Endogenous Quality Choice with Imperfectly Informed Consumers

Alexandra Charbi
The University of Texas at Austin

January 28, 2020

Abstract

In markets where consumers are imperfectly informed about the quality of products, governments intervene with quality disclosure policies and financial incentive mechanisms on the demand and the supply side, respectively, to help consumers make more informed choice decisions and to encourage quality improvement. This paper analyzes the welfare effects of quality disclosure and quality subsidies in the Medicare Advantage market. On the demand side, consumers receive information on the quality of health insurance plans through a Star Rating System (SRS) initiated in 2008. On the supply side, higher-rated insurers receive a quality-linked subsidy through a Quality Bonus Payment (QBP) program initiated in 2012. I build and estimate a full demand and supply equilibrium model that accounts for the possibility of unaware of the SRS consumers. To identify consumer awareness with respect to the SRS, I survey Medicare-eligible individuals and I find that 80% of the population is unaware of it. After I inform the unaware respondents of the rating system, I conduct a conjoint experiment to elicit preferences on quality. My conjoint estimates show that respondents who reported they were "*aware*" of the SRS assign an average monthly value to a star equal to $24, while the ones reported they were "*unaware*" of the SRS assign an average monthly value to a star equal to $19. I combine my unique survey Stated Preference choice data with Revealed Preference micro-level choice data and I estimate a Bayesian learning discrete choice model that allows for consumer heterogeneity in quality awareness and preferences. On the supply side, firms endogenously choose both price and quality endogenizing the different existing consumer types and the financial incentives they are offered. In my counterfactual analyses, I investigate the equilibrium outcomes that arise under different scenarios; first, a scenario in which consumers are fully informed, second, a scenario in which star ratings are perfectly informative, and third, a scenario of a different bonus scheme. Lastly, I investigate what is the optimal combination of those two policies from the perspective of the social planner.

Contents

1 Introduction

Delivering the highest possible quality to consumers without wasting tax-payer money in markets that are plagued by imperfect information is the reason this study is important. When consumers are imperfectly informed about the quality of a product in a market, producers have little incentive to invest in improving the quality of said product.[1] One way governments attempt to remedy the information problem is by providing consumers with information about product quality through rating systems. Such rating systems have been implemented both in the public (e.g. health and school sector) and the private sector (e.g. airline, restaurant, hotel industry).[2] In cases where consumers remain uninformed about product quality even after its disclosure, quality improvement can be encouraged by intervening on the supply side of the market. Finding out the extent to which these demand and supply side interventions improve quality and the extent to which the cost incurred is justified is important from an economic perspective. In addition, knowing what combination of demand and supply side interventions actually work can lead to better policies.

In this paper, I will focus on the market for health insurance to study the relative impacts of demand versus supply side incentives for quality improvement. A very common characteristic of health insurance markets is that there are many different dimensions of quality that characterize a health insurance plan. Some of these dimensions are easy to observe as they are marketed intensively to consumers through prices, copays, deductibles, etc., while other dimensions that refer to health outcomes, customer service, etc. are more hidden. More recently, the government has allocated considerable efforts to make these aspects of quality more salient with a variety of quality disclosure policies in hopes for better choice decisions and quality improvement.[3] In the absence of demand response to many of these disclosure policies,[4] there are cases in which the government has also tried to directly incentivize insurance providers to improve these under-appreciated attributes through financial bonuses.[5] In my analysis, I use the context of Medi-

[1]Typically, if consumers are not well aware of the price or the quality of a product in a market suppliers will have no incentive to either lower its price or incur cost to improve its quality, respectively, because this will not affect demand.

[2]Examples of quality disclosure agencies and policies are the National Committee for Quality Assurance (NCQA) that collects and provides information about health insurance plans, and the 2001 No Child Left Behind (NCLB) Act that required states to develop a program of statewide student testing and school accountability measures that would be disclosed to the public in the form of report cards. Moreover, publications like Yelp, Trip Advisor, Zagat, and Expedia provide consumer reviews on the quality of a series of products and services.

[3]Some examples of these policies are the publication of reporting reporting cards in the Medicare Advantage (MA) market during 1999-2000, and the more recent publication of star ratings in both the MA market (2008- present) and the Affordable Care Act (ACA) insurance exchanges (2017-present).

[4]This market is notorious for being plagued by many information frictions and high levels of inattention.

[5]Such an example is the Quality Bonus Payment (QBP) program implemented in the MA market during 2012-present that will be analyzed in the study. However, similar Pay For Performance (PFP) programs have been implemented in various setting

care Advantage (MA), in which a demand and a supply side intervention were implemented to improve these shrouded plan quality attributes that were difficult for consumers to observe upfront, and I study the relative impacts of each intervention for quality improvement.

MA is a subsidized program in the United States (US) that offers Medicare beneficiaries private insurance that replaces Traditional Medicare (TM). In 2008, Medicare suspected that beneficiaries were inadequately informed about some non-financial aspects of quality of the locally available health insurance plans they could enroll in. To remedy the problem of inadequate information, Medicare initiated a rating system called the Star Rating System (SRS) as a way to inform beneficiaries in hopes that more information would lead to better decisions that would also encourage competition on the supply side. In the absence of sufficient beneficiary response to the star ratings, in 2012 Medicare intervened on the supply side of the market by providing quality-linked subsidies to higher-rated insurers through a program called the Quality Bonus Payment (QBP) program as a way to directly incentivize insurers to improve quality. Within a few years, the MA market experienced massive but costly quality improvement–the cost approaching $8b in the first three years of the QBP program.

Government payments to insurers to improve quality can be justified by consumer willingness to pay for quality improvement; yet, it appears that however successful in ultimately ensuring quality the QBP might have been, its high cost call in question the efficiency of both the demand and supply side interventions–the efficiency of the SRS in guiding consumers to choose plans and in determining plan payments and the efficiency of the QBP in terms of the overall cost incurred. More specifically, on the demand side, it is unclear whether consumers found the information on quality valuable, whether they cared about these quality metrics and more importantly whether they knew these quality metrics even existed.[6] On the supply side, Layton and Ryan (2015) documented no increase in quality due to the financial incentives given to insurers from the QBP, while more recently Fioretti and Wang (2018) found that the bonus payments were retained by insurers as revenues instead of being passed to beneficiaries.[7]

To disentangle the relative impacts of the rating versus the financial incentive interventions, I build a structural demand and supply model that enables me to quantify the value consumers assign to the information they receive, to investigate whether insurance providers might overinvest or underinvest in

in the health-care market.

[6] Darden and McCarthy (2015) and Reid, Deb, Howell, and Shrank (2013) found conflicting evidence on the way the SRS affected enrollment the first years after it was implemented.

[7] Moreover, there have been many concerns that insurers could game the SRS to increase their payments.

quality, and to also investigate whether the interaction of those interventions lead to efficient equilibrium outcomes. Specifically, on the demand side, I introduce a Bayesian discrete-choice model, in the spirit of Chernew, Gowrisankaran, and Scanlon (2008), to analyze how consumers form beliefs on quality and to quantify the value they assign to the information they receive. On the supply side, in the spirit of Fan (2013), I introduce a model in which health insurers compete in two dimensions–the dimension of price and the dimension of quality–endogenizing the information environment and the incentives they are given by the two implemented policies.

A novel part of my study is that before I determine the value consumers assign to the information they are provided, I firstly determine the extent to which information reaches consumers, that is both welfare and policy relevant. So far, the literature has identified two types of frictions that may cause information not to reach consumers as intended (Handel and Schwartzstein (2018)). One refers to rational inattention, a case in which consumers rationally decide not to avail themselves of the information provided because of the costs embedded (Stigler (1961), McCall (1970), Caplin and Dean (2015)), while the other refers to the case where consumers may simply not be able to comprehend the information (Bordalo, Gennaioli, and Shleifer (2012), Bordalo, Gennaioli, and Shleifer (2013)). Cost and comprehension frictions can together or independently undermine the efficacy of information provision programs. However, an additional friction economists have not considered fully yet due to data limitations is a friction that in this analysis I define as the awareness friction. The awareness friction is a situation in which consumers are simply not aware that information available even exists and it precedes any other type of friction since consumers unaware of any kind of information cannot incur any cost to comprehend it.

