice environment on the Affordable Care Act's (ACA) Health Insurance Exchanges (HIX), focusing on the stability of the choice environment and the number of options available. The HIXs provide health insurance for nearly 10% of individuals in the United States, however, these enrollees must choose from a rapidly growing number of health plan options. Between 2019 and 2022, the number of health plan options on the HIXs quadrupled to an average of 100 plan options in 2022.[22] This large number of plans may result in choice overload among HIX enrollees, and the Centers for Medicaid and Medicare (CMS) is exploring options for reducing the number of plans offered on the HIXs in plan year 2024.[107]

Few studies have explored the choice environment and choice overload on the HIXs. Existing work in Medicare and employer-sponsored insurance suggests that offering more than 30 to 50 health plans may overload enrollees.[6, 22, 25, 44, 110] Individuals enrolled in Medicare or employer-sponsored insurance have different medical needs and health literacy and the results from these insurance markets may not generalize to the HIXs. In this dissertation we explored the number of health plan options where enrollees' benefits tapered off with new plan options. We hypothesized that enrollees' benefits would taper off before the average menu size in plan year 2022. We also hypothesized that the penetration of standardized plans in the market would affect this taper-off point. We also explored the churn in health plans offered on the HIXs.

Overall, our results suggest churn in the HIX choice environment, which could contribute to overload among enrollees and, further, that HIXs could reduce their menu sizes by up to 20 plans without negatively impacting enrollees. In Chapter 2, we explored the changes in plans offered on the HIXs over time, including health plan participation. We found that over 80% of plans remained on the HIXs for two years or less, leading to churn in the types of plans participating in HIXs. We also found that exclusive provider organization (EPO) plans became more common over time.

In Chapter 3, we evaluated the performance of various parametric and machine learning prediction models to predict future medical expenditures among HIX enrollees. The predicted expenditures from these models were used in Chapter 4. We found that linear models provided the best predictions across all enrollees, however, all models performed similarly. Ensemble models outperformed linear models when predicting future expenditures among high-cost enrollees.

In Chapter 4, we used structural choice models to identify enrollee preferences for health characteristics and then, using those revealed preferences, simulated the value enrollees had for various health plan menu sizes. Our results suggest that the benefits of having a larger health plan menu began to diminish after about 80 plan options. This threshold varied with the introduction of standardized plans; the benefits of having a larger health plan menu diminished after 20 plans when some – or all – of the health plans in the menu were standardized. Our estimates suggest that HIXs could reduce their plan offerings by up to 20-30 plans under a non-standardized HIX market and by up to 80 plans under a standardized market, without negatively affecting enrollees' wellbeing.

5.2 Policy Implications

Our work is one of the first studies to explore the choice environment on the HIXs and the first study to begin exploring choice overload. Though our study was limited to a single state-based HIX, the lessons learned from our work may be more broadly applicable to all HIXs.

First, the high churn in the HIX choice environment may reduce enrollees' willingness to shop around for new health plans and results in a high proportion of enrollees allowing themselves to auto-reenroll in their default plan. Because the HIXs allow cross-carrier and cross-tier mapping, enrollees may receive surprise medical bills if their network arrangements have changed. This may add an additional burden to enrollees who need to change their care team due to new provider networks.

Restructuring how enrollees are auto-reenrolled may improve their continuity of care and reduce surprise medical bills via enrollees being mapped to plans that are more similar to their needs and preferences. This could occur through a smart default program that ensures enrollees are not being auto-reenrolled to a dominated health plan and may also account for an enrollees' preferences and health status.[14, 128] Though the HIXs already offer decision support tools, such as total cost estimators and provider look-up tools, these tools often use out-of-date data and could be expanded as well as better personalized for enrollees.[129]

Second, the large number of health plans offered on the HIXs may not be adding much value to enrollees and instead result in enrollees using rules of thumb to choose a plan. This may lead enrollees to dread open enrollment and make health plan choices that do not align with their needs. Reducing the number of plans offered on the HIXs may not only improve enrollees' choice of insurance but may also improve price competition among insurers due to increased comparison shopping. Though CMS is considering capping the number of plans that insurance issuers may offer on the HIXs in PY2024, active purchaser models and meaningful differences standards are two additional policy avenues that CMS could consider to limit the number of plan options on HIXs.

In active purchaser models, states act as a gatekeeper to insurer participation in the HIX markets and accept active bids from insurance issuers.[125] Active purchaser models are being used on several state-based HIXs, including California, Rhode Island, and Massachusetts.[125] Notably, active purchaser models may be most feasible for state-based or state-based federal-partnership HIXs as compared to the federally-facilitated HIXs because federally-facilitated HIXs would need to consider insurer bids across 35 states; the large number of insurers participating in federally-facilitated HIXs may lead to logistical challenges in implementing active purchasing.

With meaningful differences policies, issuers must ensure that each of their plans offer a different benefit – often measured by differences in out-of-pocket costs. This policy is intended to reduce the number of duplicative plans offered in a given market.[126] Meaningful differences requirements have been previously used by CMS on the HIXs and Medicare Advantage and are currently being used by California's HIX.[126] Meaningful difference standards were removed on the HIXs in PY2019 to increase plan offerings.[130]

Third, prediction performance minimally varied across our medical cost prediction models. Linear regressions models may be the best alternative when the prediction goal is estimating expenditures across the full HIX population, such as with the HIX risk adjustment model; notably this risk adjustment model already employs linear regression. Alternatively, ensemble models may be the most appropriate prediction model when the goal is to identify medical costs within subgroups, particularly, within high-cost patients. Ensemble models may be most useful to health systems or health insurers, who may be trying to identify and plan for the future medical needs of their patient population. However, this increased subgroup accuracy comes at the cost of increased computation time and reduced interpretability of ensemble models, leading ensemble models to be undesirable for models that need to be frequently updated.

5.3 Future Directions

Though the HIXs provide insurance to over 16 million individuals in the US,[131] little is known about the health, health insurance literacy, and choice preferences of this population. Because many of the HIX enrollees were previously uninsured,[132] the health insurance knowledge and preferences of this population may differ from other studied populations. Identifying the needs of this population may be critical in ensuring their health and financial needs are met.

Our estimates correlate to the point of diminished returns rather than choice overload and the point of choice overload may likely occur at a smaller menu size than our works suggests. Future work should expand the knowledgebase of choice overload on the HIXs, by structurally separating unobserved preferences for health plan characteristics from non-rational selection of health plans. Following examples in the Medicare and employer-sponsored insurance literature, this may be achieved by identifying an HIX with a clearly dominated choice option.

Our estimates assume the observed distribution of metal tiers is constant. However, changing the distribution of metal tiers may impact enrollee valuation of health plan menus and, thereby, health plan menu sizes. Future work should explore whether the current distribution of metal tiers is optimal, whether efficiencies could be gained by modifying the distribution of tiers, and whether our point of diminishing returns varies when the distribution of metal tiers is also varied.

Future work should also consider expanding this evidence-based to federally-facilitated HIXs, which provide insurance in 35 states. State-based HIXs are permitted to set certain regulations, so the experiences of enrollees on federally-facilitated HIXs may differ from those on state-based HIXs.

5.4 CONCLUSIONS

This dissertation explored the choice environment on the HIXs and found significant churn in the number of plans offered on the HIXs along with the participation rate and types of plans offered on the HIXs overtime. Our findings also suggest that HIXs could reduce the number of plans offered without significantly impacting enrollee well-being. Future research should continue to expand our understanding of choice on the HIXs by modeling rational, heterogeneous preferences for health plans as compared to cognitive shortcuts and irrational choices made by enrollees due to an overload of choice. Understanding the choice environment and choice overload on the HIXs can help guide policies that aim to improve the health plan choice – and thereby, health and financial well-being – of HIX enrollees.

