Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Angela Choksi April 12, 2022

The Impact of Demographic Differences on Consumer Preferences for Public Versus Private Health Insurance in California

by

Angela Choksi

Dr. Evan Saltzman
Adviser

Economics

Dr. Evan Saltzman

Adviser

Dr. Stephen O'Connell

Committee Member

Dr. Holli A. Semetko

Committee Member

2022

The Impact of Demographic Differences on Consumer Preferences for Public Versus Private Health Insurance in California

By

Angela Choksi

Dr. Evan Saltzman

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics

2022

Abstract

The Impact of Demographic Differences on Consumer Preferences for Public Versus Private Health Insurance in California

By Angela Choksi

This study explores whether, and to what extent, the demographic characteristics of Californian consumers influence their health insurance election preferences, specifically with reference to private versus public health insurance options. Moreover, this work has consequential significance for private health insurance companies as it will help them tailor their marketing and advertising decisions to target specific demographics, and also observe the impact on their market share and consumer base if a public option is added to markets where it is not currently available. This research is also valuable for government policymakers, as a potential benefit of implementing a public option, which was advanced in President Biden's 2020 presidential campaign proposals, is that it will stimulate the efficiency and competitiveness of health insurance markets by motivating lower private insurance premiums. Ultimately, this study concludes that year fixed effects, which signal premium changes, are the most significant factor influencing consumer preferences, and that the likelihood of choosing the public option increases as income level decreases, is unaffected by gender and age, and increases for racial minorities and existing enrollees who are actively shopping for a new plan.

The Impact of Demographic Differences on Consumer Preferences for Public Versus Private Health Insurance in California

By

Angela Choksi

Dr. Evan Saltzman

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics

2022

Acknowledgements

I am forever indebted to Dr. Evan Saltzman, my adviser, whose unwavering support and encouragement was instrumental in helping me write and research my honors thesis. I am also grateful to my committee members, Dr. Stephen O'Connell and Dr. Holli A. Semetko, for their insight and guidance. Finally, my heartfelt appreciation extends to my mother Dimple, my family, and my friends, for always bringing a smile to my face and adding a spring to my step.

Table of Contents

1. Introduction & Motivation

According to Edward Kennedy, a former U.S. senator from Massachusetts, “Health care is not just another commodity. It is not a gift to be rationed based on the ability to pay.” To elaborate on his sentiment, health insurance is a basic human necessity that has only been exacerbated as a result of the COVID-19 pandemic that we find ourselves currently living in. One of the key objectives and responsibilities of our government is to ensure equitable and efficient access to health insurance for all citizens. However, according to the January, 2021 “U.S. Health Care Coverage and Spending Report” from the Congressional Research Service, about 9.2% (30 million individuals) of the United States’ 323 million-strong population was uninsured in 2019. The uninsured rate was relatively stable from 2008 to 2013 before dropping 6 percentage points in 2016 to 8.6%. This drop in the uninsured rate corresponds with increases in non-group coverage and Medicaid/CHIP coverage, which are associated with the implementation of various provisions of the Affordable Care Act, such as the exchanges and premium tax credits, and the Medicaid expansion. Since 2016, the uninsured rate has slowly increased.

Furthermore, according to a 2020 report from the Kaiser Family Foundation (KFF), an American non-profit organization that is not associated with the Kaiser Permanente private insurance provider, the people most at risk of being uninsured are low-income adults of color. In 2019, 73.7% of uninsured adults cited the high cost of healthcare coverage as the reason for being uninsured. Most individuals who are uninsured have been without coverage for long periods of time, and although most uninsured people have at least one worker in the family, not all employees have access to coverage through their job. Moreover, studies repeatedly demonstrate that the uninsured are less likely than those with insurance to receive preventive care and services for major health conditions and chronic diseases. Despite the large chunk of

uninsured Americans, individuals (including those who were uninsured), health insurers, and federal and state governments spent approximately $3.6 trillion on various types of health consumption expenditures (HCE) in 2019, which accounted for 16.8% of the nation's gross domestic product, according to the Congressional Research Service. In addition, according to the Organisation for Economic Co-operation and Development's (OECD) "Health at a Glance" 2017 report, U.S. healthcare spending is almost two-and-a-half times the average of the 35 OECD countries, and 25% above Switzerland, the next highest spender.

Additionally, there has been a movement in the past few years to implement a national public option, since advocates argue that it will increase the efficiency and competitiveness of health insurance markets by motivating lower private insurance premiums. Markian Hawryluk of *NPR* states that Colorado and Nevada are planning to implement public option plans in 2023 and 2026, respectively, while Connecticut, Oregon, New Jersey and New Mexico are also strongly considering following suit. While critics of a public option claim that it will drive private insurers out of the market, supporters of it disagree, citing that product differentiation, in terms of different types of policy benefits, network coverages, co-pay and deductible payment structures, etc., will ensure that private insurers coexist with public insurers, and remain in the health insurance market.

From a policy perspective, President Joe Biden's 2020 presidential campaign proposed a national public option for health insurance that resembles a "Medicare-like public insurance option." The plan is touted to "reduce costs for patients by negotiating lower prices from hospitals and other health care providers…better coordinate among all of a patient's doctors to improve the efficacy and quality of their care, and cover primary care without any co-payments… [prohibit] all brand, biotech, and abusively priced generic drugs from increasing

their prices more than the general inflation rate…offer premium-free access to the public option for the 4.9 million individuals denied access to Medicaid, [and] bring relief to small businesses struggling to afford coverage for their employees." However, the campaign was vague on the exact details of the plan, which has not been discussed or enacted by the U.S. Congress or the Biden administration, thus far, and was excluded from Biden's 2021 budget in favor of expanded Medicare coverage. Additionally, President Barack Obama abandoned the idea of a national public insurance option in order to win moderate support for the Affordable Care Act back in 2009.

I am interested in determining the role that demographics play in consumers' health insurance consumption preferences when a public option is available. Consumer demand is a key input in evaluating whether a national public option is a good government policy, and observing how consumers view the public option, and what their demand is for it, is a paramount step towards assessing the effectiveness of a public option. This research study is of consequence because it underlines the demographics that gravitate more towards a public plan versus a private plan, and will thus, be valuable to private health insurance companies when determining which of their health insurance plans might be most appealing to consumers of certain demographic segments, and help them target and tailor their marketing strategies accordingly. It will also help them dissect, and observe the impact on their market share and consumer base if a public option is added to markets where it is not currently available. In addition, this study can also help inform health insurance policy, and assist government policymakers in discerning the most beneficial type of legislation for their constituents. Specifically, future iterations and revisions to this study can assist policymakers in determining whether to expand subsidized health insurance programs such as Medicare and Medicaid, or implement a national public option to provide a

more affordable insurance option. This study will also urge the government to keep public option premiums lower than the average private insurance premium, since I suspect that premium changes are the most potent drivers shaping consumer preferences.

Public health insurance plans are those provided and subsidized by the government. The primary public health programs in the U.S. are Medicare, Medicaid, and CHIP. Contrarily, private health insurance refers to health insurance plans marketed and sold by companies in the private health insurance industry. According to the Congressional Research Service (CRS), "private health insurance is the predominant source of health insurance coverage in the United States," includes employer-sponsored plans, and can be purchased on a group basis or by individual consumers. In 2019, these markets covered an estimated 179 million individuals (55.4% of the U.S. population), and 42 million individuals (13.1% of the U.S. population), respectively. In 2019, private health insurance accounted for $1,195 billion (33.3% of overall HCE). The CRS states that Medicare is a federal health insurance program that pays for covered health care services for most people aged 65 and older and for certain permanently disabled individuals under the age of 65. An estimated 58 million individuals (18.1% of the U.S. population) were enrolled in Medicare in 2019. The program accounted for $799 billion (22.2% of overall HCE) in 2019. Contrastingly, Medicaid is a joint federal-state program that finances the delivery of primary and acute medical services to a diverse low-income population, including children, pregnant women, adults, individuals with disabilities, and people aged 65 and older. Finally, CHIP is a means-tested program that provides health coverage to targeted low-income children and pregnant women in families that have annual income above Medicaid eligibility levels but have no health insurance. An estimated 64 million individuals (19.8% of the U.S.

population) received Medicaid or CHIP in 2019, and the programs accounted for $633 billion (17.6% of overall HCE).

Private health insurance policies are typically more flexible than public health insurance policies, and give the policyholders more leeway, and options as to which doctor, specialist, or medical facility to visit. There are also more private plan choices on the market, which enable individuals to choose the most cost-effective, pocket-friendly, and efficient plan for their needs, and gives individuals a wider network of providers to choose from. However, private health insurance plans are typically more expensive than public health insurance plans, which may not be affordable to many individuals. The majority of private insurers offer both an HMO and a PPO option with differing premiums, while the California public option only offers an HMO plan. HMO or 'Health Maintenance Organization' health insurance plans are more prohibitive, because the coverage restricts patients to a particular group of physicians called a network. Contrastingly, PPO or 'Preferred Provider Organization' plans are less prohibitive and more flexible, because they allow patients to choose any physician they wish, either inside or outside of their insurance network.