To address the awareness friction, I allow my demand model to control for different consumer types in relation to the rating system: Consumers who are aware, unaware, care, or do not care about star ratings. The combination of these types leads to four main consumer types: i) consumers who are aware of the SRS and care about these ratings, ii) consumers who are aware of the SRS, but do not care about these ratings, iii) consumers who are not aware of the SRS, but had they been aware of it they would care about the ratings, and iv) consumers who are not aware of the SRS, and had they been aware they would still not care about the ratings. In the presence of these consumer types I model beneficiary preferences for star ratings with a random coefficient that is normally distributed with a mass point at zero–the mass point represents the fraction of beneficiaries that does not respond to the ratings, either because they are not aware of the rating system or because they do not care about the star ratings–and I allow each consumer to choose a plan according to their type as in Berry and Jia (2010).

A major identification challenge arises in distinguishing beneficiaries who are unaware of the SRS from beneficiaries who do not care about the ratings, because their behavior makes them to appear observationally equal in the choice data. To overcome this challenge, I design and run an electronic survey[8] that yields me unique data on 416 nationally representative Medicare-eligible individuals and enables me to measure consumer awareness of the SRS and to fully recover the joint distribution of the different consumer types and their preferences for star ratings. I measure consumer awareness by directly asking my survey respondents whether they were aware of the SRS existence. To recover the joint distribution of the different consumer types and to elicit their corresponding preferences, after I inform all different consumer types about the existence of the SRS I conduct a conjoint analysis in the spirit of Ben-Akiva, McFadden, and Train (2019).

In the conjoint analysis I randomly provide every respondent a set of 4 different choice sets, each consisting of two options among which there is a trade-off between star ratings and prices. Each respondent has to choose their most preferred option. Hence, I end up with a panel of 4 different preferred choices for each respondent, which I use to estimate the distribution of their preferences for star ratings. I model respondent preferences for star ratings with a random coefficient that is normally distributed with a mass point at zero as in the main demand model. However, the mass point at zero in this setting represents only the consumer types who do not care about star ratings and is estimated given the awareness type the respondent reports. The mass point at zero is identified by the fraction of respondents who consistently choose the cheapest available option in all the choice sets they are given.

Another novel part of my paper is the combination of my unique Stated Preference (SP) choice data coming from my survey with detailed Revealed Preference (RP) micro-level choice data coming from the Medicare Current Beneficiary Survey (MCBS) in the estimation of the main learning demand model. Specifically, after I recover the joint distribution of the different consumer-types' preferences for prices and star ratings, I rescale it with a conversion parameter in the main demand model to account for any potential bias that could arise in the SP choice data due to the different choice environment and I estimate the remaining parameters. The combination of SP and RP choice data is very advantageous in providing more accurate estimates.[9]

[8]The survey was administered by Qualtrics, a company providing a survey platform/ software for collecting, and analyzing data for market research, customer satisfaction, loyalty, etc.

[9]Results that exactly show that will be posted in a later version of the draft.

To recover the value consumers assign to the SRS, I assume that prior to the SRS beneficiaries form some prior beliefs on quality and make a choice that maximizes their current expected utility according to their type (i.e. "*carers*", "*non-carers*"). After the introduction of the SRS, beneficiaries receive a signal on quality through the star ratings, update their beliefs in a Bayesian fashion and make the choice that maximizes their current expected utility according to their type (i.e. "*aware/ carers*", "*aware/ non-carers*", "*unaware/ carers*", "*unaware/ non-carers*") with the "*unaware*" types to act as if they were in the pre-SRS period.

I use the demand estimates in the supply-side model, where insurers compete statically in both prices and star ratings endogenizing the different existing consumer types and the payment incentives they receive from Medicare, to estimate the cost structure of the insurance providers. The competitive environment in the market helps me recover the insurers' implied marginal cost of providing their services by inverting the first-order condition for prices. I assume that insurers incur some fixed cost for each different quality level they target with respect to which they are heterogeneous and I identify the parameter that captures their cost heterogeneity from the different levels of star ratings that arise in equilibrium given the competitive aspects of the market.

The model is estimated combining three main data sources. First, I use detailed administrative data on MA health insurance choices from the MCBS. The MCBS surveys on average 15000 Medicare beneficiaries every year and provides a rotating panel that tracks beneficiaries up to three years containing information on individuals' health insurance choices, the characteristics of each individual's choice of an MA plan, as well as detailed demographic information of all the sampled Medicare beneficiaries, such as health status, age, income, education, etc. Second, I use information from publicly available data sources provided by the Centers for Medicare and Medicaid Services (CMS).[10] These data sources provide data sets on aggregate level information on total enrollment levels of the MA health insurers and the main characteristics of the plans they offer, such as premiums, payments, out-of-pocket cost, copays, deductible, star ratings, drug, vision, etc. Third, I use information from the survey I conduct, that yields me unique information on the SRS awareness levels, the joint distribution of the arising different consumer types and their corresponding preferences for star ratings.

Overall, my survey finds that 80% of the Medicare population is unaware of the SRS.[11] Results from

[10]CMS is the organization that manages Medicare.

[11]This result is similar to the result of a relevant study conducted by HealthMine that found that 78% of the Medicare population was unaware of the SRS. See https://www.prnewswire.com/news-releases/only-22-of-medicare-advantage-members-are-

the conjoint analysis show that 10% of the initially aware respondents did not care about the star ratings, and that 20% of the newly aware respondents did not care about star ratings. Overall survey results show that initially aware respondents assign a $25 monthly value per star as compared to respondents who were newly aware that assigned a $20 monthly value per star.

I estimate the main model in two steps via the method of Maximum Simulated Likelihood (MSL) following Goolsbee and Petrin (2004). In the first step, I recover the mean utility levels consumers receive from their chosen plans, along with the scale parameters that govern the variation parameters recovered from the survey. In the second step, I apply a simple Instrumental Variable (IV) regression to recover the mean consumer preferences.

In my counterfactual analyses, I investigate the implications of different policy interventions: One that completely informs consumers, another that ensures a precise quality signal to consumers through more accurate star ratings, and another one that changes the bonus schemes provided to insurance providers. Finally, using the objective of the social planner, I investigate what combinations of these policy interventions lead to the optimal level of quality for consumers.

The remainder of paper proceeds as follows. Section 2 presents the basic institutional background in the MA market. Section 3 describes the literature to which this paper is related. Section 4 describes the data. Sections 5 and 6 describe the empirical framework for the demand and the supply side model, respectively. Section 7 presents the main results and section 8 presents counterfactual results. Section 9 concludes.

2 Institutional Background

Medicare is the largest public health insurance program in the U.S., comprising 2.9% of the GDP in 2018.[12] It was enacted in 1965 under the Social Security Administration (SSA) to provide health insurance for people 65 years old or older regardless of income or medical history. Over the years Medicare has undergone several changes. It is currently managed by the Centers for Medicare and Medicaid Services (CMS) and it also provides health insurance to younger people with disabilities as well as other populations.[13] It covers

familiar-with-star-ratings-of-those-stars-helped-half-choose-a-plan-healthmine-survey-300690119.html.

[12]Medicare is also known as Traditional Medicare (TM), Traditional Fee For Service (TFFS), and/ or Original Medicare (OM).

[13]Specifically, it provides health insurance to populations with End Stage Renal Disease (ESRD) and Amyotropic Lateral Sclerosis (ALS).

hospital care (Part A), medical care (Part B), and prescription drugs (Part D).[14] Providers are required to treat all Medicare beneficiaries and they are reimbursed by Medicare for their services.[15] In what follows, I describe the institutional background on the Medicare Advantage (MA) program, that provides health insurance as an alternative to TM, and the two main related-to-MA policies considered in this study.