BIBLIOGRAPHY

1. Hoffman, C. and J. Paradise, *Health insurance and access to health care in the United States.* Annals of the New York Academy of Sciences, 2008. **1136**(1): p. 149-160.
2. *Health Insurance Coverage of the Total Population*. April 9, 2023]; Available from: https://www.kff.org/other/state-indicator/total-population/?currentTimeframe=0&selectedRows=%7B%22wrapups%22:%7B%22united-states%22:%7B%7D%7D%7D&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D#notes.
3. Collins, S.R., M.Z. Gunja, and M.M. Doty, *How well does insurance coverage protect consumers from health care costs?* The Commonwealth Fund, 2017. **23**.
4. Abaluck, J. and J. Gruber, *Evolving choice inconsistencies in choice of prescription drug insurance.* American Economic Review, 2016. **106**(8): p. 2145-84.
5. Bhargava, S., G. Loewenstein, and J. Sydnor, *Do individuals make sensible health insurance decisions? Evidence from a menu with dominated options*. 2015, National Bureau of Economic Research.
6. Bhargava, S., G. Loewenstein, and J. Sydnor, *CHOOSE TO LOSE: HEALTH PLAN CHOICES FROM A MENU WITH DOMINATED OPTIONS.* Quarterly Journal of Economics, 2017. **132**(3): p. 1319-1372.
7. Abaluck, J. and J. Gruber, *Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program.* American Economic Review, 2011. **101**(4): p. 1180-1210.
8. Abaluck, J. and J. Gruber, *Improving the quality of choices in health insurance markets*. 2016, National Bureau of Economic Research.
9. Abaluck, J. and J. Gruber, *When less is more: Improving choices in health insurance markets.* The Review of Economic Studies, 2022.
10. Chernev, A., U. Böckenholt, and J. Goodman, *Choice overload: A conceptual review and meta-analysis.* Journal of Consumer Psychology, 2015. **25**(2): p. 333-358.
11. Baumol, W.J. and E.A. Ide, *Variety in retailing.* Management Science, 1956. **3**(1): p. 93-101.
12. Chernev, A. and R. Hamilton, *Assortment size and option attractiveness in consumer choice among retailers.* Journal of Marketing Research, 2009. **46**(3): p. 410-420.
13. Oppewal, H. and K. Koelemeijer, *More choice is better: Effects of assortment size and composition on assortment evaluation.* International Journal of Research in Marketing, 2005. **22**(1): p. 45-60.
14. Handel, B. and J. Kolstad, *Getting the Most from Marketplaces: Smart Policies on Health Insurance Choices*. 2015: Brookings Institution Washington (DC).
15. Handel, B. and J. Schwartzstein, *Frictions or mental gaps: what's behind the information we (don't) use and when do we care?* Journal of Economic Perspectives, 2018. **32**(1): p. 155-78.
16. Iyengar, S.S. and M.R. Lepper, *When choice is demotivating: Can one desire too much of a good thing?* Journal of personality and social psychology, 2000. **79**(6): p. 995.
17. Quincy, L. and J. Silas, *The evidence is clear: Too many health insurance choices can impair, not help, consumer decision making.* Health Policy Report. Washington, DC:

Consumers Union. Retrieved from http://consumersunion. org/research/report-the-evidence-is-clear-too-manyhealth-insurance-choices-can-impair-not-help-consumer-decision-making/. Accessed April, 2012. **19**: p. 2013.
18. Simon, H.A., *A behavioral model of rational choice.* The quarterly journal of economics, 1955: p. 99-118.
19. Huberman, G., S.S. Iyengar, and W. Jiang, *Defined contribution pension plans: determinants of participation and contributions rates.* Journal of Financial Services Research, 2007. **31**: p. 1-32.
20. Ketcham, J.D., C. Lucarelli, E.J. Miravete, and M.C. Roebuck, *Sinking, swimming, or learning to swim in Medicare Part D.* American Economic Review, 2012. **102**(6): p. 2639-73.
21. Morrin, M., S. Broniarczyk, J.J. Inman, and J. Broussard, *Saving for retirement: The effects of fund assortment size and investor knowledge on asset allocation strategies.* Journal of Consumer Affairs, 2008. **42**(2): p. 206-222.
22. Chu, R.C., J. Rudich, A. Lee, C. Peters, N. De Lew, and B.D. Sommers, *Facilitating Consumer Choice: Standardized Plans in Health Insurance Marketplaces.* 2021.
23. Quincy, L., *What's behind the door: Consumers' difficulties selecting health plans.* Washington, DC: Consumers Union. Retrieved from http://www.consumersunion.org/wp-content/uploads/2013/04/Consumer-Difficulties-Selecting-Health-Plans-Jan-2012.pdf., 2012. **19**: p. 2013.
24. Quincy, L., *Making health insurance cost-sharing clear to consumers: Challenges in implementing health reform's insurance disclosure requirements.* Issue Brief (Commonw Fund), 2011. **2**: p. 1-17.
25. McWilliams, J.M., C.C. Afendulis, T.G. McGuire, and B.E. Landon, *Complex Medicare Advantage choices may overwhelm seniors—especially those with impaired decision making.* Health Affairs, 2011. **30**(9): p. 1786-1794.
26. Heiss, F., D. McFadden, and J. Winter, *Mind the gap! Consumer perceptions and choices of Medicare Part D prescription drug plans*, in *Research findings in the economics of aging*. 2010, University of Chicago Press. p. 413-481.
27. Ho, K., J. Hogan, and F. Scott Morton, *The impact of consumer inattention on insurer pricing in the Medicare Part D program.* The RAND Journal of Economics, 2017. **48**(4): p. 877-905.
28. Marzilli Ericson, K.M., *Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange.* American Economic Journal: Economic Policy, 2014. **6**(1): p. 38-64.
29. Polyakova, M., *Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D.* American Economic Journal: Applied Economics, 2016. **8**(3): p. 165-95.
30. Heiss, F., D. McFadden, J. Winter, A. Wuppermann, and B. Zhou, *Inattention and switching costs as sources of inertia in medicare part d*. 2016, National Bureau of Economic Research.
31. Marton, J., A. Yelowitz, and J.C. Talbert, *Medicaid program choice, inertia and adverse selection.* Journal of health economics, 2017. **56**: p. 292-316.
32. Handel, B.R., *Adverse selection and inertia in health insurance markets: When nudging hurts.* American Economic Review, 2013. **103**(7): p. 2643-82.

33. Sinaiko, A.D. and R.A. Hirth, *Consumers, health insurance and dominated choices.* Journal of Health Economics, 2011. **30**(2): p. 450-457.
34. Strombom, B.A., T.C. Buchmueller, and P.J. Feldstein, *Switching costs, price sensitivity and health plan choice.* Journal of Health economics, 2002. **21**(1): p. 89-116.
35. Ketcham, J.D., C. Lucarelli, and C.A. Powers, *Paying attention or paying too much in Medicare Part D.* American Economic Review, 2015. **105**(1): p. 204-33.
36. Drakea, C., C. Ryanb, and B. Dowdc, *Sources of Consumer Inertia in the Individual Health Insurance Market.* 2019.
37. Sinaiko, A.D., R.J. Zeckhauser, I.G. Cohen, H.F. Lynch, and C.T. Robertson, *Forced to Choose, Again: The Effects of Defaults on Individuals in Terminated Health Plans*. 2015, Citeseer.
38. Forsberg, V.C., *Overview of Health Insurance Exchanges.* Congressional Research Service, June, 2018. **20**.
39. Zhu, J.M., D. Polsky, and Y. Zhang, *State-Based Marketplaces Outperform Federally-Facilitated Marketplaces.* LDI Issue Brief, 2018. **22**(1): p. 1-7.
40. *State Health Insurance Marketplace Types, 2023*. April 9, 2023]; Available from: https://www.kff.org/health-reform/state-indicator/state-health-insurance-marketplace-types/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.
41. Freed, M., A. Damico, and T. Neuman, *Medicare advantage 2022 spotlight: First look. Kaiser Family Foundation*. 2021.
42. Handel, B.R. and J.T. Kolstad, *Health insurance for" humans": Information frictions, plan choice, and consumer welfare.* American Economic Review, 2015. **105**(8): p. 2449-2500.
43. McKillop, C.N., T.M. Waters, C.M. Kaplan, E.K. Kaplan, M.P. Thompson, and I. Graetz, *Three years in–changing plan features in the US health insurance marketplace.* BMC health services research, 2018. **18**(1): p. 450.
44. Karpman, M., S.K. Long, and L. Bart, *The Affordable Care Act's marketplaces expanded insurance coverage for adults with chronic health conditions.* Health Affairs, 2018. **37**(4): p. 600-606.
45. *Washington Health Benefits Exchange State Legislation*. April 9, 2023]; Available from: https://www.wahbexchange.org/about-the-exchange/what-is-the-exchange/legislation/.
46. *Issue Brief #3: Health Benefit Exchange Functions and Responsibilities*. 2011.
47. *Enrollment Reports and Data*, W.H.B. Exchange, Editor. March 2016.
48. *Enrollment Reports & Data*, W.H.B. Exchange, Editor. March 2017.
49. *Enrollment Reports & Data*, W.H.B. Exchange, Editor. October 2014.
50. *Enrollment Reports & Data*, W.H.B. Exchange, Editor. September 2015.
51. *Enrollment Reports & Data*, W.H.B. Exchange, Editor. September 2016.
52. *Enrollment Reports & Data*, W.H.B. Exchange, Editor. September 2017.
53. *Enrollment Reports & Data*, W.H.B. Exchange, Editor. September 2018.
54. *2019 Qualified Health Plan and Qualified Dental Plan Certification Requests*. 2018, Washington Health Benefits Exchange.
55. *2022 Qualified Health Plan and Qualified Dental Plan Certfication Requests*. 2021, Washington Health Benefit Exchange.
56. *Health Insurance Marketplace Public Use Files*. Center for Consumer Information & Insurance Oversight.