Ultimately, the overarching research questions that I am interested in studying are 'What is the consumer demand for a public option plan? Is there heterogeneity in demand across consumers?' Primarily, I found that age does not seem to impact consumption patterns in any way. Secondly, I discovered that as consumers' income level increases, they become less prone to choosing the public option. Thirdly, my results show that neither gender is more prone to electing the public option than the other. Fourthly, my findings illustrate that racial minorities, specifically the Black and Hispanic populations, are more inclined to opt for public insurance. However, my results illustrate an unexplained effect of race on consumption choices, that is not

explained by income level or any of the other explanatory variables in my regression model, and which requires further analysis to accurately dissect. Finally, I found that existing enrollees who are actively choosing a new plan are more inclined towards the public plan than new enrollees. Significantly, I suspect that premium sensitivity is the main driver of consumer election preferences, however, future analysis is necessary to empirically prove this claim.

2. Literature Review

I was able to find very limited amounts of background literature and prior studies that were relevant to my research in this paper, which speaks to the novelty of my study, especially since no one has studied demographic differences in public health insurance consumption patterns yet.

Saltzman (2019) analyzed the demand for health insurance by studying consumer behavior in the California and Washington ACA exchanges. He found that exchange consumers are highly premium sensitive by estimating nested logit discrete choice models of demand for health insurance at the consumer level for both California and Washington. Furthermore, my use of the ACA dataset in this paper is supported by Saltzman's claim that the ACA exchanges provide an appealing context for analyzing health insurance demand. Firstly, the setting provides an opportunity to assess how consumers respond to both policy carrots and sticks that incentivize enrollment. Secondly, employing an analysis of the ACA setting helps to address some of the data shortcomings of examining the pre-ACA individual market, such as the measurement error of premiums, choice sets, and other key variables. Additionally, the detailed demographic information from the ACA Exchanges on variables including income, age, smoking status, and geographic residence, enabled Saltzman to precisely calculate (1) the premium that consumers

face for each plan in their choice sets; (2) the consumer-specific subsidy received for each plan; and (3) the consumer-specific penalty imposed for forgoing coverage.

Saltzman's paper is noteworthy because his empirical work uses consumer-level data along with exogenous variation in premiums in order to estimate consumer premium elasticities in two state ACA exchanges. Additionally, Saltzman's analysis does not consider the importance of provider networks, which can vary considerably between firms, in consumer decision-making. Saltzman contends that data in the ACA setting is sufficiently rich to answer key supply-side questions such as which geographic markets insurers decide to enter, and how they set premiums. Thus, a more robust understanding of both the demand-side and supply-side will help researchers characterize the competitive dynamics in the ACA exchanges, and identify which policy regimes could improve social welfare.

Tebaldi (2017) combined data on household-level enrollment and plan-level claims from the Californian ACA exchange marketplace, in which over 90% of the 1.3 million buyers received federal subsidies, with a model of insurance demand and insurers' competition, in order to study the dependence of equilibrium outcomes on how subsidies interact with three important features of private health insurance markets: demand from subsidized households, insurers' price competition, and adverse selection generated by the correlation between willingness-to-pay and expected health cost. Tebaldi's estimates highlighted a large degree of heterogeneity in consumer preferences, with comparatively older households having a significantly greater willingness to pay for individual health insurance. Moreover, within demographic groups, he found that there is a large degree of dispersion in preferences, particularly among younger households. His main counterfactual shows that one can improve upon the current scheme by providing more generous

subsidies to young buyers, with the possibility of lowering at the same time, the subsidy to the over-50 age group without making them worse off.

Drake (2019) argues that little is known about consumer tastes for network breadth, and how they affect plan selection. He estimated the demand for health insurance plans in California's Marketplace, 'Covered California,' by developing a geospatial measure of network breadth that reflects the physical locations of households and network providers, using 2017 individual enrollment data and provider network directories. He found that households are sensitive to network breadth in their plan choices, meaning that households gravitated more towards health insurance plans with larger network coverages. Variation in the willingness-to-pay indicates that a selection mechanism exists whereby older households sort into broader network plans. He also found that households are highly premium sensitive, which may be a result of plan standardization in Covered California.

Sen et al. (2018), investigated the effects of Medicaid expansion on private health insurance markets using data on the plans offered through the health insurance ACA Exchanges. They employed geographic matching to compare premiums for private plans in neighboring counties that straddled expansion and non-expansion states, and found that premiums of marketplace plans are 11% lower in Medicaid expansion states, while controlling for demographic and health characteristics as well as measures of healthcare access. These results are consistent with evidence on the composition of the private insurance risk pool in expansion versus non-expansion states, and the associated differences in expected health spending.

Polsky et al. (2016), used data from all the silver plans offered in the 2014 health insurance exchanges in the fifty states and the District of Columbia, in order to estimate the association between the breadth of a provider network and plan premiums, since the introduction

of health insurance marketplaces under the Affordable Care Act has been associated with the growth of restricted provider networks. They found that within a market, for plans of otherwise equivalent design, and controlling for issuer-specific pricing strategy, a plan with an extra-small network had a monthly premium that was 6.7% less expensive than that of a plan with a large network. They concluded that since narrow networks remain an important strategy available to insurance companies to offer lower-cost plans on health insurance marketplaces, the success of health insurance coverage expansions may be tied to the successful implementation of narrow networks.

3. Empirical Methodology

I utilized individual consumer-level data from the California Affordable Care Act (Obamacare) Exchange, procured from Dr. Evan Saltzman at Emory University, in order to study my research questions. Specifically, the data contains California health insurance records from the 2014 through 2019 plan years, and, has a sample size of about 10 million enrollee-year records. Californian counties are segregated into 19 different rating areas or health insurance markets, as visible in figure 1 in the Appendix. Three markets in California, namely Santa Clara County, Los Angeles County East, and Los Angeles County West, provided a public insurance option, while 16 did not. Contra Costa County only offered a public option in 2014, which is why I did not include it in my analysis. Contra Costa eliminated the public option since the municipality was unable to comply with some of the regulations of the offering in terms of providing the same product both within and outside the ACA exchange. The dataset contains 43 different variables including enrollment year, premium, deductible, and metal level, as well as demographic variables such as race, gender, and income level.

Metal-level or tier is an important variable to elaborate on, since there are four tiers of health insurance plans, named after metals, available on the market including bronze, silver, gold, and platinum. The tiers differ based on how the cost of healthcare services is split between the consumer and his or her insurer. Bronze plans typically have the lowest premiums, and the highest out-of-pocket expenses, while platinum plans typically have the highest premiums, and the lowest out-of-pocket expenses. Silver insurance plans are the most commonly chosen because consumers who are eligible for cost sharing reductions (CSRs), which are discounts that lower the amount that one has to pay for co-payments, deductibles, and co-insurance, must purchase a silver plan in order to receive CSRs, which is why my analysis is based on silver premiums.

Moreover, I conducted a multiple linear regression analysis in order to investigate and examine my research questions, particularly, by engaging one primary overall specification, as well as eight sub-specifications, with seven excluding one demographic variable each, and one jointly excluding two demographic variables, in order to discern their impact (or lack thereof) on the dependent variable. I determined that the exclusion of the year fixed effects, as well as both the year and market fixed effects together, had the most significant impact on the R^2 or coefficient of determination of the regression model, since the Year variable, which is a proxy for premium changes, explained the majority of the variation in the likelihood of consumers who elect the public option.

4. Descriptive Data & Summary Statistics

Consolidating the descriptive data and summary statistics of the Obamacare Exchange dataset was crucial to my analysis, since it helped me explore the characteristics of the data, and the demographics of each insurer's consumer base, as well as delineate many paramount and

vital trends in California's health insurance market. The three California counties where a public health insurance option is available include Santa Clara County (Rating Area 7), Los Angeles County East (Rating Area 15), and Los Angeles County West (Rating Area 16).

Tables 1 through 3 below spotlight the demographic breakdown of each insurance provider's consumer base in each of the three above-mentioned counties.