2.1 The Medicare Advantage Program

Medicare Advantage (MA) is a form of managed care competition that originated in the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA) in an effort to cut costs and provide a richer pool of options to Medicare beneficiaries.[16] The program has undergone both name and structural changes over the years. It was first called Medicare+Choice (Part C), and then was renamed MA after the Medicare prescription Drug, Improvement, and Modernization Act of 2003.[17] MA allows beneficiaries to opt out of TM and enroll in plans administered by private insurers who are allowed to provide insurance to beneficiaries in exchange for a per patient monthly fee provided that they fully accept the risks each patient carries. A beneficiary enrolled in MA receives her medical coverage from her private insurer exclusively.

The program did not attract a lot of attention at first, but its popularity increased in the last twenty years. From the perspective of the supply side, availability of the program has dramatically increased. In 2018, 95% of Medicare beneficiaries had the option of choosing an MA plan in their county. From the perspective of the demand side, the enrollment level rates have also followed an increasing path. As of the same year, MA was serving 33% of the Medicare population. However, its popularity has been accompanied with higher costs for the government.

Like TM, MA insurers have to cover Parts A and B, and as Medicare beneficiaries, MA enrollees pay yearly for Part B. There are advantages and disadvantages in choosing MA over TM. The advantage of choosing an MA plan is that it usually offers a more generous cost sharing, and often also offers supplemental benefits, such as dental, vision, hearing, and drug coverage. A disadvantage is that the network of healthcare providers MA insurers offer is usually more restricted.

MA plans are offered by different organizations. The most common are Health Maintenance Orga-

[14]The Part D was added in 2006.

[15]Patients may be charged extra fees, such as co-payments/ coinsurance, which they pay directly to their providers.

[16]The economic rationale for managed care competition is that the economic environment will be competitive enough and the profit maximization motives will be such that they will induce firms to price close to marginal cost and to provide the optimal level of quality.

[17]See McGuire, Newhouse, and Sinaiko (2011) for details in the history of MA.

nizations (HMOs) and Preferred Provider Organizations (PPOs). One less common is the Private Fee-For-Service (PFFS) organization type. HMOs deliver care through providers who work directly for the insurance firm. Beneficiaries who buy these plans are required to choose their primary provider and they must have a referral to visit a specialist. HMO enrollees may see a non-network provider, but will have to pay extra. HMOs have also restrictions on the number of visits, tests, or treatments their enrollees can receives. PPOs are similar to HMOs, but they are more flexible in terms of networks, patients' ability to see specialists, plan costs, and coverage for out-of-network services. Finally, PFFS plans resemble Original Medicare plans in access, benefits, and reimbursement to providers; PFFS offers a flat reimbursement rate per procedure, and providers can choose on a case-by-case or service-by-service basis whether or not to accept patients.

As insurance providers in a market, MA insurers compete with each other on premiums. Research has shown that they also compete on benefit design, size of network, and more recently on quality (Miller, Petrin, Town, and Chernew (2019), So (2019)). In addition to competing on the differences in plans, lately, MA firms have begun engaging in marketing very actively.

Although there have been many alterations in the way the subsidies of MA insurers are determined, the main payment structure has remained stable over time. In general, MA insurers receive a Risk-Adjusted (RA) payment from the government per enrollee they serve.[18] Every year CMS calculates benchmark (capitation) rates that represent beneficiaries' health cost estimates for each county in the nation.[19] This calculation is mainly based on the per capita FFS Medicare spending within the county.[20] After benchmark rates are revealed, every insurer submits a bid for each of the plans she offers depending on her cost.[21] If the submitted bid is above the benchmark of the county where a plan is offered, the insurer receives from CMS (for that specific plan) an amount equal to the county benchmark. To cover the difference the insurer adds extra charges to the standard Medicare Part B and/ or to any supplemental premiums. If the submitted bid is below the county benchmark the insurer receives an amount equal to the submitted bid plus an additional rebate (as a reward), the amount of which is a fixed percentage share of the difference between the bid and the benchmark. Insurers are required to pass the actuarial value of the rebates to enrollees

[18]Risk, here, refers to the risk beneficiaries carry depending on their heath status.

[19]Due to the significant variation in the health status and thus health cost of enrollees across the nation, there is also high variation in the benchmark rates.

[20]The final benchmark will be realized after the addition of the average Part A spending, the average Part B spending and a following ranking rule that will place counties into different quartiles according to which the final benchmark will be determines. For more details see Miller, Petrin, Town, and Chernew (2019).

[21]For different plans she offers, an insurer can receive different levels of reimbursement.

through additional benefits.

Beneficiaries are offered plans, study them and select one during an annual Open Enrollment period in the fall of each year. Beneficiaries must enroll in a plan within this period, except in cases when they first become eligibles or when a major life change happens. Recently, plans of higher quality can enroll beneficiaries at any time of the year.[22]

Although MA was initiated to improve the Medicare market, there have been many concerns around it. One recent concern has been the low level of consumer-awareness of non-financial, but still important, aspects of plan quality, such as health outcomes of the people who enroll in the plan, the way plans help enrollees manage their chronic conditions, members' experiences with the plan, access to medical care, customer service, etc., that can help beneficiaries find the plans that could best match their needs.[23] Importantly, consumer-unawareness of quality not only affects consumer choices, but also it dis-incentivizes insurers to invest and further improve quality. In response to these concerns, the government in 2008 initiated an information provision policy, known as Star Rating System (SRS), through which it started rating MA plans to inform consumers on plan-quality. In 2012, this policy was accompanied with a rating-dependent bonus scheme, known as Bonus Quality Payment (BQP) program, that would financially incentivize insurers to further improve quality. The next two sections describe these policies in detail.

2.1.1 The Star Rating System

To improve health outcomes and to help beneficiaries find plans that best match their needs, every October since 2008 CMS has been publishing Part C and D star ratings on a 1-to-5 scale for the locally available MA plans. The rules governing the assignment of a final star rating to an insurer are complicated and have changed over time.[24] Broadly speaking, a variety of data sources is used to collect information on a set of 30 to 44 performance measures that span 5 to 9 categories that are aligned with the government's goals.[25] These metrics are non-financial and refer to health outcomes of the people who enroll in the plan, the way plans help their enrollees to manage their chronic conditions, member experience with the plan, access

[22] Since 2012 plans that receive 5 star ratings can enroll beneficiaries during any time of the year.

[23] It is worth to note here that as is common in every health insurance market, MA is also a market that provides a large amount of information that consumers have to absorb, understand, and interpret before they choose a plan. Thus, comparing information across the many different existing plans, and making a rational decision based on it is often cumbersome. Literature has documented that due to the large amount of information, consumers can be rationally or irrationally inattentive (Handel, Kolstad, and Spinnewijn (2018), Handel and Schwartzstein (2018), Brown and Jeon (2019)).

[24] Section A.3.1 in the appendix analyzes in detail the main rules that govern the assignment of star ratings to insurers.

[25] These data sources are mainly the Consumer Assessment of Healthcare Providers and Systems (CAHPS) and the Healthcare Effectiveness Data and Information Set (HEDIS) surveys.

to medical care as well as customer service. CMS is then using a weighting system to summarize these metrics into an average "summary rate" that is rounded[26] to the nearest half star. Table 5 of the appendix shows the exact metrics along with the weights used in 2015. CMS releases only these discretized versions of the average summary rates as well as the individual performance metrics.[27] The period between the health care delivery, the survey data collection and the final release of star ratings lasts two years.[28] Star ratings are assigned at a contract[29] level and every plan under the same contract gets the same star rating. Contracts receiving less than two stars are excluded from the market until they improve their quality.