57. Oversight, T.C.f.C.I.I. *Health Plan Finder Data*. September 6, 2019]; Available from: https://www.cms.gov/CCIIO/Resources/Data-Resources/health_plan_finder_data.html.
58. Oversight, T.C.f.C.I.I. *Health Insurance State-based Exchange Public Use Files*. September 6, 2019]; Available from: https://www.cms.gov/CCIIO/Resources/Data-Resources/sbm-puf.html.
59. Cox, C., G. Claxton, L. Levitt, M. Long, S. Gonzales, and N. Sroczynski, *Analysis of 2017 Premium Changes and Insurer Participation in the Affordable Care Act's Health Insurance Marketplaces*. 2016: Henry J. Kaiser Family Foundation.
60. Parys, J.V., *ACA marketplace premiums grew more rapidly in areas with monopoly insurers than in areas with more competition.* Health Affairs, 2018. **37**(8): p. 1243-1251.
61. Medicare, C.f. and H. Medicaid Services *Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2017.* Federal register, 2016. **81**(45).
62. Barker, A.R., T.D. McBride, L.M. Kemper, and K. Mueller, *Geographic variation in premiums in health insurance marketplaces.* Rural Policy Brief, 2014(2014 10): p. 1-4.
63. Branham, D.K. and T. DeLeire, *Zero-Premium Health Insurance Plans Became More Prevalent In Federal Marketplaces In 2018.* Health Affairs, 2019. **38**(5): p. 820-825.
64. Crespin, D. and T. DeLeire, *As insurers exit Affordable Care Act marketplaces, so do consumers.* Health Affairs, 2019. **38**(11): p. 1893-1901.
65. Dafny, L., J. Gruber, and C. Ody, *More insurers lower premiums: Evidence from initial pricing in the health insurance marketplaces.* American Journal of Health Economics, 2015. **1**(1): p. 53-81.
66. McDermott, D. and C. Cox. *Insurer Participation on the ACA Marketplaces, 2014-2021*. March 30, 2021]; Available from: https://www.kff.org/private-insurance/issue-brief/insurer-participation-on-the-aca-marketplaces-2014-2021/.
67. Abraham, J.M., A.B. Royalty, and C. Drake, *Plan choice and affordability in the individual and small-group markets: policy and performance—past and present.* Health Affairs, 2019. **38**(4): p. 675-683.
68. Forsberg, V.C., *Overview of Health Insurance Exchanges*, C.R. Services, Editor. 2021.
69. *Health Insurance Exchanges Open Enrollment Period Final Report*, in *April*. 2018, Centers for Medicare and Medicaid Services. p. 2018.
70. Zeng, X., Y. Feng, S. Moosavinasab, D. Lin, S. Lin, and C. Liu. *Multilevel self-attention model and its use on medical risk prediction*. in *PACIFIC SYMPOSIUM ON BIOCOMPUTING 2020*. 2019. World Scientific.
71. Iezzoni, L.I., *Risk adjustment for measuring health care outcomes*. Vol. 3. 2003: Health Administration Press Chicago.
72. van Kleef, R. and T. McGuire, *Risk adjustment, risk sharing and premium regulation in health insurance markets*, in *Theory and practice*. 2018, Academic Press London.
73. Park, S. and A. Basu, *Alternative evaluation metrics for risk adjustment methods.* Health Economics, 2018. **27**(6): p. 984-1010.
74. Pope, G.C., J. Kautter, R.P. Ellis, A.S. Ash, J.Z. Ayanian, L.I. Iezzoni, . . . J. Robst, *Risk adjustment of Medicare capitation payments using the CMS-HCC model.* Health care financing review, 2004. **25**(4): p. 119.
75. Pope, G.C., J. Kautter, M.J. Ingber, S. Freeman, R. Sekar, and C. Newhart, *Evaluation of the CMS-HCC risk adjustment model: final report.* RTI International, 2011. **127**.
76. Hielman, G.R., S.M. Mehmud, and M.A. Rosenberg, *Risk Scoring in Health Insurance: A Primer*. 2016, Society of Actuaries Health Section.

77. Javaid, M., A. Haleem, R. Pratap Singh, R. Suman, and S. Rab, *Significance of machine learning in healthcare: Features, pillars and applications.* International Journal of Intelligent Networks, 2022. **3**: p. 58-73.
78. Yang, C., C. Delcher, E. Shenkman, and S. Ranka, *Machine learning approaches for predicting high cost high need patient expenditures in health care.* BioMedical Engineering OnLine, 2018. **17**(1): p. 131.
79. Kan, H.J., H. Kharrazi, H.-Y. Chang, D. Bodycombe, K. Lemke, and J.P. Weiner, *Exploring the use of machine learning for risk adjustment: A comparison of standard and penalized linear regression models in predicting health care costs in older adults.* PLOS ONE, 2019. **14**(3): p. e0213258.
80. Kautter, J., G.C. Pope, M. Ingber, S. Freeman, L. Patterson, M. Cohen, and P. Keenan, *The HHS-HCC risk adjustment model for individual and small group markets under the Affordable Care Act.* Medicare & medicaid research review, 2014. **4**(3): p. mmrr2014-004-03-a03.
81. Bynum, J. and V. Lewis, *Value-based payments and inaccurate risk adjustment—who is harmed?* JAMA internal medicine, 2018. **178**(11): p. 1507-1508.
82. Schone, E. and R. Brown, *Risk Adjustment: What Is the Current State of the Art and How Can it Be Improved?* POLICY, 2013. **1**(6).
83. AbuKhousa, E. and P. Campbell. *Predictive data mining to support clinical decisions: An overview of heart disease prediction systems*. in *2012 International Conference on Innovations in Information Technology (IIT)*. 2012. IEEE.
84. Lee, W., N. Schwartz, A. Bansal, S. Khor, N. Hammarlund, A. Basu, and B. Devine, *A Scoping Review of the Use of Machine Learning in Health Economics and Outcomes Research: Part 1—Data From Wearable Devices.* Value in Health, 2022.
85. Park, S. and A. Basu, *Improving risk adjustment with machine learning: accounting for service-level propensity scores to reduce service-level selection.* Health Services and Outcomes Research Methodology, 2021. **21**(3): p. 363-388.
86. Rose, S., *A machine learning framework for plan payment risk adjustment*, in *Health Services Research*. 2016. p. 2358+.
87. Rose, S., *Intersections of machine learning and epidemiological methods for health services research.* Int J Epidemiol, 2021. **49**(6): p. 1763-1770.
88. de Ruijter, U.W., Z. Kaplan, W.M. Bramer, F. Eijkenaar, D. Nieboer, A. van der Heide, . . . W.A. Bax, *Prediction Models for Future High-Need High-Cost Healthcare Use: a Systematic Review.* Journal of General Internal Medicine, 2022: p. 1-8.
89. Rosso, R.J., *US health care coverage and spending.* Congressional Research Service (CRS). https://fas. org/sgp/crs/misc/IF10830. pdf, 2020.
90. *What is the Exchange?* September 18, 2022]; Available from: https://www.wahealthplanfinder.org/content/wahbe/global/en/about-the-exchange/what-is-the-exchange.html.
91. *Washington State All Payer Claims Database (WA-APCD).* September 18, 2022]; Available from: https://www.hca.wa.gov/about-hca/data-and-reports/washington-state-all-payer-claims-database-wa-apcd.
92. Wilks, A. and J. Gildner, *Measuring Health Care Utilization in Medicare Advantage Encounter Data: Methods, Estimates, and Considerations for Research.* 2019.