Table 1: Santa Clara County - Insurer Breakdown by Consumer Demographic Characteristics

	Anthem	Blue Shield	Health Net	Kaiser	Small Insurer	Public Insurer	Overall
Total Enrollment	156,961	47,981	6,749	168,669	NA	52,536	432,896
Bronze	41%	22%	42%	36%	NA	26%	35%
Silver	54%	61%	41%	46%	NA	64%	53%
Gold	3%	11%	6%	11%	NA	5%	7%
Platinum	1%	6%	4%	6%	NA	2%	4%
HMO	5%	6%	NA	100%	NA	100%	53%
PPO	95%	94%	100%	NA	NA	NA	47%
2014	24%	16%	46%	8%	NA	4%	15%
2015	25%	21%	17%	14%	NA	5%	18%
2016	22%	19%	19%	17%	NA	8%	18%
2017	15%	16%	18%	19%	NA	17%	17%
2018	8%	15%	NA	22%	NA	31%	17%
2019	6%	12%	NA	20%	NA	36%	15%
138% FPL or less	4%	3%	2%	3%	NA	3%	3%
138% FPL to 150% FPL	13%	11%	7%	10%	NA	11%	11%
150% FPL to 200% FPL	29%	21%	19%	26%	NA	32%	27%
200% FPL to 250% FPL	18%	12%	14%	18%	NA	21%	17%
250% FPL to 400% FPL	26%	27%	33%	30%	NA	26%	28%
400% FPL or greater	11%	26%	24%	13%	NA	7%	13%
Subsidized	87%	72%	73%	84%	NA	91%	84%

Unsubsidized	13%	28%	27%	16%	NA	9%	16%
0 to 17 Years	5%	12%	12%	7%	NA	3%	6%
18 to 25 Years	12%	9%	11%	11%	NA	13%	11%
26 to 34 Years	17%	13%	14%	19%	NA	16%	17%
35 to 44 Years	15%	14%	13%	15%	NA	15%	15%
45 to 54 Years	24%	22%	22%	20%	NA	23%	22%
55 to 64 Years	26%	29%	27%	28%	NA	29%	27%
65+ Years	1%	1%	1%	1%	NA	1%	1%
Female	51%	53%	53%	52%	NA	51%	52%
Male	49%	47%	47%	48%	NA	49%	48%
Asian	56%	35%	32%	37%	NA	54%	46%
Black/African American	1%	1%	1%	2%	NA	2%	2%
Hispanic	14%	8%	12%	19%	NA	22%	16%
Other Race	8%	11%	10%	10%	NA	9%	9%
White	21%	45%	44%	32%	NA	13%	27%
English	73%	93%	92%	86%	NA	66%	80%
Spanish	5%	1%	2%	5%	NA	10%	5%
Other Language	21%	6%	7%	9%	NA	24%	15%
New Enrollee	50%	48%	70%	46%	NA	49%	49%
Existing Enrollee	50%	52%	30%	54%	NA	51%	51%

Table 2: Los Angeles County East - Insurer Breakdown by Consumer Demographic Characteristics

	Anthem	**Blue Shield**	**Health Net**	**Kaiser**	**Small Insurer**	**Public Insurer**	**Overall**
Total Enrollment	77,557	445,104	330,227	172,933	48,330	115,974	1,190,125
Bronze	35%	12%	8%	38%	47%	23%	19%
Silver	55%	78%	82%	42%	47%	64%	70%
Gold	4%	6%	5%	10%	3%	10%	7%
Platinum	4%	3%	3%	8%	1%	3%	4%
HMO	31%	7%	88%	100%	89%	100%	57%
PPO	69%	93%	12%	NA	11%	NA	43%

2014	24%	15%	18%	9%	1%	9%	14%
2015	27%	15%	22%	16%	2%	7%	17%
2016	32%	21%	14%	15%	5%	5%	17%
2017	16%	17%	12%	18%	64%	13%	17%
2018	NA	18%	16%	21%	14%	33%	18%
2019	NA	15%	17%	22%	15%	34%	17%
138% FPL or less	4%	4%	5%	5%	4%	5%	4%
138% FPL to 150% FPL	15%	26%	24%	12%	20%	20%	22%
150% FPL to 200% FPL	31%	30%	38%	30%	35%	35%	33%
200% FPL to 250% FPL	16%	13%	15%	17%	18%	17%	15%
250% FPL to 400% FPL	22%	17%	14%	24%	16%	17%	18%
400% FPL or greater	12%	10%	5%	12%	7%	7%	8%
Subsidized	86%	88%	93%	84%	91%	92%	89%
Unsubsidized	14%	12%	7%	16%	9%	8%	11%
0 to 17 Years	6%	6%	3%	7%	4%	5%	5%
18 to 25 Years	12%	11%	13%	11%	12%	12%	12%
26 to 34 Years	18%	16%	14%	22%	20%	14%	17%
35 to 44 Years	15%	16%	15%	15%	15%	14%	15%
45 to 54 Years	24%	24%	26%	20%	24%	25%	24%
55 to 64 Years	25%	26%	28%	24%	24%	29%	26%
65+ Years	0%	1%	1%	0%	1%	1%	1%
Female	52%	53%	52%	52%	49%	52%	52%
Male	48%	47%	48%	48%	51%	48%	48%
Asian	35%	49%	44%	23%	39%	26%	40%
Black/African American	2%	2%	2%	4%	2%	3%	2%
Hispanic	25%	14%	34%	37%	38%	45%	28%
Other Race	9%	9%	7%	9%	7%	8%	8%
White	28%	26%	13%	27%	14%	18%	21%
English	79%	72%	60%	87%	63%	65%	70%
Spanish	7%	3%	17%	8%	18%	23%	11%

Other Language	13%	25%	23%	5%	18%	11%	19%
New Enrollee	57%	43%	46%	50%	53%	47%	47%
Existing Enrollee	43%	57%	54%	50%	47%	53%	53%

Table 3: Los Angeles County West - Insurer Breakdown by Consumer Demographic Characteristics

	Anthem	**Blue Shield**	**Health Net**	**Kaiser**	**Small Insurer**	**Public Insurer**	**Overall**
Total Enrollment	177,915	309,411	327,496	286,133	181,233	162,559	1,444,747
Bronze	23%	15%	5%	41%	48%	25%	24%
Silver	65%	67%	85%	40%	45%	62%	62%
Gold	5%	11%	6%	10%	4%	10%	8%
Platinum	5%	6%	3%	7%	1%	3%	4%
HMO	40%	3%	93%	100%	76%	100%	67%
PPO	60%	97%	7%	NA	24%	NA	33%
2014	27%	16%	21%	7%	1%	12%	14%
2015	26%	18%	24%	15%	5%	7%	17%
2016	28%	17%	17%	15%	22%	5%	17%
2017	19%	17%	11%	19%	35%	11%	18%
2018	NA	17%	15%	23%	20%	31%	18%
2019	NA	15%	11%	20%	17%	33%	16%
138% FPL or less	4%	4%	5%	4%	4%	4%	4%
138% FPL to 150% FPL	15%	16%	21%	10%	13%	16%	16%
150% FPL to 200% FPL	29%	25%	39%	27%	34%	37%	32%
200% FPL to 250% FPL	14%	13%	14%	17%	19%	19%	16%
250% FPL to 400% FPL	22%	21%	14%	25%	20%	17%	20%
400% FPL or greater	17%	21%	7%	15%	10%	7%	13%
Subsidized	81%	76%	92%	80%	87%	91%	84%
Unsubsidized	19%	24%	8%	20%	13%	9%	16%
0 to 17 Years	7%	9%	3%	6%	4%	4%	6%
18 to 25 Years	10%	8%	11%	10%	11%	11%	10%
26 to 34	21%	22%	15%	26%	23%	17%	21%

Years							
35 to 44 Years	17%	20%	15%	17%	16%	15%	17%
45 to 54 Years	22%	21%	26%	19%	23%	24%	22%
55 to 64 Years	22%	20%	28%	21%	23%	29%	24%
65+ Years	0%	0%	1%	0%	1%	1%	1%
Female	52%	53%	51%	51%	48%	51%	51%
Male	48%	47%	49%	49%	52%	49%	49%
Asian	15%	14%	23%	13%	12%	13%	15%
Black/African American	4%	3%	4%	6%	4%	4%	4%
Hispanic	20%	12%	43%	30%	48%	54%	33%
Other Race	11%	13%	7%	10%	8%	7%	10%
White	50%	59%	23%	40%	28%	21%	38%
English	89%	94%	62%	89%	70%	62%	79%
Spanish	7%	2%	27%	9%	27%	33%	16%
Other Language	5%	4%	11%	2%	3%	4%	5%
New Enrollee	56%	45%	47%	46%	47%	44%	47%
Existing Enrollee	44%	55%	53%	54%	53%	56%	53%

Since the California public option only offers an HMO plan, I decided to omit the PPO insurers from the following graphs and tables in this paper for the purpose of visual clarity and aesthetic simplicity.