Every year before the enrollment period begins CMS officially discloses star ratings along with other plans characteristics at Medicare's Plan Finder website. Sometimes, insurers also disclose the star ratings they are assigned either in their websites or the mailings they send to their beneficiaries. Although insurers disclose information on their assigned star rating, it is not always easy for consumers to find this information comparing to finding information for other benefits. After observing star ratings along with other plan characteristics, consumers can sign up to the plan they prefer.[30]

The SRS has been criticized for its use both on the demand side and the supply side. On the demand side, the main concerns refer to its use as a guide to consumers choosing among MA plans. Specifically, because information on star ratings is not disclosed in all possible sources beneficiaries use to inform themselves on the plan availability, it is possible that they sign up for plans without having seen the star ratings. Moreover, the complexity of the SRS and the abundance of information a final star rating may convey make it difficult for consumers to understand and process it so that they can use it effectively in their choice decision. On the supply side, the main concerns refer to the extent to which insurers can "manipulate" the SRS as a quality metric system. Specifically, although CMS has been changing the rules it follows to assign the final star ratings,[31] insurers are aware of the main algorithm that it follows along

[26]For example a plan with a 3.75 summary score is rounded up to a 4-star rating, while a plan with a 3.74 summary score is rounded down to a 3.5-star rating. For MA plans offering Part D benefits (MA-PD), Part C and D "summary rates" are combined to create an "overall rate". For MA plans not offering Part D benefits, the "summary rate" is also the "overall rate".

[27]Using data on all the individual metrics along with the algorithm that CMS follows to calculate the final star rating, I recover the average summary rates. In section A.3.2 of the appendix, I describe in detail the process I follow for their recovery. Figure 4 shows the recovered summary rates along with the published star ratings. Complete alignment of the recovered summary rates and the final star ratings is not possible due to some data incompleteness from CMS' side.

[28]At the end of the section, I provide an example describing in detail the timing.

[29]A contract refers to a specific product type, such as HMO, PPO, or PFFS, that covers a specific geographical area, whereas a plan refers to a specific package including coverage, premium, copayments, coinsurance and other benefits. A contract can exist in multiple counties and can offer multiple plans that may provide different sets of benefits to the final enrollees.

[30]Beneficiaries sign up for Medicare either automatically if they get Social Security or Railroad Retirement Board (RRB) benefits or by calling the Medicare line and submit an application.

[31]These changes refer to the weights each metric is assigned with, the specific measures it uses, etc. Sometimes these changes are pre-announced, whereas other times they are sudden.

with the metrics that are heavily weighted. As a result, it is often the case that insurers can invest in quality targeting specific metrics that will affect significantly the final star rating.

2.1.2 The Quality Bonus Payment program

The Quality Bonus Payment (QBP) program originates in the amendments of the 2010 Patient Protection and Affordable Care Act (PPACA) that targeted quality improvement of health insurance plans. The QBP is the first Pay For Performance (PFP) program in the market for health insurance and it awards insurers a certain level of bonus depending on the star rating they get.[32] The bonus scheme of the QBP has changed over time. Table 1 shows how the bonus scheme has evolved over time. According to the original plan of the QBP, contracts with 4 or more star ratings would receive a 5% higher benchmark rate than all other contracts. However, the program started with a three-year period (2012-2014) "demonstration project" that provided bonuses to contracts with 3-3.5 star ratings, as well.[33] Bonus payments are paid per enrollee and are calculated as a share of the MA benchmarks, which vary by county.[34]

TABLE 1. **Quality Bonus Payment by Star Rating**

Year	2.5 Stars	3 Stars	3.5 Stars	4 Stars	4.5 Stars	5 Stars
2008 - 2011	0%	0%	0%	0%	0%	0%
2012 - 2014	0%	3%	3.5%	4%	4.5%	5%
2015 - present	0%	0%	0%	5%	5%	5%

Notes: This table shows the quality bonus payments by star rating across time.

After the implementation of the QBP program, quality as reflected by star ratings started increasing. Figure 5 of the appendix shows how the distribution of the continuous quality of plans (as I constructed it) weighted by enrollment has evolved over time. To abstract from concerns that the evolution of quality observed in the graph is driven by potential entry or exit of insurance plans, I exclude plans that entered

[32] PFP programs are not new. They have been used in many industries and more recently they started being used in the healthcare sector to incentivize providers to improve their quality.

[33] The goal behind the "demonstration project" was to incentivize more insurers to increase quality. This project also allowed bonuses for contracts in some counties with special demographic characteristics to be double.

[34] Bonuses for new contracts, for which there is not enough information, are 3.5% of benchmark. Plans failing to report their quality are treated as less than 3.5 stars and do not receive any bonus payments.

or exited the market during the period the policies were implemented and I keep only the plans that have always existed in the market. Many interesting points arise from looking at this graph carefully. First, it is obvious that quality has improved over time with more consumers on the right side of the distribution over time, too, implying potential welfare improvement over time. Second, the fact that the distributions are multi-modal with the modes being around the thresholds required to receive a certain level of star rating implies that insurance providers put effort to reach the corresponding thresholds. Third, and more interestingly, the fact that higher modes of the distributions started being realized closer to the thresholds where the plan reimbursement based on quality would occur after 2012–when the QBP program was implemented–implies that insurance providers were effectively motivated by the financial incentive mechanism. Specifically, we see that during 2012-2014 the modes of the distributions are almost equally spread around the different thresholds, whereas after 2014, when the QBP program took its final form, we see higher masses of the distributions to be realized around the 4-star rating threshold.[35] The fact that we do not see only one big mass around the 4-star rating threshold implies that there is some plan cost heterogeneity that prevents some insurance providers to get 4 star ratings. Lastly, the fact that we do not see the modes to be exactly at the threshold levels of the corresponding star rating levels reflects that there is some uncertainty that prevents insurance providers to be at the exact point that will guarantee them the quality level they target.

Among the many concerns around the QBP program has always been its cost. The first three years of the program increased the size of the bonus payments as well as the number of contracts receiving them, providing bonuses to the vast majority of MA contracts with total spending approaching \$8.35 billion. Part of this cost can be rationalized by the consumer willingness to pay for high-star plans. However, if consumers are not aware of the existence of the SRS or if they are not willing to pay for high-star plans, then the question that arises is whether there are more efficient ways through which the same quality outcome can be realized at a lower cost.

2.1.3 Summary of the payment structure after the implementation of SRS and QBP

Overall, there is a sequence of interrelated events taking place before the official release of star ratings and the final determination of the CMS payments to the MA insurers. In what follows I provide an example of

[35] It is interesting to see that there are also insurers that target even higher levels of quality. That reflects the fact that 5-star plans attend some extra benefits comparing to lower-star plans. For example, they can enroll beneficiaries any time during the year, whereas all the other insurer can do so, only during the formal enrollment period.

this sequence that covers a two-year period between January, 2016-December, 2017 within which insurers deliver health care services, data are collected, stars are released, benchmark rates are determined, and finally, enrollment and payments are realized. More specifically,

1. From January, 2016 to August, 2016 insurers deliver health care services to their enrollees.

2. Based on the insurer-performance during this period, CAHPS and HEDIS measures are released in September, 2016.[36]

3. From January, 2017 to September, 2017 CMS collects these data, and calculates the star ratings for the enrollment period of 2018. During this period, it also announces changes that may happen regarding the calculation of the star ratings in the future.

4. In October, 2017 the star ratings for the plans that will serve the market in 2018 are released.

5. After the release of the star ratings, the benchmark rates will be finalized, and insurers will submit their bids.

6. The period between October, 15th - December, 7th, 2017 the official enrollment period begins. Consumers can see the plans offered in the area they live, they observe prices, star ratings and other plan characteristics and they decide which plan they will enroll in.