93. *How to Report Income*. September 9, 2022]; Available from: https://www.wahealthplanfinder.org/content/wahbe/global/en/new-customers/application-quick-tips/steps-to-apply/how-to-report-income.html#Verify.
94. Kuhn, M., *Building predictive models in R using the caret package.* Journal of statistical software, 2008. **28**: p. 1-26.
95. Hastie, T., R. Tibshirani, J.H. Friedman, and J.H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. 2009: Springer.
96. Friedman, J.H. and C.B. Roosen, *An introduction to multivariate adaptive regression splines.* Statistical Methods in Medical Research, 1995. **4**(3): p. 197-217.
97. Morgan, J., *Classification and regression tree analysis.* Boston: Boston University, 2014. **298**.
98. Beam, A.L., A.K. Manrai, and M. Ghassemi, *Challenges to the Reproducibility of Machine Learning Models in Health Care.* JAMA, 2020. **323**(4): p. 305-306.
99. Morid, M.A., K. Kawamoto, T. Ault, J. Dorius, and S. Abdelrahman, *Supervised Learning Methods for Predicting Healthcare Costs: Systematic Literature Review and Empirical Evaluation.* AMIA ... Annual Symposium proceedings. AMIA Symposium, 2018. **2017**: p. 1312-1321.
100. Dixon, J., P. Smith, H. Gravelle, S. Martin, M. Bardsley, N. Rice, . . . M. De Lorenzo, *A person based formula for allocating commissioning funds to general practices in England: development of a statistical model.* Bmj, 2011. **343**.
101. Jones, A.M., *625 Models for Health Care*, in *The Oxford Handbook of Economic Forecasting*, M.P. Clements and D.F. Hendry, Editors. 2011, Oxford University Press. p. 0.
102. Bertsimas, D., M.V. Bjarnadóttir, M.A. Kane, J.C. Kryder, R. Pandey, S. Vempala, and G. Wang, *Algorithmic prediction of health-care costs.* Operations Research, 2008. **56**(6): p. 1382-1392.
103. Frees, E.W., J. Gao, and M.A. Rosenberg, *Predicting the frequency and amount of health care expenditures.* North American Actuarial Journal, 2011. **15**(3): p. 377-392.
104. Hero, J.O., A.D. Sinaiko, J. Kingsdale, R.S. Gruver, and A.A. Galbraith, *Decision-Making Experiences Of Consumers Choosing Individual-Market Health Insurance Plans.* Health Aff (Millwood), 2019. **38**(3): p. 464-472.
105. Frank, R.G. and K. Lamiraud, *Choice, price competition and complexity in markets for health insurance.* Journal of Economic Behavior & Organization, 2009. **71**(2): p. 550-562.
106. Collins, S.R., P.W. Rasmussen, S. Beutel, and M.M. Doty, *The problem of underinsurance and how rising deductibles will make it worse. Findings from the Commonwealth Fund Biennial Health Insurance Survey, 2014.* Issue Brief (Commonw Fund), 2015. **13**: p. 1-20.
107. *HHS Notice of Benefit and Payment Parameters for 2024, Proposed Rule*. April 18, 2023]; Available from: https://www.cms.gov/newsroom/fact-sheets/hhs-notice-benefit-and-payment-parameters-2024-proposed-rule.
108. Barch Jr, D.H., *Assessing Models of Subjective Probability Judgment*. 2015, Tufts University.
109. Tversky, A. and D. Kahneman, *The framing of decisions and the psychology of choice.* science, 1981. **211**(4481): p. 453-458.

110. Hewitt, M., *Facilitating state health exchange communication through the use of health literate practices: workshop summary*. 2012: National Academies Press.
111. Chandra, A., B. Handel, and J. Schwartzstein, *Behavioral economics and health-care markets*, in *Handbook of Behavioral Economics: Applications and Foundations 1*. 2019, Elsevier. p. 459-502.
112. Bhargava, S. and G. Loewenstein, *Choosing a health insurance plan: complexity and consequences.* Jama, 2015. **314**(23): p. 2505-2506.
113. Hanoch, Y., T. Rice, J. Cummings, and S. Wood, *How much choice is too much? The case of the Medicare prescription drug benefit.* Health Services Research, 2009. **44**(4): p. 1157-1168.
114. *Health Coverage Enrollment Report*. Spring 2019, Washington Health Benefits Exchange.
115. *Mapping hierarchy for cross-carrier to qualified health plans*. 2018 September 20, 2019]; Available from: https://www.insurance.wa.gov/mapping-hierarchy-cross-carrier-qualified-health-plans.
116. Determann, D., D. Gyrd-Hansen, G.A. de Wit, E.W. de Bekker-Grob, E.W. Steyerberg, M.S. Lambooij, and L. Bjørnskov Pedersen, *Designing Unforced Choice Experiments to Inform Health Care Decision Making: Implications of Using Opt-Out, Neither, or Status Quo Alternatives in Discrete Choice Experiments.* Medical Decision Making, 2019. **39**(6): p. 681-692.
117. Train, K., *Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples.* Journal of choice modelling, 2015. **16**: p. 15-22.
118. Train, K.E., *Discrete choice methods with simulation*. 2009: Cambridge university press.
119. McFadden, D. and K. Train, *Mixed MNL models for discrete response.* Journal of applied Econometrics, 2000. **15**(5): p. 447-470.
120. Royalty, A.B. and J.M. Abraham, *Health insurance and labor market outcomes: Joint decision-making within households.* Journal of Public Economics, 2006. **90**(8): p. 1561-1577.
121. Higuera, L., C.S. Carlin, and B. Dowd, *Narrow Provider Networks and Willingness to Pay for Continuity of Care and Network Breadth.* Journal of Health Economics, 2018.
122. Campbell, D. and S. Erdem, *Including opt-out options in discrete choice experiments: issues to consider.* The Patient-Patient-Centered Outcomes Research, 2019. **12**: p. 1-14.
123. *§ 147.102 Fair health insurance premiums.*, D.o.H.a.H. Services, Editor.
124. Natsis, A., *A Summary of the 2020 Actuarial Value Calculator*. 2019, Society of Actuaries.
125. Shepard, M. and E. Forsgren, *Do insurers respond to active purchasing? Evidence from the Massachusetts health insurance exchange.* Journal of Risk and Insurance, 2022.
126. Rasmussen, P.W. and E.A. Taylor, *What Can the Federal Government Learn From States About Health Insurance Plan Standardization?* JAMA Health Forum, 2021. **2**(11): p. e213478-e213478.
127. Heiss, F., A. Leive, D. McFadden, and J. Winter, *Plan selection in Medicare Part D: Evidence from administrative data.* Journal of Health Economics, 2013. **32**(6): p. 1325-1344.

128. Anderson, D.M., P.W. Rasmussen, and C. Drake. *Estimated plan enrollment outcomes after changes to US health insurance marketplace automatic renewal rules*. in *JAMA Health Forum*. 2021. American Medical Association.
129. Wong, C.A., D.A. Asch, C.M. Vinoya, C.A. Ford, T. Baker, R. Town, and R.M. Merchant, *Seeing Health Insurance and HealthCare.gov Through the Eyes of Young Adults.* Journal of Adolescent Health, 2015. **57**(2): p. 137-143.
130. *HHS Notice of Benefit and Payment Parameters for 2019, Proposed Rule*. April 18, 2018]; Available from: https://www.govinfo.gov/content/pkg/FR-2018-04-17/pdf/2018-07355.pdf.
131. *Marketplace Enrollment, 2014-2023*. May 26, 2023]; Available from: https://www.kff.org/health-reform/state-indicator/marketplace-enrollment/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.
132. Goldman, A.L., D. McCormick, J.S. Haas, and B.D. Sommers, *Effects of the ACA's health insurance marketplaces on the previously uninsured: a quasi-experimental analysis.* Health Affairs, 2018. **37**(4): p. 591-599.