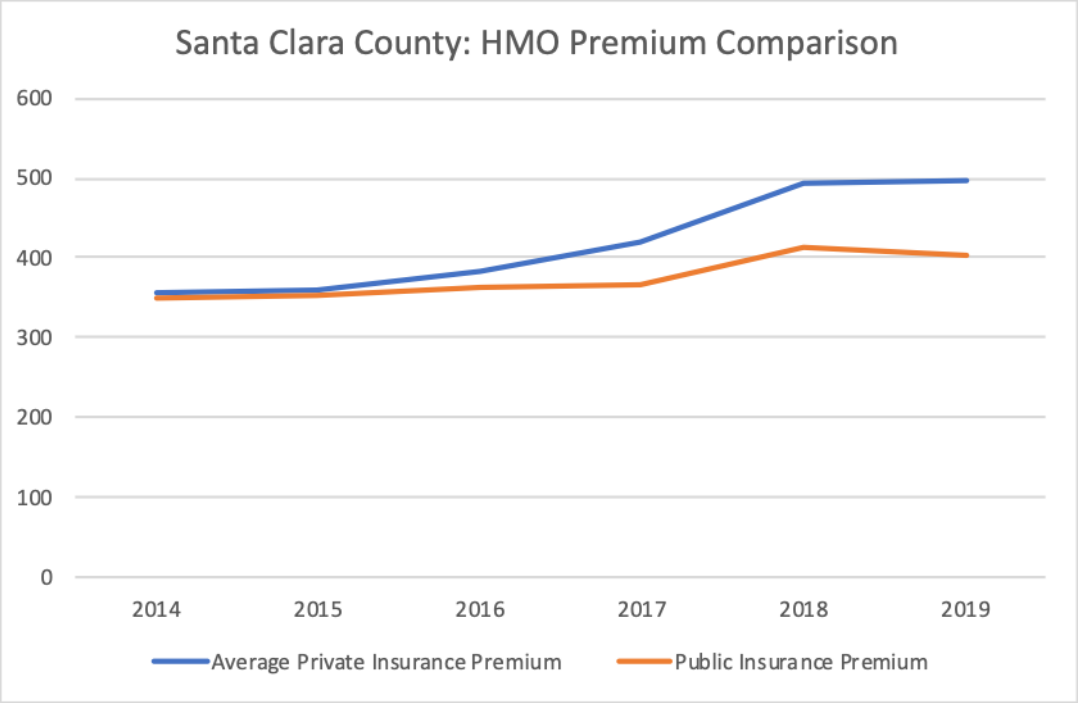
Graphs 1, 3, and 5 below individually illustrate changes in each of the 3 aforementioned counties' average HMO private insurance premiums and the public insurance premium from 2014 to 2019. Graphs 2, 4, and 6 below individually demonstrate changes in each of the 3 aforementioned counties' HMO private insurer market share and the public insurer market share from 2014 to 2019. I have placed each county's market share changes graph directly below its premium changes graph because premium changes directly impact insurers' market shares. It is critical to note that the visible patterns and trends in the premium and market share changes

graphs in each of the 3 counties are relatively similar on average, despite the exact values and numbers being different. The descriptive similarity in consumption patterns across each of the 3 counties is supported by the relatively insignificant impact of the Region variable on the regressand in the multiple linear regression output detailed later in the paper.

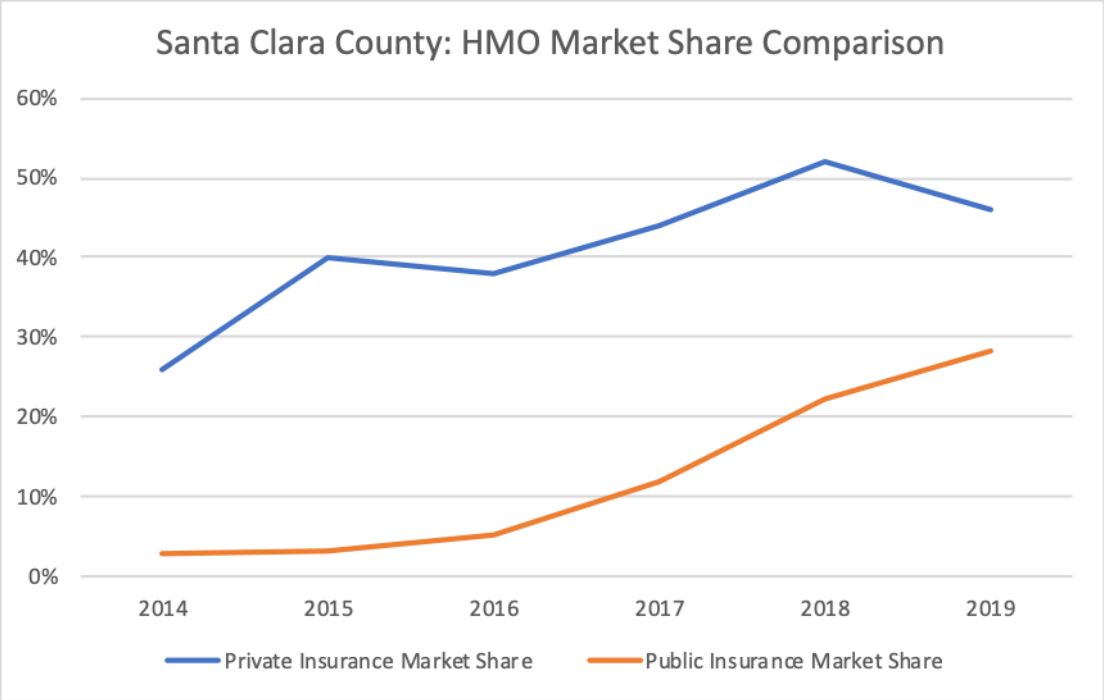
The important takeaways from the below graphs are that the public insurer's premium is lower than the average premium during each year, in each of the 3 counties, except in 2014, when the public insurer's premium was slightly higher in Los Angeles County East, as visible in Graph 3. However, that could have simply been due to a statistical discrepancy or reporting error, or because it was the first year that the public option was introduced to the California market. Furthermore, in each of the 3 counties, all insurers' HMO premiums drastically increased in 2017, and slightly decreased or stabilized from 2018 onwards. 2017 is a pivotal year since one of the private insurers, Anthem eliminated its HMO offering, and a new one, Blue Shield, introduced an HMO option in the California market in 2017. Unlike Blue Shield, which had the highest HMO private insurance premium, the public insurer was able to absorb a lot of Anthem's market share, presumably due to its lower relative premium. Additionally, the public insurance premium increases at a much slower pace than the average premium in each of the 3 counties, thus, contributing to a larger gap between the two premiums every succeeding year.