7. The final CMS subsidy is given to every insurer.

3 Related Literature

This paper contributes to the literature that investigates the welfare effects of information provision by allowing for the possibility that there might be consumers who are not aware of the information available. There are two main categories in the literature that analyze the effects of information on welfare. First, there are studies that focus on information provision for quality in the form of reporting cards and ratings (Beaulieu (2002), Wedig and Tai-Seale (2002), Jin and Leslie (2003), Jin and Sorensen (2006), Dafny and Dranove (2008), Lewis and Zervas (2016)). Second, there are studies that focus on the welfare effects of information provision more broadly in the form of advertising (Ackerberg (2003), Shum (2004)). Most

[36]As mentioned in the previous section the star ratings-calculation depends on these measures.

of these studies conduct their analysis under the assumption that consumers are fully aware of the available information. This paper is unique in taking into account the existence of consumers who might not be aware of the available information. In this sense, while most papers quantify the effects of the intention of information provision to treat consumers, this paper quantifies the effects of the treatment of the information provision on the people who are actually treated.

This paper also contributes to the literature that analyzes the different types of frictions that prevent information from fully reaching out consumers leading to inefficient choice decisions in the health insurance market.[37] The existence of different kind of asymmetries in this market has made it very challenging for researchers to distinguish and identify which frictions prevent information from reaching out consumers.[38] Many researchers have identified high switching or searching costs (Nosal (2012), Handel (2013), Ericson (2014) , Polyakova (2016), Yeo and Miller (2018)), while others have identified to rational or irrational inattention (Abaluck and Gruber (2011), Ketcham, Lucarelli, Miravete, and Roebuck (2012), Ketcham, Lucarelli, and Powers (2015), Abaluck and Gruber (2016), Ho, Hogan, and Morton (2017), Heiss, Mcfadden, and Winter) as the main frictions that prevent information from reaching out to consumers. More recently, Handel and Kolstad (2015), Handel, Kolstad, and Spinnewijn (2018) separately identify the role risk preferences, information frictions and hassle costs play in the demand for health insurance plan and their resulting welfare effects. My paper focuses on a different type of information friction and the arising equilibrium effects from the ones analyzed so far. I call this friction the awareness friction. The awareness friction is simply a situation in which consumers are not aware of the information available. In this sense, this paper is closer to Brown (2018) that addresses the fact that there might be limited price transparency that affects consumer decisions in their demand for healthcare services; although Brown (2018) exploits variation from website traffic logs, this paper exploits variation from unique survey data.

This paper is also related to a very rich literature that has studied the MA program. Methodologically, I build my model upon Town and Liu (2003), and Curto, Einav, Levin, and Bhattacharya (2018) that have used a full demand and supply model to estimate the welfare associated with the program under the form it took on the corresponding years of their analysis.[39] Studies in the MA literature have also focused on the pass-through rates of the MA subsidies to the final consumers (Song, Landrum, and Chernew (2013),

[37]This literature goes back to Arrow (1963) who points out that imperfect information is a main friction that prevents consumers from making efficient choice decisions. For an extended literature review on the quality of consumer decision making in the market for health insurance see Kolstad and Chernew (2009).

[38]For an extended literature review on this topic see Handel and Schwartzstein (2018).

[39]Dunn (2010) also analyzes the effects of insurance coverage on the demand for the MA plans.

Duggan, Starc, and Vabson (2016), Cabral, Geruso, and Mahoney (2018)), while others have focused on adverse selection and its welfare implications in the market (Lustig (2010), So (2019)). While most of these studies have focused on financial aspects of quality that takes the form of benefit design, this paper focus on non-financial aspects of quality that are also welfare relevant. This paper is also closely related to Aizawa and Kim (2018) that analyzes the impact of the information provision in the form of advertising on the MA market and to Miller, Petrin, Town, and Chernew (2019) that determines the optimal subsidy schedule for the MA program. Lastly, the study is related to Ericson (2014), Miller (2015), Wu (2016) in the sense that it takes into account the existing demand side information frictions when analyzing the supply side.

This paper also contributes to a more recent literature that has focused on the SRS and QBP programs by analyzing the relative impact of the each program on the total welfare of the MA market. Most studies have focused on the effects of each policy separately from the other. Darden and McCarthy (2015) study the effects of the MA plan ratings reporting on enrollment and find conflicting evidence across the different years of their analysis. McCarthy and Darden (2017) studied the effects of the SRS on the premiums set by insurers and the total welfare. Layton and Ryan (2015) studied the effect of the QBP bonus scheme on plan quality, and Sun (2017) analyzed the equilibrium effects of the QBP program during the first years that was implemented. This is the first paper that builds a full demand and supply equilibrium model that takes into account both policies effectively and thus quantifies the relative impact of the SRS versus the QBP program.[40]

This paper integrates techniques from the behavioral (Deaton (1990), McFadden (2013)), marketing (Green, Krieger, and Wind (2001), McFadden, Bemmaor, Caro, Dominitz, Jun, Lewbel, Matzkin, Molinary, Schwartz, Willis, and Winter (2005), Ben-Akiva, McFadden, and Train (2019)) and transport economics (Kroes and Sheldon (1998), Catalano, Casto, and Migliore (2008), Yang, Choudhury, and Ben-Akiva (2009), Cherchi and Hensher (2015)) literature by collecting unique data via a survey and by using a conjoint analysis to elicit preferences of the survey respondents. Combining stated and revealed preference choice data in discrete choice estimation is advantageous because it provides more accurate results. With the exception of Handel and Kolstad (2015) in the health economics literature, this is the first paper that combines stated and revealed preferences data to estimate a full equilibrium model in the IO literature.[41]

[40]Recently, there has also been a growing literature that focuses on potential risk selection that can arise due to either the SRS or the QBP program either in the demand or the supply side of the market (Madeira (2015), Decarolis, Guglielmo, and Luscombe (2017), Decarolis and Guglielmo (2017), Fioretti and Wang (2018)). However, this study does not focus on any type of selection.

[41]Studies that combine stated and revealed preferences data are usually found in the behavioral, marketing and transport economics literature as mentioned. Some studies are Ben-Akiva, Bradley, Morikawa, Benjamin, Novak, Oppewal, and Rao (1994),

Empirically, the study also uses techniques from the Bayesian learning literature (Ackerberg (2003), Chernew, Gowrisankaran, and Scanlon (2008), Erden, Keane, and Sun (2008), Ching (2010), Grennan and Town (2019)). Specifically, it is close to Chernew, Gowrisankaran, and Scanlon (2008) in the way it quantifies the value consumers assign to the information they receive if they are aware of it by allowing these consumer types to update their beliefs in a Bayesian fashion based on the quality signals they receive. Lastly, the paper is related to Berry and Jia (2010) in the way it allows different consumer types to have different preferences reflected by a different utility specification in the demand model and to Fan (2013) in the way it allows the suppliers to compete in more than one dimension.

4 Data

I combine data from three data sources. First, I use individual level data on individual MA plan choices and demographic characteristics that I find in the Medicare Current Beneficiary Survey (MCBS) for the years 2006-2015.[42] Second, I use publicly available data sets provided by CMS to get information on aggregate enrollment level data as well as plan characteristics for the years 2006-2015. Third, I use unique data that I collect through the survey I conduct in August, 2019.

4.1 Medicare Current Beneficiary Survey Individual-Level Data (2006-2015)

The MCBS is a rotating panel survey that tracks on average 15000 nationally representative Medicare beneficiaries on annual basis for up to 3-4 years. The survey is sponsored by CMS and executed by Webstat. It provides information on beneficiaries' annual enrollment decisions as well as their relevant demographic information, such as age, health status, education, income, sex, race, and location.[43] This data set is also linked to Medicare administrative data that provide information on the MA insurance plans the beneficiaries have chosen and on the claims of the TM enrollees only.[44]

Brownstone, Bunch, and Train (2000), Earnhart (2002), Train and Wilson (2008), Carlsson (2010), Phaneuf, Taylor, and Braden (2013), Morgan, Whitehead, Huth, and Martin (2013).