APPENDICES

Appendix A. SUMMARY STATISTICS FOR ALL PREDICTORS AVAILABLE IN THE DATA AND DECISION FOR PREDICTOR INCLUSION IN THE MODELS

Variable Name	Description	Mean	Standard Deviation	Predictor	Reason for drop (if applicable)
Enrollee Characteristics					
Male	Sex of enrollee	0.46	0.50	Yes	-
Age	Age of enrollee (years)	47.93	14.31	Yes	-
Family Size	Number of family members	2.24	1.29	Yes	-
Family Count	Number of family members enrolled in a plan	1.75	0.97	Yes	-
Premium Amount	Total premium paid by a subscriber	788.96	433.35	Yes	-
Year	Plan year	-	-	Yes	-
Subscriber Indicator	Indicates whether the enrollee is primary subscriber under the plan	0.76	0.42	Yes	-
Spouse Indicator	Indicates whether the enrollee is the spouse of the primary subscriber	0.24	0.43	Yes	-
Child Indicator	Indicates whether the enrollee is the child of the primary subscriber	0.07	0.25	Yes	-

Number of Spouses	Indicates total number of spouses enrolled under the plan	0.45	0.52	Yes	-
Number of Children	Indicates total number of children enrolled under the plan	0.24	0.62	Yes	-
Race					
American Indian & Alaskan Native	Enrollee race/ethnicity	0.01	0.08	No	Near-zero variance
Asian or Indian	Enrollee race/ethnicity	0.02	0.14	No	Near-zero variance
Black or African American	Enrollee race/ethnicity	0.01	0.11	No	Near-zero variance
Cambodian	Enrollee race/ethnicity	0.00	0.04	No	Near-zero variance
Chinese	Enrollee race/ethnicity	0.03	0.18	Yes	-
Filipino	Enrollee race/ethnicity	0.01	0.09	No	Near-zero variance
Guamanian	Enrollee race/ethnicity	0.00	0.02	No	Near-zero variance
Hawaiian	Enrollee race/ethnicity	0.00	0.02	No	Near-zero variance
Japanese	Enrollee race/ethnicity	0.01	0.08	No	Near-zero variance
Korean	Enrollee race/ethnicity	0.02	0.14	No	Near-zero variance
Laotian	Enrollee race/ethnicity	0.00	0.03	No	Near-zero variance
Other Asian or Pacific Islander	Enrollee race/ethnicity	0.00	0.07	No	Near-zero variance
Other	Enrollee race/ethnicity	0.02	0.15	Yes	-
Samoan	Enrollee race/ethnicity	0.00	0.01	No	Near-zero variance
Thai	Enrollee race/ethnicity	0.00	0.04	No	Near-zero variance
Vietnamese	Enrollee race/ethnicity	0.02	0.15	Yes	-
White	Enrollee race/ethnicity	0.61	0.49	Yes	-
Relationship to Primary Subscriber					
Child	Enrollee's relationship with primary plan subscriber	0.07	0.25	Yes	-

Domestic Partner	Enrollee's relationship with primary plan subscriber	0.00	0.01	No	Near-zero variance
Self	Enrollee's relationship with primary plan subscriber	0.69	0.46	Yes	-
Spouse	Enrollee's relationship with primary plan subscriber	0.24	0.43	Yes	-
Stepchild	Enrollee's relationship with primary plan subscriber	0.00	0.03	No	Near-zero variance
Other	Enrollee's relationship with primary plan subscriber	0.00	0.00	No	Near-zero variance
Rate Classification					
Couple	Rate classification for plan premium	0.33	0.47	Yes	-
Couple with 1 dependent	Rate classification for plan premium	0.06	0.24	Yes	-
Couple with 2 dependents	Rate classification for plan premium	0.03	0.18	Yes	-
Couple with 3 Dependents	Rate classification for plan premium	0.01	0.10	No	Near-zero variance
Individual	Rate classification for plan premium	0.50	0.50	Yes	-
Individual with 1 dependent	Rate classification for plan premium	0.03	0.18	Yes	-
Individual with 2 dependents	Rate classification for plan premium	0.02	0.12	Yes	-
Individual with 3 dependents	Rate classification for plan premium	0.00	0.06	No	Near-zero variance
CSR Arrangement					
73% Actuarial Value Variant	Federal cost-sharing reduction subsidy (CSR) received	0.10	0.30	Yes	-

87% Actuarial Value Variant	Federal cost-sharing reduction subsidy (CSR) received	0.17	0.38	Yes	-
94% Actuarial Value Variant	Federal cost-sharing reduction subsidy (CSR) received	0.08	0.28	Yes	-
Zero Cost-Sharing; American Indian/Alaska Native	Federal cost-sharing reduction subsidy (CSR) received	0.00	0.06	No	Near-zero variance
Limited Cost-Sharing: American Indian/Alaska Native	Federal cost-sharing reduction subsidy (CSR) received	0.00	0.03	No	Near-zero variance
CSR Eligibility					
No Cost-Sharing Reduction Subsidy (CSR)	Federal CSR eligibility	0.47	0.50	Yes	-
73% Actuarial Value Variant	Federal CSR eligibility	0.17	0.37	Yes	-
87% Actuarial Value Variant	Federal CSR eligibility	0.22	0.41	Yes	-
94% Actuarial Value Variant	Federal CSR eligibility	0.10	0.30	Yes	-
Zero Cost-Sharing; American Indian/Alaska Native	Federal CSR eligibility	0.00	0.06	No	Near-zero variance
Limited Cost-Sharing: American Indian/Alaska Native	Federal CSR eligibility	0.00	0.03	No	Near-zero variance
County					
Asotin	County of Residence	0.00	0.05	No	Near-zero variance
Benton	County of Residence	0.02	0.14	Yes	-
Chelan	County of Residence	0.01	0.11	No	Near-zero variance
Clallam	County of Residence	0.02	0.13	No	Near-zero variance
Clark	County of Residence	0.07	0.25	Yes	-
Columbia	County of Residence	0.00	0.03	No	Near-zero variance
Cowlitz	County of Residence	0.01	0.11	No	Near-zero variance
Douglas	County of Residence	0.00	0.07	No	Near-zero variance

Ferry	County of Residence	0.00	0.03	No	Near-zero variance
Franklin	County of Residence	0.01	0.08	No	Near-zero variance
Garfield	County of Residence	0.00	0.02	No	Near-zero variance
Grant	County of Residence	0.01	0.09	No	Near-zero variance
Grays Harbor	County of Residence	0.01	0.10	No	Near-zero variance
Island	County of Residence	0.02	0.13	No	Near-zero variance
Jefferson	County of Residence	0.01	0.10	No	Near-zero variance
King	County of Residence	0.32	0.47	Yes	-
Kitsap	County of Residence	0.04	0.19	Yes	-
Kittitas	County of Residence	0.01	0.08	No	Near-zero variance
Klickitat	County of Residence	0.01	0.07	No	Near-zero variance
Lewis	County of Residence	0.01	0.10	No	Near-zero variance
Lincoln	County of Residence	0.00	0.04	No	Near-zero variance
Mason	County of Residence	0.01	0.09	No	Near-zero variance
Okanogan	County of Residence	0.01	0.09	No	Near-zero variance
Pacific	County of Residence	0.01	0.07	No	Near-zero variance
Pend Oreille	County of Residence	0.00	0.04	No	Near-zero variance
Pierce	County of Residence	0.07	0.26	Yes	-
San Juan	County of Residence	0.01	0.10	No	Near-zero variance
Skagit	County of Residence	0.02	0.15	Yes	-
Skamania	County of Residence	0.00	0.05	No	Near-zero variance
Snohomish	County of Residence	0.10	0.30	Yes	-
Spokane	County of Residence	0.07	0.25	Yes	-
Stevens	County of Residence	0.01	0.08	No	Near-zero variance
Thurston	County of Residence	0.03	0.18	Yes	-
Wahkiakum	County of Residence	0.00	0.03	No	Near-zero variance
Walla Walla	County of Residence	0.01	0.09	No	Near-zero variance
Whatcom	County of Residence	0.06	0.23	Yes	-
Whitman	County of Residence	0.00	0.07	No	Near-zero variance
Yakima	County of Residence	0.02	0.13	No	Near-zero variance

CSR Arrangement					
73% Actuarial Value Variant	Federal cost-sharing reduction subsidy (CSR) received	0.10	0.30	Yes	-
87% Actuarial Value Variant	Federal cost-sharing reduction subsidy (CSR) received	0.17	0.38	Yes	-
94% Actuarial Value Variant	Federal cost-sharing reduction subsidy (CSR) received	0.08	0.28	Yes	-
Zero Cost-Sharing; American Indian/Alaska Native	Federal cost-sharing reduction subsidy (CSR) received	0.00	0.06	No	Near-zero variance
Limited Cost-Sharing: American Indian/Alaska Native	Federal cost-sharing reduction subsidy (CSR) received	0.00	0.03	No	Near-zero variance
Family Income	Verified monthly family income from WAHBE	4161.06	5430.06	No	High missingness (> 10%)
Federal Poverty Level	Family's income as a percent of the federal poverty level	285.64	356.06	No	High missingness (> 10%)
Advanced Payment Tax Credit (APTC) Eligibility Amount	Federal premium subsidy	362.71	387.24	No	Near-zero variance
Advanced Payment Tax Credit (APTC) Received	Federal premium subsidy received during the plan year	538.63	420.51	No	Large missingness
Utilization					
Total Medical Expenditures	Sum of all medical expenditures accrued in the predictor year	3977.74	16110.55	Yes	-
Outpatient Expenditures	Sum of outpatient expenditures accrued in the predictor year	911.48	6425.99	Yes	-