Santa Clara County

Graph 1: Santa Clara County – HMO Premium Comparison

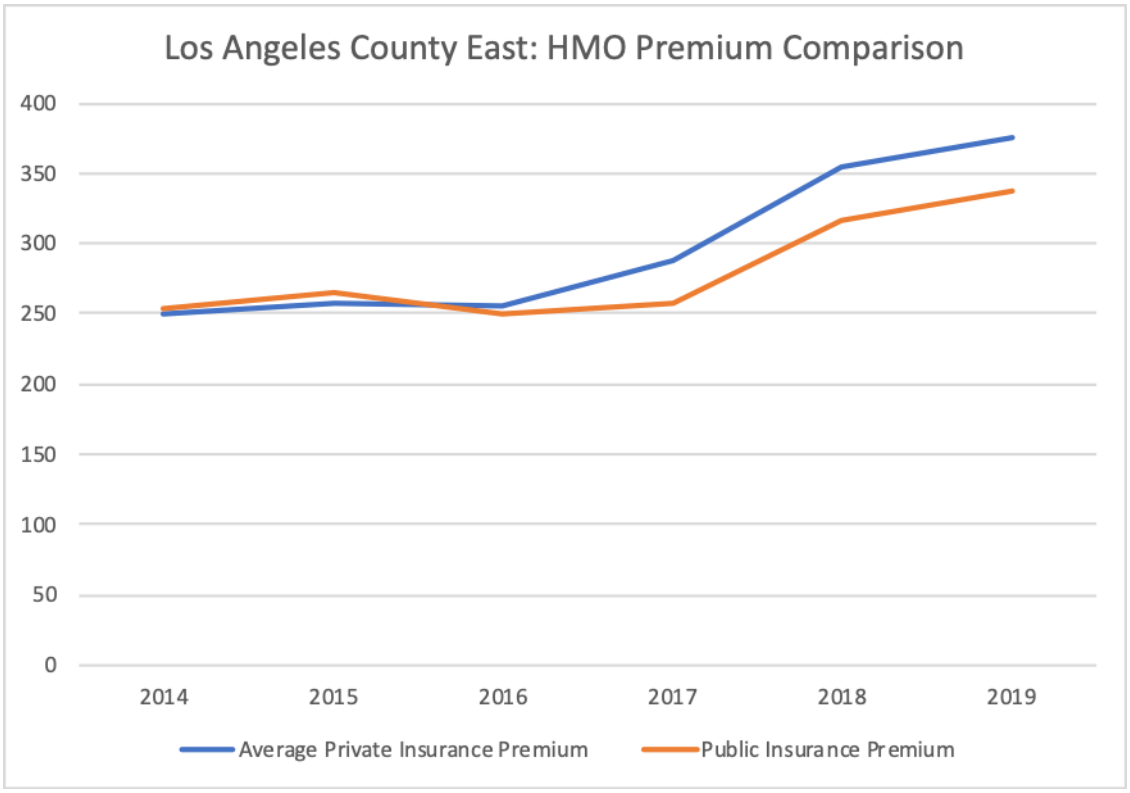


Graph 2: Santa Clara County – HMO Market Share Comparison

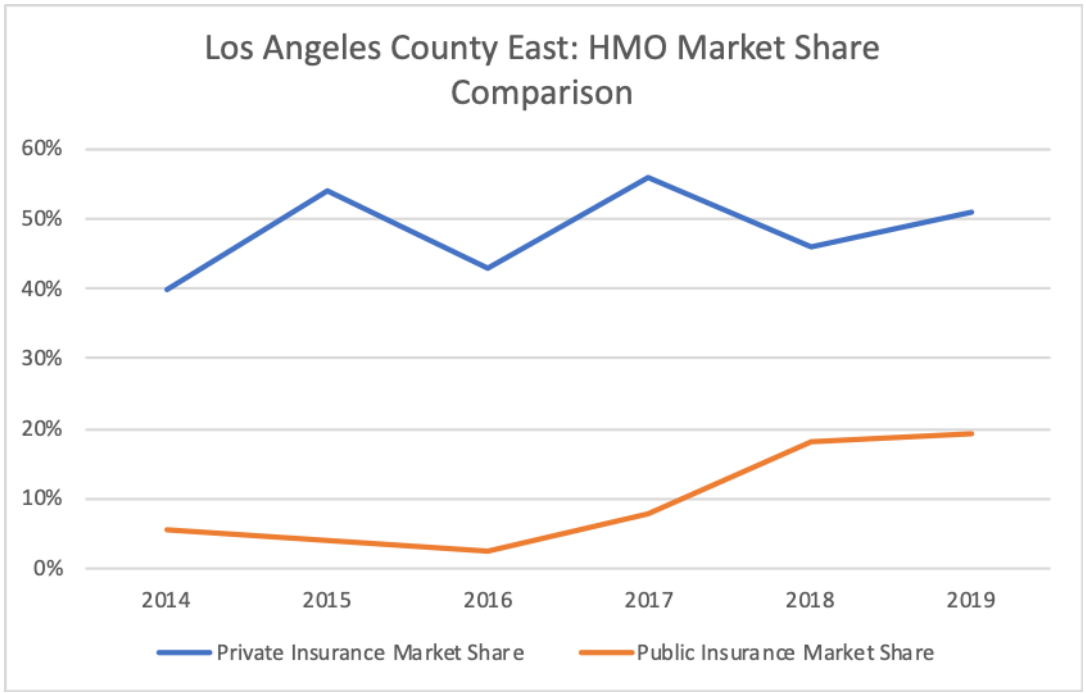


Los Angeles County East

Graph 3: Los Angeles County East – HMO Premium Comparison

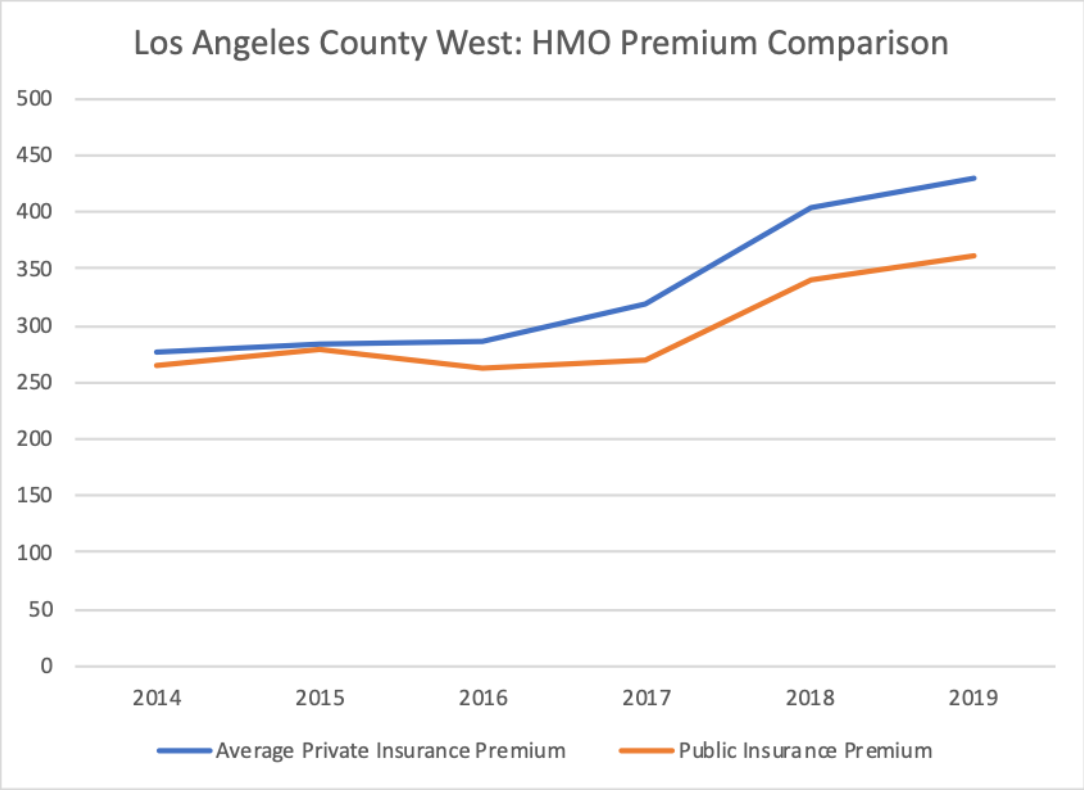


Graph 4: Los Angeles County East – HMO Market Share Comparison

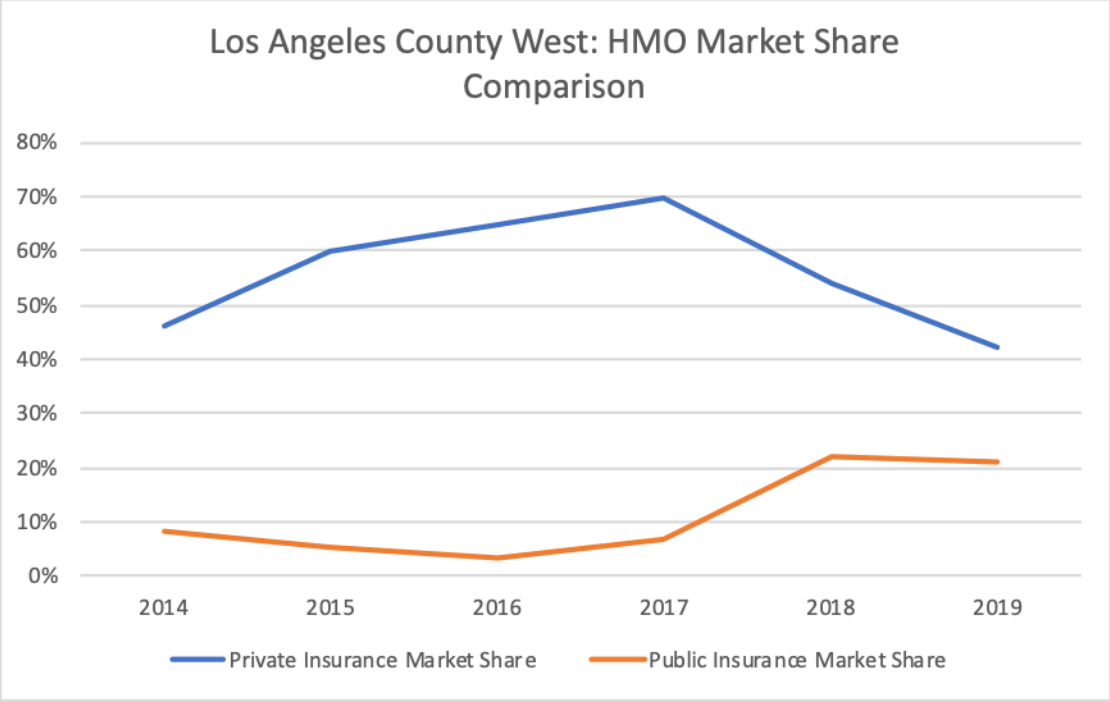


Los Angeles County West

Graph 5: Los Angeles County West – HMO Premium Comparison

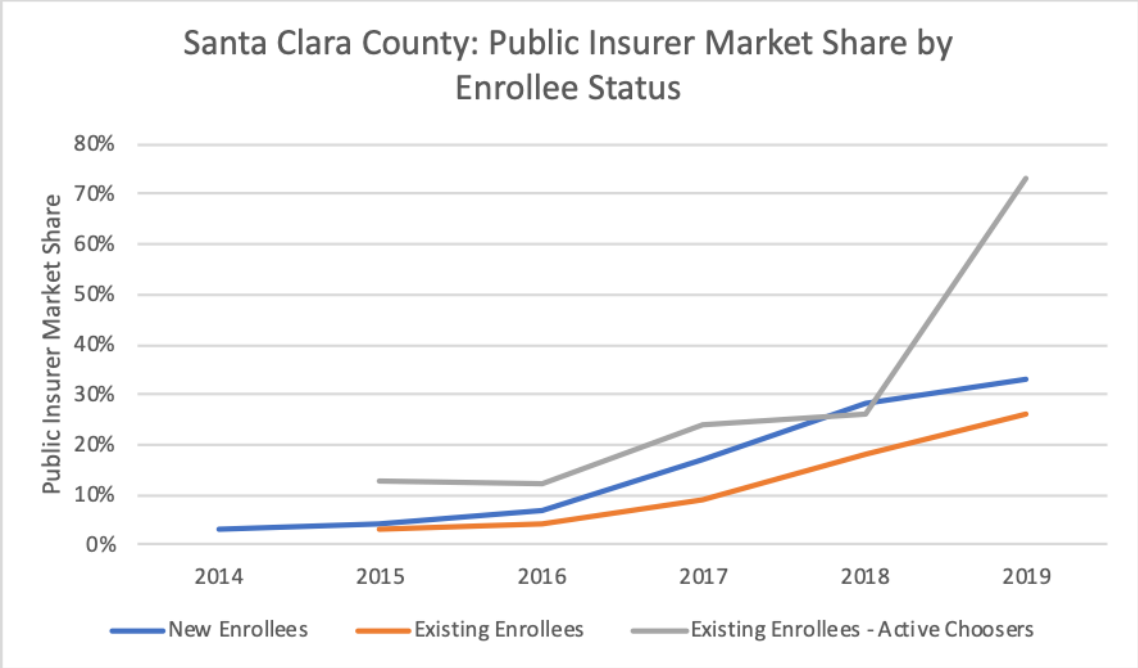


Graph 6: Los Angeles County West – HMO Market Share Comparison

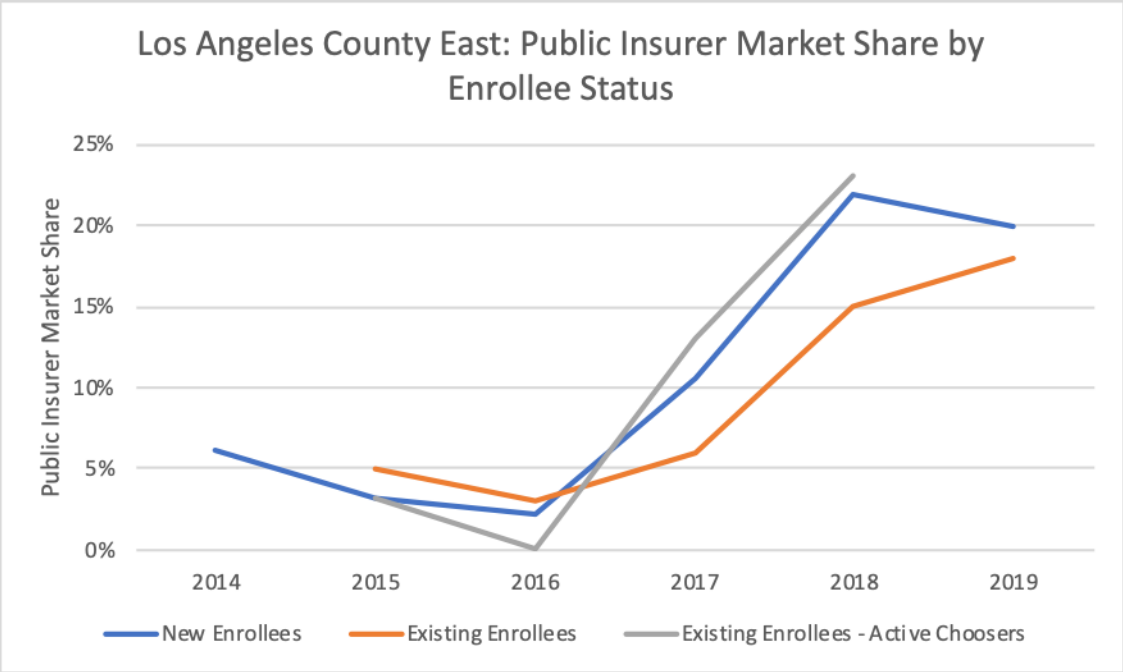


Graphs 7 through 9 below illustrate the public insurer market share in each of the three counties with a public option, from 2014 to 2019, broken up by the consumers' enrollee status. The essential revelations are that the public insurance market share drastically increased from 2016 onwards in LA counties East and West, and from 2018 onwards in Santa Clara County, fueled by existing enrollees that are active choosers as well as new enrollees. Moreover, active choosers opted for the public plan to a much larger extent than new enrollees in Santa Clara County, and to roughly an equal extent in Los Angeles counties East and West. Los Angeles counties East and West appear to have a much larger share of existing public insurance enrollees than Santa Clara County, especially from 2017 onwards.

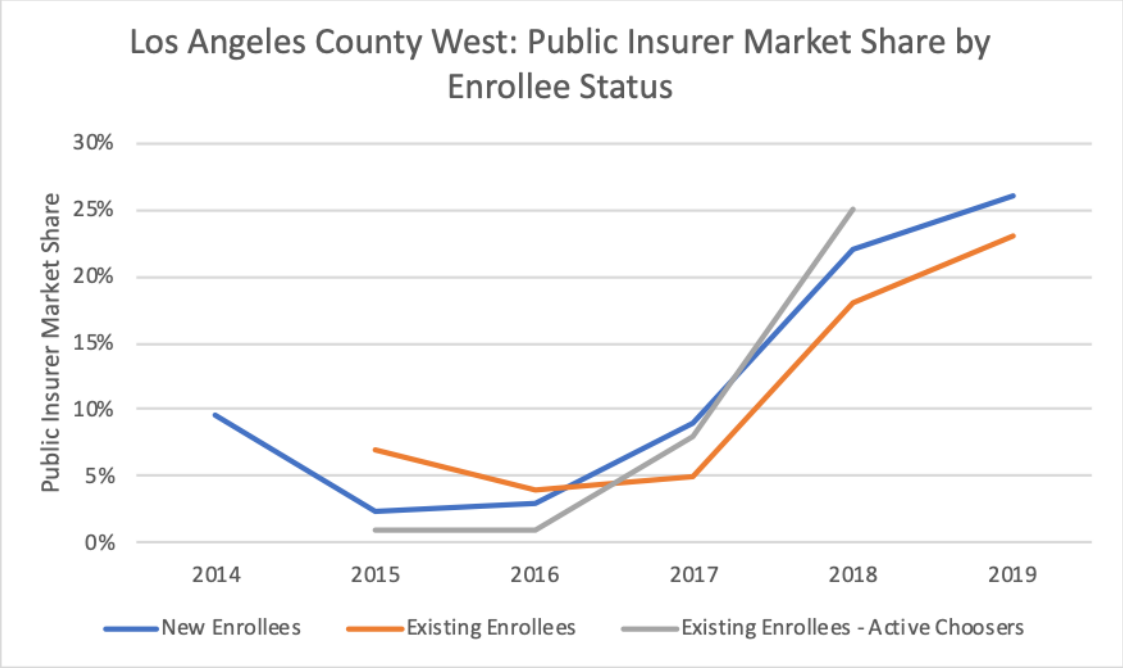
Graph 7: Santa Clara County – Public Insurer Market Share by Enrollee Status



Graph 8: Los Angeles County East – Public Insurer Market Share by Enrollee Status



Graph 9: Los Angeles County West – Public Insurer Market Share by Enrollee Status



5. Regression Equations & Model

1. Complete Regression

$$ChosePublic = \beta_0 + \beta_1 IncomeLevel + \beta_2 Gender + \beta_3 EnrolleeStatus + \beta_4 Race + \beta_5 Age + \beta_6 HouseholdSize + \beta_7 Year + \beta_8 Region + u$$

My complete multiple linear regression model includes eight independent variables, however it is important to note that each of its betas, except for those of household size, and the binary variable gender, is a vector of betas, since each of the independent variables except for gender and household size, is broken up into different categories.