[42]MCBS did not release data for the year 2014.

[43]To ensure considerable variation in plan enrollment and demographics, the MCBS samples Medicare beneficiaries using a clustering procedure in multiple levels. First, it randomly selects a few Metropolitan Statistical Areas (MSAs) and groups of nearby non-metropolitan areas to represent the existing diversity across urban and rural populations. Second, within each randomly selected area it randomly selects Medicare eligible individuals. Lastly, the MCBS provides sampling weights that I use to transform the information provided and the relevant results into a nationally representative form.

[44]MCBS does not provide the MA plan choice explicitly for some years. Instead, it provides information on the main insurance firm each individual has chosen along with information on premiums and characteristics of their chosen plan. Following Aizawa and Kim (2018) and Miller, Petrin, Town, and Chernew (2019), I compare the characteristics of each chosen MA plan for every MA enrollee with the characteristics of plans offered by the chosen MA insurance firm and I match each MA enrollee in the

From the original MCBS sample I exclude individuals with End Stage Renal Disease because different MA rules are applied to this group of people. Because Part B enrollment is required for MA enrollment, I exclude individual-year pairs who did not enroll in both Medicare Parts A and B. I also exclude any individuals who were dually eligible for both Medicare and Medicaid as well as those with missing address information. Lastly, I drop individuals whose insurance plans were employer-sponsored. After all these exclusions the final sample size of my data set is 53267 individual-year observations covering 1959 county-year markets and 28131 unique individuals.

Table 6 displays the summary statistics of the resulted sampled Medicare population. The first two columns present means and standard deviations across all observations. The average age of individuals in my sample is 72.4. 55% of my sample size is represented by females. Around 90% of the individuals are White with 7.9% and 1.9% to be Black and Hispanic, respectively. Almost 34% of the sample reports that has graduated high school, while 16% and 20% report that have a bachelor degree or have attended college, respectively. More than 50% of the sample size have reported that their health status is at least good, while 17% and 7.2% have reported fair and poor health status, respectively. The last two columns split the sample by MA enrollment. On average MA enrollees have lower income, are slightly older, less likely to be White, and have slightly lower education attainment. Lastly, MA enrollees in the sample have reported on average slightly better health status. Lastly, table 7 shows the proportion of the MA enrollees that have chosen a plan of a certain star rating level in the sample.[45] On average more than 50% of the MA enrollees choose a plan with at least 3-3.5 star ratings.

4.2 Centers for Medicare and Medicaid Services Aggregate-Level Data (2006-2015)

I combine multiple data sets that are available in the databases of the CMS to create a data set at an aggregate level. Specifically, I use information on contract/ plan enrollment level data by state and county and eligibility for MA for the period 2006-2015, which I use to construct market shares of MA plans. I also observe pricing information for 2006-2015 which is composed of three different types of premiums– standard Part B, C, and D premiums. In these data sources I do not observe insurers' bids. However, I observe information on benchmark rates, payment rates and rebates for the period 2006-2016. I combine this piece of information with the pricing information and I construct the submitted bids. I also use data

MCBS to a specific plan within the insurer the enrollee chose.

[45]The number of observations in this table reduces to 3013 because the SRS begins in 2008 (while the data set covers the period 2006-2016) and because star ratings are not available for all the MA plans.

on the out-of-pocket cost an enrollee incurs in a plan she enrolls depending on her health status. The CMS data sources also provide information on the different characteristics as well as the types of all the available plans in every market.

I use information on star ratings for the period 2008-2015. I observe star ratings for all individual metrics, all domains separately, Parts C and D "summary" rates as well as "overall" rates. I do not observe the average summary quality rates that result into the final star ratings. However, CMS provides detailed instructions for their constructions, which I use to recover them.[46]

Table 8 shows summary statistics on plan characteristics by each different level of Star Rating. There are 13862 observations–an observation is a year-contract-plan. Information on plans receiving less than 2.5 star ratings is not reported because of limited availability of such plans in the data. Overall, plans that receive higher star ratings are also priced at a higher level. On average a beneficiary pays \$36 Part C premium, \$23 Part D premium, and \$342 on oopc for a 4-star plan. The data do not reveal any particular pattern between other plan characteristics or plan types with the star rating levels.

4.3 New Survey Data on Medicare Population (2019)

The existing structure of the available data (MCBS and CMS) do not provide information on the different consumer types (e.g. "*aware*", "*unaware*", "*carers*", "*non-carers*"). To overcome this challenge I design and run an electronic survey that yields me unique information on the joint distribution of the consumer types (i.e. "*aware/ carers*", "*aware/non-carers*", "*unaware/ carers*" and "*unaware/ non-carers*") and further demographic information of the relevant sample.

The survey was executed by Qualtrics, which is a privately held experience management company that provides a survey platform/ software for collecting and analyzing data for market research, customer satisfaction, loyalty etc.[47] The survey was distributed to 415 nationally representative respondents in 2019. Because the population of interest was the MA beneficiaries, I stratified the sample and oversampled the MA population by distributing the survey to 309 MA and 106 TM enrollees. Respondents gained access by invitation only and recruits had to go through an additional layer of validation (age, eligibility, and employer-sponsoring) to ensure that they represented the target population. Once respondents were fil-

[46]See the appendix for details on the construction of the average summary rates and figure 4 to observe the recovered variable and how it matches with the given overall star rating. The overlaps shown in the graph are due to data limitations that prevent perfect matching between the resulting average summary rates and the available star ratings data.

[47]Qualtrics has 6m members in N. America and it recruits respondents by providing them an incentive the strength of which depends on the length of the survey and on the target level acquisition difficulty.

tered and got access to the main part of the survey, they had to answer a set of questions regarding their current plan choices, their awareness with respect to the SRS, and their preferences for star ratings.[48]

Table 9 compares information on demographics between the sample population in the MCBS and Qualtrics' (i.e. my own survey's) data.[49] The first two columns present information on the demographics of the MCBS sampled population split by MA enrollment, while the last two columns present information on the demographics of my survey's sampled population split by MA enrollment. Overall the sample in my survey's sample is slightly younger.[50] Slightly more than 50% of my sample is represented by males, while slightly more than 50% of the MCBS sample is represented by females. My survey does not have considerable Black and/ or Hispanic representation comparing to the MCBS sample. The education attainment is similar between the two samples.[51] Lastly, respondents in my survey reported that they are slightly healthier comparing to the health status the MCBS respondents reported. Overall, such differences are expected because the techniques and methodologies that are followed for data collection are not exactly the same between MCBS and Qualtrics. However, these differences are not of a concern because they are not either big in magnitude or statistically significant[52] and because in my analysis I control for potential selection due to such observable characteristics.

5 Demand for Health Insurance Plans

In this section, I introduce a learning Bayesian discrete choice model in the spirit of Chernew, Gowrisankaran, and Scanlon (2008), while also controlling for different consumer types–controlling for different consumer types is welfare relevant and an aspect that distinguishes my model from the standard learning models in the literature. The consumer types I control for are consumers who might or might not be aware of the SRS and the information it provides and consumers who might or might not care about the part of plan quality the SRS is concerned with.

The model builds upon the recent demand models of Town and Liu (2003), Curto, Einav, Levin, and Bhattacharya (2018) and Miller, Petrin, Town, and Chernew (2019) and consists of two periods; the pre-

[48]For the full set of questions I included in the survey see website-TBD

[49]MA observations are more than TM observations in my survey's sample comparing to the relevant proportion in the MCBS sample because of the sample stratification I imposed. Also, information for the "college attainment" variable is not reported for the Qualtrics sample as the relevant variable did not exist in the corresponding question I constructed.

[50]This is not surprising since this is an electronic survey and the people who use electronic devices are in general younger.

[51]The "college attainment" variable is subsumed in the "bachelor degree" variable in my data, because I did not include the option "college attainment" in the education related question in my survey.