Professional Expenditures	Sum of professional expenditures accrued in the predictor year	1455.58	5461.96	Yes	-
Pharmacy Expenditures	Sum of pharmacy expenditures accrued in the predictor year	975.48	6884.76	Yes	-
Inpatient Expenditures	Sum of inpatient expenditures accrued in the predictor year	635.21	7359.77	No	Near-zero variance
Clinical Status					
HHS-Hierarchical Condition Categories					
HIV/AIDS	HHS-HCC	0.00	0.07	No	Near-zero variance
Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	HHS-HCC	0.00	0.04	No	Near-zero variance
Central Nervous System Infections, Except Viral Meningitis	HHS-HCC	0.00	0.02	No	Near-zero variance
Viral or Unspecified Meningitis	HHS-HCC	0.00	0.00	No	Near-zero variance
Opportunistic Infections	HHS-HCC	0.00	0.02	No	Near-zero variance
Metastatic Cancer	HHS-HCC	0.00	0.04	No	Near-zero variance
Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia	HHS-HCC	0.00	0.04	No	Near-zero variance
Non-Hodgkin's Lymphomas and Other Cancers and Tumors	HHS-HCC	0.00	0.04	No	Near-zero variance

Colorectal, Breast (Age < 50), Kidney, and Other Cancers	HHS-HCC	0.00	0.05	No	Near-zero variance
Breast (Age 50+) and Prostate Cancer, Benign/Uncertain Brain Tumors, and Other Cancers and Tumors	HHS-HCC	0.01	0.08	No	Near-zero variance
Thyroid Cancer, Melanoma, Neurofibromatosis, and Other Cancers and Tumors	HHS-HCC	0.00	0.05	No	Near-zero variance
Pancreas Transplant Status/Complications	HHS-HCC	0.00	0.01	No	Near-zero variance
Diabetes with Acute Complications	HHS-HCC	0.00	0.02	No	Near-zero variance
Diabetes with Chronic Complications	HHS-HCC	0.02	0.15	Yes	-
Diabetes without Complication	HHS-HCC	0.04	0.19	Yes	-
Protein-Calorie Malnutrition	HHS-HCC	0.00	0.02	No	Near-zero variance
Mucopolysaccharidosis	HHS-HCC	0.00	0.00	No	Near-zero variance
Lipidoses and Glycogenosis	HHS-HCC	0.00	0.01	No	Near-zero variance
Congenital Metabolic Disorders, Not Elsewhere Classified	HHS-HCC	0.00	0.04	No	Near-zero variance
Amyloidosis, Porphyria, and Other Metabolic Disorders	HHS-HCC	0.00	0.02	No	Near-zero variance

Adrenal, Pituitary, and Other Significant Endocrine Disorders	HHS-HCC	0.00	0.07	No	Near-zero variance
Liver Transplant Status/Complications	HHS-HCC	0.00	0.01	No	Near-zero variance
End-Stage Liver Disease	HHS-HCC	0.00	0.02	No	Near-zero variance
Cirrhosis of Liver	HHS-HCC	0.00	0.04	No	Near-zero variance
Chronic Hepatitis	HHS-HCC	0.00	0.00	No	No variation: all observations had same value
Acute Liver Failure/Disease, Including Neonatal Hepatitis	HHS-HCC	0.00	0.02	No	Near-zero variance
Intestine Transplant Status/Complications	HHS-HCC	0.00	0.00	No	No variation: all observations had same value
Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis	HHS-HCC	0.00	0.03	No	Near-zero variance
Intestinal Obstruction	HHS-HCC	0.00	0.03	No	Near-zero variance
Chronic Pancreatitis	HHS-HCC	0.00	0.02	No	Near-zero variance
Acute Pancreatitis/Other Pancreatic Disorders and Intestinal Malabsorption	HHS-HCC	0.00	0.05	No	Near-zero variance
Inflammatory Bowel Disease	HHS-HCC	0.00	0.07	No	Near-zero variance
Necrotizing Fasciitis	HHS-HCC	0.00	0.00	No	Near-zero variance
Bone/Joint/Muscle Infections/Necrosis	HHS-HCC	0.00	0.03	No	Near-zero variance

Rheumatoid Arthritis and Specified Autoimmune Disorders	HHS-HCC	0.01	0.08	No	Near-zero variance
Systemic Lupus Erythematosus and Other Autoimmune Disorders	HHS-HCC	0.00	0.06	No	Near-zero variance
Osteogenesis Imperfecta and Other Osteodystrophies	HHS-HCC	0.00	0.01	No	Near-zero variance
Congenital/Developmental Skeletal and Connective Tissue Disorders	HHS-HCC	0.00	0.02	No	Near-zero variance
Cleft Lip/Cleft Palate	HHS-HCC	0.00	0.00	No	Near-zero variance
Major Congenital Anomalies of Diaphragm, Abdominal Wall, and Esophagus, Age < 2	HHS-HCC	0.00	0.01	No	Near-zero variance
Hemophilia	HHS-HCC	0.00	0.00	No	Near-zero variance
Myelodysplastic Syndromes and Myelofibrosis	HHS-HCC	0.00	0.00	No	Near-zero variance
Aplastic Anemia	HHS-HCC	0.00	0.00	No	Near-zero variance
Acquired Hemolytic Anemia, Including Hemolytic Disease of Newborn	HHS-HCC	0.00	0.01	No	Near-zero variance
Sickle Cell Anemia (Hb-SS)	HHS-HCC	0.00	0.00	No	No variation: all observations had same value
Thalassemia Major	HHS-HCC	0.00	0.01	No	Near-zero variance
Combined and Other Severe Immunodeficiencies	HHS-HCC	0.00	0.01	No	Near-zero variance

Disorders of the Immune Mechanism	HHS-HCC	0.00	0.04	No	Near-zero variance
Coagulation Defects and Other Specified Hematological Disorders	HHS-HCC	0.00	0.05	No	Near-zero variance
Drug Psychosis	HHS-HCC	0.00	0.01	No	Near-zero variance
Drug Dependence	HHS-HCC	0.00	0.05	No	Near-zero variance
Schizophrenia	HHS-HCC	0.00	0.03	No	Near-zero variance
Major Depressive and Bipolar Disorders	HHS-HCC	0.01	0.10	No	Near-zero variance
Reactive and Unspecified Psychosis, Delusional Disorders	HHS-HCC	0.00	0.02	No	Near-zero variance
Personality Disorders	HHS-HCC	0.00	0.03	No	Near-zero variance
Anorexia/Bulimia Nervosa	HHS-HCC	0.00	0.03	No	Near-zero variance
Prader-Willi, Patau, Edwards, and Autosomal Deletion Syndromes	HHS-HCC	0.00	0.00	No	No variation: all observations had same value
Down Syndrome, Fragile X, Other Chromosomal Anomalies, and Congenital Malformation Syndromes	HHS-HCC	0.00	0.02	No	Near-zero variance
Autistic Disorder	HHS-HCC	0.00	0.02	No	Near-zero variance
Pervasive Developmental Disorders, Except Autistic Disorder	HHS-HCC	0.00	0.02	No	Near-zero variance
Traumatic Complete Lesion Cervical Spinal Cord	HHS-HCC	0.00	0.00	No	No variation: all observations had same value
Quadriplegia	HHS-HCC	0.00	0.00	No	Near-zero variance