The first explanatory variable is income level, which highlights the economic standing of each consumer in terms of his income's percentage of the federal poverty level (FPL), and includes the categories '250% FPL or less,' '250% to 400% FPL,' and '400% FPL or greater.' The second variable is gender, which is a binary or dummy variable for the consumer's sex assigned at birth. The third variable is enrollee status, which asserts whether a consumer is an existing enrollee with an established health insurance plan, an existing enrollee who is shopping for a new plan, or an entirely new enrollee. The fourth variable is race, which includes the Asian, Hispanic, Black/African American, White, and Other Race categories. The fifth variable is age, which includes interval ranges starting from 0 to 17 years, all the way up to 65 years and older. The sixth variable is household size, which cites the number of members in a particular consumer's primary home. The seventh variable is the year in which the consumer is purchasing a health insurance plan, and it spreads across 2014 to 2019. The eighth and final variable is the region, rating area, or county wherein the consumer resides, and encompasses Los Angeles County West, Los Angeles County East, and Santa Clara County. Therefore, by using a cross-sectional approach, I examined the impact that these eight variables have on the dependent variable, namely the likelihood of choosing the public health insurance option, through a multiple linear regression analysis.

I chose to dissect these particular explanatory variables because they were readily available in the dataset that I used, and because I believe them to have the most relevant, and noteworthy impact on consumer choices and preferences based on my research and the aforementioned literature review. Additionally, I chose to eliminate the Language variable, which indicates the primary spoken language of the consumer, and includes the categories English, Spanish, and Other Language, from my regression model because there was a high degree of correlation and collinearity between it and the Race variable, presumably since the majority of Spanish speakers are part of the Hispanic population, and thus, gravitated more towards the public plan than English speakers or speakers of other languages.