[52]A table that confirms this will be posted in a later draft of the paper.

SRS period and the post-SRS period. In the pre-SRS period consumers are uncertain about some parts of the quality of the locally available insurance plans, they form some prior beliefs about them, and then they choose their most preferred plan. In the post-SRS period consumers who are aware of the SRS resolve their uncertainty, update their prior beliefs in a Bayesian fashion, and then choose their most preferred plan. Consumers who are unaware of the SRS act as if they were in the pre-SRS period and choose a plan based on their prior beliefs.

5.1 Assumptions/ Primitives

As it is standard in the literature, I define markets at a county level and I assume that in every market, $m = 1, 2, ..., M$, there is a set of insurers, $k_m = 1, ..., K_m$, each offering a set of plans, $j_{km} = 1, ..., J_{km}$. Plans in every market are differentiated with respect to premium, p_{kjmt}, quality, q_{kjmt}, and other plan characteristics, x_{kjmt}. Individual i gets in market m at time t and chooses a plan j offered by an insurer k.

I assume that the quality of services, q_{kjmt}, a plan offers is composed of two parts. The first part of quality, ξ_{kjmt}, refers to quality as it is determined by plan characteristics that are observed by the consumers, but are unobserved by the econometrician. Examples of these aspects of quality are the network size of insurance plans, marketing campaigns the plans engage in, etc. The second part of quality, q_{kt}, refers to non-financial aspects of quality that are unobserved both by the consumers and the econometrician until the introduction of the SRS.[53] Examples of these aspects of quality are health outcomes of the people who enroll in the plan, the way plans help enrollees manage their chronic conditions, members' experiences with the plan, access to medical care, customer service, etc. I assume that the quality, q_{kt}, of services an insurer, k offers at time t is the same for all the plans she offers and is distributed as follows,

$$q_{kt} \sim N(\mu_0, \sigma_0^2). \tag{1}$$

I allow for two different sets of consumer types the combination of which leads to four main consumer types. The first set refers to consumers who might or might not be aware of the SRS–the "*aware*" and the "*unaware*" types, respectively. Let I_i be a discrete random variable that indicates whether a consumer is

[53] I define q_{kt} at an insurer-time level because the star ratings are assigned at this level. Each plan gets the star rating the insurer it is offered by gets.

aware of the SRS and follows the Bernoulli distribution, $I_i \sim Bern(\iota)$, as follows,

$$I_i = \begin{cases} 1 & \text{w/ prob. } \iota \\ 0 & \text{w/ prob. } 1-\iota \end{cases}. \tag{2}$$

The second set refers to consumers who might or might not care about the part of quality the SRS provides information on–the "*carer*" and the "*non-carer*" types, respectively. Let Φ_i be a discrete random variable that indicates whether a consumer cares for the part of quality, q_{kt}, the SRS is concerned with and follows the Bernoulli distribution, $\Phi_i \sim Bern(\phi)$, as follows,

$$\Phi_i = \begin{cases} 1 & \text{w/ prob. } \phi \\ 0 & \text{w/ prob. } 1-\phi \end{cases}. \tag{3}$$

Finally, the combination of these two sets of distinct consumer types leads to the following final consumer types: "*aware/carers*", "*aware/non-carers*", "*unaware/carers*", "*unaware/non-carers*".[54] Because consumers' knowledge and preferences can affect each other, I allow the two discrete variables, I_i and Φ_i, to be correlated and governed by a joint probability distribution as follows,[55]

$$f(\iota_i, \phi_i) = P(I = \iota_i, \Phi = \phi_i). \tag{4}$$

Lastly, I assume that consumers are risk neutral, that their preferences are stable over time, and that they are aware of prices and other plan characteristics (except for q_{kt}). With this set of assumptions, caring about prices, quality of services and other plan characteristics, each consumer chooses the insurance plan that maximizes her current expected utility.[56]

[54]For the pre-SRS period there are only two of the four original consumer types, the "*carers*" and the "*non-carers*" since the SRS does not exist.

[55]There are two reasons these two variables might be correlated. First, it might be the case that the more likely a consumer cares about quality the more intense her research for the right insurance plan for her is and the more likely it is for her to become aware of the SRS. Conversely, it might be the case that being aware of the SRS makes consumers more conscious with respect to quality. This is actually an empirical question and In my results I find evidence of the latter case. Either way it is important to account for the fact that these two variables are not statistically independent.

[56]By maximizing their current expected utility, consumers are assumed to be static utility maximizers like in Aizawa and Kim (2018) and Miller, Petrin, Town, and Chernew (2019). This is not an unreasonable assumption as Miller, Petrin, Town, and Chernew (2019) also state. First, a model in which consumers would be dynamic utility maximizers would be computationally more intensive than what the model already is. Second, such a model would likely require to assume that individuals make their choices based on neoclassical preferences with a discount factor close to one, a behavior not representing Medicare beneficiaries' behavior as Dalton, Gowrisankaran, and Town (2018) state. Third, Nosal (2012) estimates such a model and finding extremely

Let u_{ikjmt} denote consumer i's utility when enrolled in plan j offered by insurer, k in market, m at time, t. The expected utility she gains from this plan conditional on the plan characteristics as well as her personal taste is given by,

$$E[u_{ikjmt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] = \alpha - \alpha_i p_{kjmt} + \beta_i E[q_{kt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt}, \text{[57]} \tag{5}$$

where $\overline{x}_{kjmt} = (p_{kjmt}, x_{kjmt}, \xi_{kjmt})$ is a vector of the observed by the consumer plan characteristics, $\theta_i = (\alpha_i, \beta_i)'$ is a vector that represents consumer preferences with respect to price and quality, respectively, and ϵ_{ikjmt} is a zero mean i.i.d. stochastic error term that follows the Type I extreme-value distribution across plans and consumers and represents consumers' idiosyncratic tastes.

I allow consumers to have heterogeneous preferences for price and quality[58] that are distributed as follows,

$$\theta_i = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{cases} \begin{pmatrix} \alpha_{i,c} \\ \beta_{i,c} \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{\alpha}_c \\ \overline{\beta}_c \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha,c} & \rho \\ \rho & \sigma_\beta \end{pmatrix} \right) & \text{w/ prob. } \phi \\ \\ \begin{pmatrix} \alpha_{i,nc} \\ \beta_{i,nc} \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{\alpha}_{nc} \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha,nc} & 0 \\ 0 & 0 \end{pmatrix} \right) & \text{w/ prob. } 1-\phi \end{cases}, \tag{7}$$

where ϕ is the probability that consumer i cares about quality, q_{kt} (as defined in eq. 3) and has some preferences for it, and the remaining $1 - \phi$ is the probability that the consumer does not care at all about quality, q_{kt}. In this way, I allow for the "*carer*" and "*non-carer*" (for q_{kt}) consumer types. The "*carer* " types have preferences for prices and quality, $(\alpha_{i,c}, \beta_{i,c})'$, that are jointly normally distributed, while the "*non-carer*" types have only preferences for prices, $(\alpha_{i,nc}, \beta_{i,nc})'$ that are normally distributed.[59]

high switching costs. Fourth, controlling for switching behavior as Miller, Petrin, Town, and Chernew (2019) do and controlling for different consumer types, my model takes into account the inertia that is prevalent in the market.

[57]It would be useful to rewrite the utility u_{ikjmt} into a product level mean

$$\delta_{kjmt} = \alpha - \overline{\alpha} p_{kjmt} + \overline{\beta} E[q_{kt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] + \gamma x_{kjmt} + \xi_{kjmt} \tag{6}$$

following Berry, Levinshon, and Pakes (1995). However, because the utility specification will change depending both on the period–before or after SRS–and on the consumer types, for ease of exposition I will follow the original notation.

[58]With this assumption, I allow for more flexible substitution patterns that are not contaminated by the Independence from Irrelevant Alternatives (IIA) property.