Traumatic Complete Lesion Dorsal Spinal Cord	HHS-HCC	0.00	0.00	No	No variation: all observations had same value
Paraplegia	HHS-HCC	0.00	0.01	No	Near-zero variance
Spinal Cord Disorders/Injuries	HHS-HCC	0.00	0.03	No	Near-zero variance
Amyotrophic Lateral Sclerosis and Other Anterior Horn Cell Disease	HHS-HCC	0.00	0.01	No	Near-zero variance
Quadriplegic Cerebral Palsy	HHS-HCC	0.00	0.01	No	Near-zero variance
Cerebral Palsy, Except Quadriplegic	HHS-HCC	0.00	0.02	No	Near-zero variance
Spina Bifida and Other Brain/Spinal/Nervous System Congenital Anomalies	HHS-HCC	0.00	0.01	No	Near-zero variance
Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy	HHS-HCC	0.00	0.03	No	Near-zero variance
Muscular Dystrophy	HHS-HCC	0.00	0.01	No	Near-zero variance
Multiple Sclerosis	HHS-HCC	0.00	0.04	No	Near-zero variance
Parkinson's, Huntington's, and Spinocerebellar Disease, and Other Neurodegenerative Disorders	HHS-HCC	0.00	0.03	No	Near-zero variance
Seizure Disorders and Convulsions	HHS-HCC	0.00	0.06	No	Near-zero variance

Hydrocephalus	HHS-HCC	0.00	0.01	No	Near-zero variance
Non-Traumatic Coma, Brain Compression/Anoxic Damage	HHS-HCC	0.00	0.02	No	Near-zero variance
Respirator Dependence/Tracheostomy Status	HHS-HCC	0.00	0.01	No	Near-zero variance
Respiratory Arrest	HHS-HCC	0.00	0.00	No	Near-zero variance
Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes	HHS-HCC	0.00	0.04	No	Near-zero variance
Heart Assistive Device/Artificial Heart	HHS-HCC	0.00	0.00	No	Near-zero variance
Heart Transplant	HHS-HCC	0.00	0.01	No	Near-zero variance
Congestive Heart Failure	HHS-HCC	0.01	0.07	No	Near-zero variance
Acute Myocardial Infarction	HHS-HCC	0.00	0.04	No	Near-zero variance
Unstable Angina and Other Acute Ischemic Heart Disease	HHS-HCC	0.00	0.04	No	Near-zero variance
Heart Infection/Inflammation, Except Rheumatic	HHS-HCC	0.00	0.02	No	Near-zero variance
Hypoplastic Left Heart Syndrome and Other Severe Congenital Heart Disorders	HHS-HCC	0.00	0.00	No	Near-zero variance
Major Congenital Heart/Circulatory Disorders	HHS-HCC	0.00	0.03	No	Near-zero variance
Atrial and Ventricular Septal Defects, Patent	HHS-HCC	0.00	0.03	No	Near-zero variance

Ductus Arteriosus, and Other Congenital Heart/Circulatory Disorders					
Specified Heart Arrhythmias	HHS-HCC	0.01	0.09	No	Near-zero variance
Intracranial Hemorrhage	HHS-HCC	0.00	0.02	No	Near-zero variance
Ischemic or Unspecified Stroke	HHS-HCC	0.00	0.03	No	Near-zero variance
Cerebral Aneurysm and Arteriovenous Malformation	HHS-HCC	0.00	0.02	No	Near-zero variance
Hemiplegia/Hemiparesis	HHS-HCC	0.00	0.02	No	Near-zero variance
Monoplegia, Other Paralytic Syndromes	HHS-HCC	0.00	0.01	No	Near-zero variance
Atherosclerosis of the Extremities with Ulceration or Gangrene	HHS-HCC	0.00	0.02	No	Near-zero variance
Vascular Disease with Complications	HHS-HCC	0.00	0.02	No	Near-zero variance
Pulmonary Embolism and Deep Vein Thrombosis	HHS-HCC	0.00	0.05	No	Near-zero variance
Lung Transplant Status/Complications	HHS-HCC	0.00	0.01	No	Near-zero variance
Cystic Fibrosis	HHS-HCC	0.00	0.01	No	Near-zero variance
Chronic Obstructive Pulmonary Disease, Including Bronchiectasis	HHS-HCC	0.01	0.08	No	Near-zero variance
Asthma	HHS-HCC	0.02	0.16	Yes	-
Fibrosis of Lung and Other Lung Disorders	HHS-HCC	0.00	0.03	No	Near-zero variance

Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections	HHS-HCC	0.00	0.02	No	Near-zero variance
Kidney Transplant Status	HHS-HCC	0.00	0.02	No	Near-zero variance
End Stage Renal Disease	HHS-HCC	0.00	0.02	No	Near-zero variance
Chronic Kidney Disease, Stage 5	HHS-HCC	0.00	0.02	No	Near-zero variance
Chronic Kidney Disease, Severe (Stage 4)	HHS-HCC	0.00	0.02	No	Near-zero variance
Ectopic and Molar Pregnancy, Except with Renal Failure, Shock, or Embolism	HHS-HCC	0.00	0.01	No	Near-zero variance
Miscarriage with Complications	HHS-HCC	0.00	0.00	No	Near-zero variance
Miscarriage with No or Minor Complications	HHS-HCC	0.00	0.03	No	Near-zero variance
Completed Pregnancy With Major Complications	HHS-HCC	0.00	0.01	No	Near-zero variance
Completed Pregnancy With Complications	HHS-HCC	0.00	0.03	No	Near-zero variance
Completed Pregnancy with No or Minor Complications	HHS-HCC	0.00	0.06	No	Near-zero variance
Chronic Ulcer of Skin, Except Pressure	HHS-HCC	0.00	0.04	No	Near-zero variance
Hip Fractures and Pathological Vertebral or Humerus Fractures	HHS-HCC	0.00	0.02	No	Near-zero variance

Pathological Fractures, Except of Vertebrae, Hip, or Humerus	HHS-HCC	0.00	0.01	No	Near-zero variance
Extremely Immature Newborns, Birthweight < 500 Grams	HHS-HCC	0.00	0.00	No	No variation: all observations had same value
Extremely Immature Newborns, Including Birthweight 500-749 Grams	HHS-HCC	0.00	0.00	No	Near-zero variance
Extremely Immature Newborns, Including Birthweight 750-999 Grams	HHS-HCC	0.00	0.00	No	Near-zero variance
Premature Newborns, Including Birthweight 1000-1499 Grams	HHS-HCC	0.00	0.00	No	Near-zero variance
Premature Newborns, Including Birthweight 1500-1999 Grams	HHS-HCC	0.00	0.00	No	Near-zero variance
Premature Newborns, Including Birthweight 2000-2499 Grams	HHS-HCC	0.00	0.00	No	Near-zero variance
Other Premature, Low Birthweight, Malnourished, or Multiple Birth Newborns	HHS-HCC	0.00	0.01	No	Near-zero variance
Term or Post-Term Singleton Newborn, Normal or High Birthweight	HHS-HCC	0.00	0.01	No	Near-zero variance
Stem Cell, Including Bone Marrow, Transplant Status/Complications	HHS-HCC	0.00	0.01	No	Near-zero variance

Artificial Openings for Feeding or Elimination	HHS-HCC	0.00	0.03	No	Near-zero variance
Amputation Status, Lower Limb/Amputation Complications	HHS-HCC	0.00	0.02	No	Near-zero variance
Prescription Drug Categories					
Anti-HIV Agents	HHS-HCC	0.01	0.08	No	Near-zero variance
Anti-Hepatitis C (HCV) Agents	HHS-HCC	0.00	0.03	No	Near-zero variance
Antiarrhythmics	HHS-HCC	0.00	0.04	No	Near-zero variance
Phosphate Binders	HHS-HCC	0.00	0.01	No	Near-zero variance
Inflammatory Bowel Disease Agents	HHS-HCC	0.00	0.05	No	Near-zero variance
Insulin	HHS-HCC	0.01	0.11	No	Near-zero variance
Anti-Diabetic Agents, Except Insulin and Metformin	HHS-HCC	0.02	0.13	No	Near-zero variance
Multiple Sclerosis Agents	HHS-HCC	0.00	0.03	No	Near-zero variance
Immune Suppressants and Immunomodulators	HHS-HCC	0.00	0.06	No	Near-zero variance
Cystic Fibrosis Agents	HHS-HCC	0.00	0.01	No	Near-zero variance
Ammonia Detoxicants	HHS-HCC	0.01	0.08	No	Near-zero variance
Diuretics, Loop and Select Potassium-Sparing	HHS-HCC	0.01	0.11	No	Near-zero variance
# Brand Prescriptions	Unique number of brand prescriptions in predictor year	2.91	5.11	Yes	-
# Generic Prescriptions	Unique number of generic prescriptions in predictor year	1.26	3.73	Yes	-
# Emergency Room Visits	Unique number of emergency room visits in predictor year	0.25	1.23	Yes	-