Furthermore, I estimated eight additional sub-specifications, with seven excluding one demographic variable each, and one jointly excluding the Year and Region variables, in order to discern their impact, or lack thereof, on the R^2 or the coefficient of determination estimates. The exclusion of the year fixed effects, as well as both the year and market fixed effects together, had the most significant impact on the R^2 estimate, which is why I included those two models below, in order to compare their estimates and results with those of my complete regression model.

2. Regression Excluding Year & Market Fixed Effects

$$ChosePublic = \beta_0 + \beta_1 IncomeLevel + \beta_2 Gender + \beta_3 EnrolleeStatus + \beta_4 Race + \beta_5 Age + \beta_6 HouseholdSize + u$$

3. Regression Excluding Year Fixed Effects

$$ChosePublic = \beta_0 + \beta_1 IncomeLevel + \beta_2 Gender + \beta_3 EnrolleeStatus + \beta_4 Race + \beta_5 Age + \beta_6 HouseholdSize + \beta_7 Region + u$$

6. Results & Findings

It is pertinent to note that I ran each of the 3 regressions only in Los Angeles County East (Rating Area 15), Los Angeles County West (Rating Area 16), and Santa Clara County (Rating

Area 7), and that there were some missing data values in the Race variable, which explains why 7.5 million fewer observations were incorporated during my regression analysis. Furthermore, all of the regressors are statistically significant, presumably due to the large number of observations.

The Year explanatory variable, which indicates year fixed effects, and thereby is a proxy for measuring premium changes over time, is the most significant factor influencing consumer choices in health insurance, as evidenced by the regression output, which portrays an R^2 of 0.024 in the regression without the year fixed effects, referred to forthwith as regression 3, while the R^2 of the complete regression including the Year variable, referred to as regression 1 henceforth, is 0.085. This finding reveals that without the year fixed effects, only 2.4% of the observed variation can be explained by the regression model's remaining inputs, however, the inclusion of the year fixed effects helps explain 8.5% of the observed variation in the response variable. Hence, the miniscule proportion of the variance in the dependent variable explained by the remaining independent variables in the regression model emphasizes the importance of the year fixed effects in the model. Furthermore, the regression 1 output demonstrates that consumers become more inclined to opt for the public plan over time. This discovery coincides with premium changes, since the public insurance premium increased at a slower rate than the average private insurance premium, and remained lower than the average private insurance premium at all points in time as highlighted in graphs 1, 3, and 5 above. This finding leads me to suspect that consumers are highly premium-sensitive, and that their consumption decisions are greatly influenced by premium changes, however, further analysis is necessary to empirically cement this claim. Additionally, the year 2014 comprises the discarded category in order to prevent multicollinearity, since no existing enrollees or existing enrollees that are active choosers existed in this year, and all consumers were new enrollees by default.

The regression output demonstrates that the exclusion of the Region independent variable, which denotes market fixed effects, does not affect the R^2 of the model significantly, since it only decreases to 0.023 in the regression ignoring both the year and market fixed effects, referred to forthwith as regression 2. Furthermore, consumption decisions are not largely impacted by the Region variable, since consumers' likelihood of choosing the public option is relatively similar in both Los Angeles County East, and Los Angeles County West in both regressions 1 and 3. Interestingly, as evidenced by regression 2, only 0.1% less of the variation in the dependent variable is explained by the exclusion of the market fixed effects. In addition, Santa Clara County represents the dropped category in order to avert multicollinearity.

The Income Level covariate has a notable impact on consumer choice in each of the 3 regressions. Specifically, consumers become marginally more inclined to elect the public option as their income level decreases, or the closer they are to the federal poverty level. Additionally, the impact of the Income Level covariate on the dependent variable becomes larger in regressions 2 and 3. This finding implies that when year fixed effects, and to a much smaller extent, market fixed effects, are eliminated from the regression equation, a larger amount of the variation in the dependent variable is explained by income level fluctuations. Moreover, the '250% FPL and below' tier represents the discarded category in order to prevent multicollinearity.

The Gender predictor variable has a negligible impact on consumer choice in each of the 3 regressions, since the regression output highlights that the inclination of both males and females towards the public option is roughly equal. This finding illustrates that gender does not play any role in determining consumer health insurance preferences. Since Gender is a binary variable, it was unnecessary to drop a category from it to avoid multicollinearity.

The Enrollee Status covariate has a noteworthy influence on consumer election in each of the 3 regressions because the regression output illustrates that existing enrollees who are actively choosing a new plan, and new enrollees are more prone to electing the public option. Moreover, the impact of the Enrollee Status covariate on the dependent variable increases in regressions 2 and 3. This discovery illustrates that when year fixed effects, and to a more miniscule extent, market fixed effects, are discarded from the regression equation, a greater amount of the variation in the dependent variable is explained by the enrollee status of the consumers. Furthermore, the existing enrollees who are not shopping for a new health insurance plan constitutes the discarded category in order to evade multicollinearity.

The Race explanatory variable has a conspicuous impression on consumer preferences in each of the 3 regressions because the regression output illuminates that Black and Hispanic consumers are more willing to opt for the public option than those of other races, excluding White consumers, which constitutes the dropped category in the Race variable to avert multicollinearity. This discovery illuminates that racial minorities are more inclined towards the public option. Furthermore, the Race predictor variable explains slightly more of the variation in the response variable in regressions 2 and 3 due to the exclusion of the year and market fixed effects. It is important to note that much of the effect of race on consumption choices is not explained by income level or any of the other explanatory variables, however, I suspect that there is a slight correlation or collinearity between the Hispanic population, and the low-income population. One conjecture I have is that racial minorities, regardless of their income level, gravitate more towards the public plans, because they are specifically marketed and advertised towards them, and especially towards the Hispanic population through Spanish ads. This would

also help explain the high degree of correlation between the Race and Language variables, and why Spanish speakers are more inclined towards the public option.

The Age covariate has an inconsequential impact on consumer preferences in each of the 3 regressions, with no consistent consumption trends or patterns noticeable as consumers' age progresses, based on the regression output. The category of consumers aged 55 years and older was dropped from the regression model in order to avert multicollinearity.

The Household Size predictor variable has a negligible impact on consumer inclinations across all 3 regressions, since the regression output highlights that households of different sizes have similar consumption patterns, and that an additional household member does not significantly sway consumer health insurance elections. The Household Size variable explains a slightly greater amount of the variation in the response variable when market and year fixed effects are omitted in regressions 2 and 3, than it does in the complete regression. Since the Household Size variable was not broken up into sub-categories, it inadvertently avoids multicollinearity.

It is evident that the year fixed effects, which coincide with, and signal premium effects, explain the majority of the variation in the response variable, and that when the Year variable is dropped from the regression equation, most of the other explanatory variables are compelled to explain marginally more of the variation in the dependent variable than they previously did in the complete regression 1.

Table 1: Multiple Linear Regression Output - Influences on Health Insurance Policy Choice

	Dependent Variable		
	Chose the Public Option		
	Complete	**Excluding Year & Market Fixed Effects**	**Excluding Year Fixed Effects**

	(1)	(2)	(3)
Income Level			
250% FPL to 400% FPL	-0.018***	-0.006***	-0.010***
	(0.001)	(0.001)	(0.001)
400% FPL or greater	-0.043***	-0.034***	-0.038***
	(0.001)	(0.001)	(0.001)
Gender			
Male	0.004***	0.004***	0.003***
	(0.0004)	(0.0004)	(0.0004)
Enrollee Status			
Existing Enrollees - Active Choosers	0.060***	0.115***	0.115***
	(0.002)	(0.002)	(0.002)
New Enrollees	0.026***	-0.003***	-0.004***
	(0.0004)	(0.0004)	(0.0004)
Race			
Asian	0.018***	0.019***	0.021***
	(0.001)	(0.001)	(0.001)
Black	0.050***	0.050***	0.050***
	(0.001)	(0.001)	(0.001)
Hispanic	0.103***	0.102***	0.104***
	(0.001)	(0.001)	(0.001)
Other Race	0.020***	0.024***	0.025***
	(0.001)	(0.001)	(0.001)
Age			
0 to 17 years	-0.021***	-0.011***	-0.011***
	(0.001)	(0.001)	(0.001)
18 to 25 years	0.002**	0.002***	0.001
	(0.001)	(0.001)	(0.001)
26 to 34 years	-0.022***	-0.020***	-0.020***
	(0.001)	(0.001)	(0.001)
35 to 44 years	-0.013***	-0.013***	-0.014***
	(0.001)	(0.001)	(0.001)
45 to 54 years	0.001**	-0.003***	-0.002***
	(0.001)	(0.001)	(0.001)
Household Size	-0.003***	-0.007***	-0.006***

	(0.0002)	(0.0002)	(0.0002)
Year			
2015	-0.018***		
	(0.001)		
2016	-0.023***		
	(0.001)		
2017	0.021***		
	(0.001)		
2018	0.140***		
	(0.001)		
2019	0.176***		
	(0.001)		
Region			
Los Angeles County East	-0.053***		-0.047***
	(0.001)		(0.001)
Los Angeles County West	-0.029***		-0.025***
	(0.001)		(0.001)
Constant	0.063***	0.094***	0.123***
	(0.001)	(0.001)	(0.001)
Observations	2,420,104	2,420,104	2,420,104
R^2	0.085	0.023	0.024
Adjusted R^2	0.085	0.023	0.024
Residual Std. Error	0.300 (df = 2420081)	0.311 (df = 2420088)	0.310 (df = 2420088)
F Statistic	10,261.720*** (df = 22; 2420081)	3,783.696*** (df = 15; 2420088)	3,972.768*** (df = 15; 2420088)
Note:			*p**p***p<0.01

7. Inference & Hypothesis Testing

I conducted a hypothesis test for the overall joint significance of my regression, with the Null Hypothesis H_0 stating that the independent variables in my model are not jointly significant to the regression, and the Alternative Hypothesis H_1 stating that the independent variables in my model are jointly significant to the regression.