# Generic Prescriptions	Unique number of generic prescriptions in predictor year	0.06	0.36	Yes	-
# Generic Prescriptions	Unique number of generic prescriptions in predictor year	0.03	0.43	No	Near-zero variance
# Generic Prescriptions	Unique number of generic prescriptions in predictor year	0.01	0.56	No	Near-zero variance
# Generic Prescriptions	Unique number of generic prescriptions in predictor year	0.67	2.25	No	Near-zero variance
# Generic Prescriptions	Unique number of generic prescriptions in predictor year	5.25	9.85	Yes	-
Plan Information					
Actuarial Value	Health Plan's Actuarial Value	54.96	26.76	Yes	-
Combined Medical/Drug Deductible	Indicator for whether a health plan had separate deductibles for medical care and pharmacy claims	0.90	0.30	Yes	-
Multiple Tiers Indicator	An indicator of whether multiple tiers exist for in-network costs	0.04	0.19	Yes	-
Out-of-Pocket Maximum	Maximum amount that enrollees will pay out-of-pocket for covered medical and pharmaceutical services	6330.82	938.50	Yes	-
Deductible (Individual)	Health plan's deductible for a given individual	4183.80	1949.71	Yes	-
Plan Variant Indicator	Indicator for whether a plan was a cost-sharing reduction variant	0.85	0.35	Yes	-

New Plan Indicator	Indicator for whether a health plan had been previously offered on WAHBE	0.41	0.49	Yes	-
Wellness Program	Indicator for whether a health plan offers wellness programs	0.10	0.30	Yes	-
Out of county coverage	Indicator for whether the plan covers medical costs accrued outside the US	0.53	0.50	Yes	-
Out of service coverage	Indicator for whether the plan covers medical costs outside the service area	0.73	0.44	Yes	-
National Network	Indicator for whether the plan is supported by a national network of health services provider companies	0.65	0.48	Yes	-
Health Savings Account (HSA) Eligible	Indicator for whether the plan is eligible for a health savings account	0.60	0.49	Yes	-
Market Type	Indicator for whether a plan was offered on-exchange, off-exchange, or both	0.59	0.49	Yes	-
Issuer	Health Plan Issuer	60007	16990	Yes	-
Plan Tier					
Catastrophic	Metal tier of health plan	0.00	0.07	Yes	-
Gold	Metal tier of health plan	0.08	0.27	Yes	-
Platinum	Metal tier of health plan	0.00	0.03	Yes	-
Silver	Metal tier of health plan	0.59	0.49	Yes	-
Rating Area					
Rating Area 2	Identifier for health plan rating area	0.42	0.49	Yes	-

Rating Area 3	Identifier for health plan rating area	0.07	0.26	Yes	-
Rating Area 4	Identifier for health plan rating area	0.08	0.27	Yes	-
Rating Area 5	Identifier for health plan rating area	0.10	0.31	Yes	-
Plan Type					
HMO	Indicator for plans under a health maintenance organization (HMO)	0.47	0.50	Yes	-
PPO	Indicator for plans using a preferred provider network (PPO)	0.34	0.47	Yes	-
Network Identifier					
WAN002	Designation of the health plan's provider network	0.01	0.11	No	Near-zero variance
WAN003	Designation of the health plan's provider network	0.01	0.10	No	Near-zero variance
WAN004	Designation of the health plan's provider network	0.00	0.06	No	Near-zero variance
WAN005	Designation of the health plan's provider network	0.04	0.19	Yes	-
WAN006	Designation of the health plan's provider network	0.01	0.09	No	Near-zero variance
WAN007	Designation of the health plan's provider network	0.00	0.05	No	Near-zero variance
WAN201	Designation of the health plan's provider network	0.04	0.19	Yes	-
Service Area Identifier					
WAS002	Designation of the health plan's service area	0.03	0.18	Yes	-

WAS003	Designation of the health plan's service area	0.04	0.20	Yes	-
-WAS004	Designation of the health plan's service area	0.01	0.08	No	Near-zero variance
WAS005	Designation of the health plan's service area	0.01	0.12	No	Near-zero variance
WAS006	Designation of the health plan's service area	0.13	0.33	Yes	-
WAS007	Designation of the health plan's service area	0.01	0.10	No	Near-zero variance
WAS201	Designation of the health plan's service area	0.04	0.19	Yes	-
Formulary Identifier					
WAF002	Designation of the health plan's prescription drug formulary	0.18	0.39	Yes	-
WAF003	Designation of the health plan's prescription drug formulary	0.20	0.40	Yes	-
WAF004	Designation of the health plan's prescription drug formulary	0.11	0.31	Yes	-
WAF005	Designation of the health plan's prescription drug formulary	0.09	0.28	Yes	-
WAF006	Designation of the health plan's prescription drug formulary	0.11	0.32	Yes	-
WAF007	Designation of the health plan's prescription drug formulary	0.09	0.28	Yes	-
WAF008	Designation of the health plan's prescription drug formulary	0.01	0.10	No	Near-zero variance
WAF009	Designation of the health plan's prescription drug formulary	0.01	0.11	No	Near-zero variance
WAF010	Designation of the health plan's prescription drug formulary	0.00	0.01	No	Near-zero variance

WAF013	Designation of the health plan's prescription drug formulary	0.00	0.04	No	Near-zero variance
WAF014	Designation of the health plan's prescription drug formulary	0.01	0.10	No	Near-zero variance
WAF015	Designation of the health plan's prescription drug formulary	0.00	0.05	No	Near-zero variance
WAF017	Designation of the health plan's prescription drug formulary	0.00	0.05	No	Near-zero variance
WAF018	Designation of the health plan's prescription drug formulary	0.00	0.06	No	Near-zero variance
WAF019	Designation of the health plan's prescription drug formulary	0.00	0.03	No	Near-zero variance
WAF020	Designation of the health plan's prescription drug formulary	0.00	0.01	No	Near-zero variance
WAF201	Designation of the health plan's prescription drug formulary	0.00	0.05	No	Near-zero variance
WAF202	Designation of the health plan's prescription drug formulary	0.01	0.10	No	Near-zero variance
WAF203	Designation of the health plan's prescription drug formulary	0.02	0.15	Yes	-
Essential Health Benefits Apportionment	The EHB apportionment for pediatric dental.	1.00	-	No	No variation: all observations had same value
Maximum Coinsurance for Specialty Drugs	Maximum amount that enrollee pays for specialty drugs	-	-	No	High missingness (> 10%)
Maximum Inpatient Days	Maximum number of days that a patient will be charged a co-payment for an inpatient stay	4.00	-	No	No variation: all observations had same value

# of Primary Care Visits before cost-sharing	The number of fully covered primary care visits before cost-sharing begins	3.00	-	No	No variation: all observations had same value
# of Copays before coinsurance or deductible	The number of copays a patient pays before cost-sharing begins	3.00	-	No	No variation: all observations had same value
Deductible (Family)	Health plan's deductible for the family	9322.92	3990.92	No	High missingness (> 10%)
Coinsurance	Percent of a covered medical bill that an enrollee is responsible to pay	-	-	No	High missingness (> 10%)
Out-of-Network Deductible	Health plan's deductible for out-of-network care	7437.24	3328.87	No	High missingness (> 10%)

VITA

Diana Poehler is a doctoral candidate in the Department of Health Systems and Population Health at the University of Washington. She earned her bachelor's degree in economics and in mathematics from the University of North Carolina at Greensboro in 2013. She is anticipated to receive her PhD in Health Services from the University of Washington School of Public Health in August of 2023, with an emphasis in Health Economics. Diana has worked as a research assistant across a variety of research organizations since 2011. She began as an undergraduate research assistant at the University of North Carolina at Greensboro, in the Department of Hospitality and Tourism Management, evaluating safer nightlife drinking environments and surveying tourism to North Carolina wineries. She continued her research portfolio with her postgraduate job at RTI International, where she worked on economic and policy evaluations of chronic illnesses, before beginning her PhD. While a PhD student, Diana served as a research assistant for Drs. Beth Devine and Edwin Wong, to assist in Department of Veterans Affairs (VA) sponsored research on amputation-level decision making and the network breadth of VA community care. Diana was funded in the first two years of her doctoral work as T32 Pre-Doctoral Trainee Grant fellow from the Agency for Healthcare Research and Quality.