H_0: $B_j = 0$
H_1: $B_j \neq 0$
$j = \{1 \ldots 7\}$

The Null Hypothesis is rejected for each of the 3 regressions, since the F Statistic of the complete regression 1 is 10,261.72, which is larger than the corresponding critical value of 1.42, the F Statistic of regression 2 is 3,783.70, which is larger than the corresponding critical value of 1.49, and the F Statistic of regression 3 is 3,972.77, which is larger than the corresponding critical value of 1.49. Thus, the explanatory variables in all 3 of the aforementioned regression models are jointly significant at the 1%, 5%, and 10% significance levels.

8. Limitations & Shortcomings

One of the major limitations of this study is that I was unable to utilize a McFadden logit discrete choice model in order to study my research question as a result of a dearth of time, and instead, had to resort to conducting a multiple linear regression analysis. The logit choice model would have allowed me to more accurately predict and anticipate health insurance consumption preferences based on demographic characteristics. My inability to employ a logit choice model prevented the inclusion of a 'premium' regressor in my regression model, because of which I had to resort to utilizing the Year variable as a proxy for premium changes. The logit choice model would have also helped me factor in variation within the state of California, since some constituents are exposed to the public option, while others are not, and allowed me to calculate consumers' willingness to pay (WTP) for a public versus a private health insurance plan.

Additionally, since the dataset that I employed only includes observations from 2014 to 2019, it ignores the most recent health insurance consumption data from the past 2 years, which could have possibly swayed my findings. However, the involuntary exclusion of the past two

years from the dataset bestowed the benefit of my analysis not being marred or confused by the onset of the COVID-19 pandemic in March, 2020, which generated a global healthcare crisis.

Finally, since the Obamacare Exchange dataset only includes data from California, there is a limit to the degree of generalizability of the results of my study to the rest of the country, due to demographic differences in the populations.

9. Possible Future Extensions

A worthwhile extension of my research would be the addition of more explanatory variables, including sexual orientation, religious preference, birth country, highest level of education, profession, immigration status, and marital status, to the regression model, since they would help present a clearer picture of the demographics that lean more towards public health insurance, and might help explain more of the variation in the regressand.

Although my research is comprehensively descriptive, it provides a strong foundation to base additional research on, in order to dissect the nuances underlying the particular demographic trends and consumer preference patterns discovered in my paper. Some worthwhile research topics that can be pursued in the future by utilizing my research as scaffolding, include examining whether there is a negative or biased connotation or prejudice associated with the public plan about its quality and extent of coverage, delving into the extent of the public's level of education and knowledge regarding the nature of the public option and what it entails, surveying if changing the branding of the public option, and giving it a unique name will sway consumer preference in any way, ascertaining if there is a difference between consumers' willingness to pay (WTP) for a public versus a private plan, studying whether tailored advertising techniques or schemes have been utilized in order to target or market the public option to certain demographics, inspect if inertia or the tendency to stick with one's existing

provider because of the financial or bureaucratic switching costs plays a role in swaying consumer preferences, and investigating how the public option might fare in a different state or region of the United States. However, the last suggestion is clearly conditional on the public option actually being introduced in a different state market in the United States. A final possible area of study is conducting a cross-sectional geographic analysis in order to compare the public option offering in different regions or countries around the world so as to divine the factors or situations that contributed to its success or failure in different circumstances, and in populations with differing demographic characteristics.

In addition, in a future revision of this paper, I would endeavor to employ a McFadden logit discrete choice model instead of a multiple linear regression model, in order to better understand the nuances underlying health insurance consumption patterns, and thereby, be able to better predict consumers' health insurance choices, based on their demographic characteristics. The logit choice model would also allow me to include a 'premium' regressor in my regression model, and thus, examine the variation within the California market, as several different types of private and public insurance plans exist, and only the constituents in Santa Clara County, Los Angeles County East, and Los Angeles County West are exposed to the public option, while those in the remaining Californian counties are not.

10. Conclusive Reflections

Ultimately, based on my findings, I conclude that age does not appear to play any role in influencing consumer health insurance preferences. Secondly, my results highlight that as consumers' income level rises, they become less willing to elect the public plan. Thirdly, my findings illustrate that gender does not play a role in swaying consumer health insurance elections, and that women are equally as likely to lean towards the public plan, as men. Fourthly,

I discovered that racial minorities, especially the Black and Hispanic populations, are more willing to choose the public option. Finally, I found that active choosers are examined to be more inclined towards the public plan than new enrollees.

It is evident that year fixed effects, which signal premium changes, are the most potent drivers behind consumer health insurance preferences since the Year explanatory variable explains most of the variation in the response variable. However, race, and income level also play a large role in helping anticipate and predict consumer health insurance choices when a public option is available. This finding is germane for private health insurance companies, since it can help them modify their current advertising practices, and focus on highlighting the competitive nature of their premiums. By a similar token, private companies that already have low insurance premiums that are comparable to the public plan premium, should target consumers of lower income strata, as well as racial minority populations, since they are suspected to be the most premium-sensitive, according to the Obamacare Exchange dataset employed in this paper. Furthermore, private companies can also modify their existing plan offerings to reduce their premiums in an effort to attract more consumers, and increase their consumer base.

The detections in this paper also have policy consequences for the U.S. government, since it informs legislators about the effectiveness of the public health insurance option in California, and the specific demographics that elect it the most. Thus, the U.S. government can continue to keep public plan premiums lower than average private insurance premiums, and advertise their lower premiums to lower-income and racial minority populations, in order to increase the public option's market share in California. The U.S. government can also consider implementing a national public option or expanding the public option to other American states in

an effort to increase the efficiency and competitiveness of health insurance markets by motivating lower private insurance premiums. A national public option would also help provide U.S. constituents with a more affordable health insurance option, and thereby, potentially decrease the proportion of uninsured Americans, and reduce the disproportionate impact of prohibitive healthcare coverage costs on low-income communities of color.

Lastly, it is clear that additional regressors are necessary in order to better explain the variation in public insurance election patterns. Albeit future work and study is necessary in order to more fully explicate the nuances and logical rationale underlying my findings, my thesis is still consequential and valuable, with far-reaching policy implications, especially from a descriptive analysis perspective.

11. Bibliography

Biden-Harris, Democrats. “Healthcare.” https://joebiden.com/healthcare/

Drake, C. “What are consumers willing to pay for a broad network health plan?: Evidence from covered California.” *Journal of Health Economics*. 2019. Volume 65, pp. 63 - 67.

“Health at a Glance 2017: OECD Indicators.” Organisation for Economic Co-operation and Development. 2017. https://www.oecd.org/unitedstates/Health-at-a-Glance-2017-Key-Findings-UNITED-STATES.pdf

Hawryluk, Markian. “The 1st public option health plan in the U.S. struggles to gain traction.” *NPR*. February 21, 2022. https://www.npr.org/sections/health-shots/2022/02/21/1081913184/the-first-public-option-health-plan-in-the-u-s-struggles-to-gain-traction#:~:text=Colorado%20and%20Nevada%2C%20which%20are,are%20likely%20to%20follow%20suit.

Orgera, K., Damico, A., and Tolbert, J. “Key Facts about the Uninsured Population.” *Kaiser Family Foundation*. 6 November, 2020. https://www.kff.org/uninsured/issue-brief/key-facts-about-the-uninsured-population/

Polsky, D., Z. Cidav, and A. Swanson. “Marketplace plans with narrow physician networks feature lower monthly premiums than plans with larger networks.” *Health Affairs*. October 2016.

Saltzman, E. “Demand for health insurance: Evidence from the California and Washington ACA exchanges.” *Journal of Health Economics*. Volume 63, January 2019, pp. 197 - 222.

Sen, A. P. and T. DeLeire. “How does the expansion of public health insurance affect risk pools and premiums in the market for private health insurance? Evidence from Medicaid and the affordable care act marketplaces.” *Health Economics*. 2018. Volume 27, pp. 1877 - 1903.

Tebaldi, P. “Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the ACA.” August 2017. BFI Working Paper.

“U.S. Health Care Coverage and Spending.” *Congressional Research Service*. 26 January, 2021. https://sgp.fas.org/crs/misc/IF10830.pdf

12. Appendix

The map in figure 1 below highlights all the different counties in California, and how they are broken up into 19 different rating regions or health insurance markets.

Figure 1: California Rating Area Map