[59]Basically, $\beta_{i,nc} = 0$.

Lastly, α represents the mean utility a consumer gains from choosing an MA plan as opposed to the outside option, which in this market is represented by TM.

5.2 Choice of Health Insurance Plan

In this section, I present the demand model and the form it takes before and after the introduction of the SRS. Before the introduction of the SRS there are only the "*carer*" and the "*non-carer*" consumer types, while after the introduction of the SRS there is the full set of consumer types; the "*aware/ carer*", the "*aware/non-carer*", the "*unaware/ carer*", and the "*unaware/ non-carer*" consumer types. The utility specification changes depending on the type as in Berry and Jia (2010).

5.2.1 Choice of Health Insurance Plan Before SRS

In this subsection, I introduce the model of demand for health insurance plans before the introduction of the SRS. I assume that consumers know the distribution from which quality, q_{kt}, is drawn and that when forming their prior beliefs there is no extra information that can help them predict it.[60] Hence, the mean of their prior is given by,

$$E[q_{kt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] = \mu_0. \tag{8}$$

Under this assumption and conditional on prices and other plan characteristics, consumers choose the plan that maximizes their current expected utility which takes the following form,

$$E[u_{ikjmt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] = \alpha - \alpha_i p_{kjmt} + \beta_i \mu_0 + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt}. \tag{9}$$

[60] This is a technical simplifying assumption. In reality, prices and other plan attributes can affect consumer beliefs. However, in the MA market prices are mostly zero and can barely signal quality. Even non-zero prices can barely signal quality, since what determines quality, q_{kt}, is a combination of factors difficult for insurers to control, and thus also difficult to price.

Given the distributional assumption on ϵ_{ikjmt}, the observed choice probability of an individual i enrolled in plan j offered by insurer k in market m at time t is given by,

$$
\begin{aligned}
&s_{ikjmt}(\overline{x}_{kjmt}, \epsilon_{ikjmt}; \theta_{pre}) \equiv Pr(i \text{ chooses } j) = \sum_{\phi_i \in \{0,1\}} Pr(\Phi_i = \phi_i) \cdot s_{ikjmt}(\overline{x}_{kjmt}, \epsilon_{ikjmt}; \theta_{pre}|\phi_i) = \\
&\phi \cdot \int\limits_{\alpha_{i,c}, \beta_{i,c}} \frac{exp(\alpha - \alpha_{i,c} p_{kjmt} + \beta_{i,c}\mu_0 + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum\limits_{k \in K} \sum\limits_{\substack{j' \in J \\ j' \neq j}} exp(\alpha - \alpha_{i,c} p_{kj'mt} + \beta_{i,c}\mu_0 + \gamma x_{kj'mt} + \xi_{kj'mt})} d(\alpha_{i,c}, \beta_{i,c}) + \\
&(1-\phi) \cdot \int\limits_{\alpha_{i,nc}} \frac{exp(\alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum\limits_{k \in K} \sum\limits_{\substack{j' \in J \\ j' \neq j}} exp(\alpha - \alpha_{i,nc} p_{kj'mt} + \gamma x_{kj'mt} + \xi_{kj'mt})} d\alpha_{i,nc},
\end{aligned}
\tag{10}
$$

where $\theta_{pre} = (\alpha, \overline{\alpha}_c, \sigma_{\alpha,c}, \overline{\beta}_c, \sigma_{\beta,c}, \rho, \overline{\alpha}_{nc}, \sigma_{\alpha,nc}, \gamma, \phi, \mu_0)$ the set of parameters to be estimated. The expected consumer surplus for an individual i given prices and other plan attributes is given by,

$$
\begin{aligned}
CS_i(\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}) = \phi \cdot \frac{1}{\alpha_{i,c}} log \left(1 + \sum_{k \in K} \sum_{j \in J} exp(\alpha - \alpha_{i,c} p_{kjmt} + \beta_{i,c}\mu_0 + \gamma x_{kjmt} + \xi_{kjmt} \right) + \\
+ (1-\phi) \cdot \frac{1}{\alpha_{i,nc}} log \left(1 + \sum_{k \in K} \sum_{j \in J} exp(\alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} \right).
\end{aligned}
\tag{11}
$$

5.2.2 Choice of Health Insurance Plan After SRS

In this subsection, I introduce the model of demand for health insurance plans after the introduction of the SRS. When the SRS is introduced signals in the form of star ratings, r_{kt},[61] are given to the individuals. I assume that the signal through the star rating is distributed normally around the true quality as follows,

$$
r_{kt} \sim N(q_{kt}, \sigma_r^2), \tag{12}
$$

where σ_r^2 represents the precision of the signal the star ratings send to consumers. By Bayes' rule, consumers' posterior beliefs about quality, q_{kt}, are also normally distributed. The mean of the posterior is given by,

$$
E[q_{kt}|\overline{x}_{kjmt}, r_{kt}, \theta_i, \epsilon_{ikjmt}] = w r_{kt} + (1-w)\mu_0, \tag{13}
$$

[61]The star ratings are assigned at the insurer level. All plans offered by a specific insurer get the same star rating.

where w represents the weight consumers put on the signal and is defined as $w = \frac{\sigma_0^2}{\sigma_0^2+\sigma_r^2}$. If the star ratings are very informative for consumers then $\sigma_r^2 \to 0$ and $w \to 1$ and thus when forming their posterior beliefs, consumers put all the weight to the information that is provided by the star ratings. Conversely, if the star ratings are not very informative for consumers then $\sigma_r^2 \to \infty$ and $w \to 0$ implying that when forming their posterior beliefs, consumers place no weight on the star ratings.[62]

Consumers who are aware of the SRS receive the signals from the star ratings and update their beliefs in a Bayesian fashion. Consumers who are not aware of the SRS behave as if they were in the pre-SRS period. All consumers will choose the plan that maximizes their current expected utility, which, now, takes the following form,

$$E[u_{ikjmt}|\overline{x}_{kjmt}, r_{kt}, \theta_i, \epsilon_{ikjmt}, \iota_i] = \begin{cases} \alpha - \alpha_i p_{kjmt} + \beta_i(wr_{kt} + (1-w)\mu_0) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 1 \\ \alpha - \alpha_i p_{kjmt} + \beta_i \mu_0 + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 0 \end{cases}. \tag{14}$$

Given the distributional assumption on ϵ_{ikjmt} and the distributional assumption that governs the existence of the difference consumer types, the observed choice probability of an individual i enrolled in plan j offered

[62]Under the assumption that both the prior beliefs and the signal are normally distributed, the variance of the posterior beliefs is $Var[q_{kt}] = w\sigma_r^2$. However, since I have assumed that consumers are risk neutral, I cannot identify this variance.

by insurer k in market m at time t is given by,

$$s_{ikjmt}(\overline{x}_{kjmt}, r_{kt}, \epsilon_{ikjmt}; \theta_{post}) \equiv Pr(i \text{ chooses } j) = \sum_{\iota_i, \phi_i \in \{0,1\}} Pr(I_i = \iota_i, \Phi_i = \phi_i) \cdot s_{ikjmt}(\overline{x}_{kjmt}, r_{kt}, \epsilon_{ikjmt}; \theta_{post} | \iota_i, \phi_i)$$

$$= Pr(I_i = 1, \Phi_i = 1) \cdot \int_{\alpha_{i,c}, \beta_{i,c}} \frac{exp(\alpha - \alpha_{i,c} p_{kjmt} + \beta_{i,c}(wr_{kt} + (1-w)\mu_0) + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum_{k \in K} \sum_{\substack{j' \in J \\ j' \neq j}} exp(\alpha - \alpha_{i,c} p_{kj'mt} + \beta_{i,c}(wr_{kt} + (1-w)\mu_0) + \gamma x_{kj'mt} + \xi_{kj'mt})} d(\alpha_{i,c}, \beta_{i,c})$$

$$+ Pr(I_i = 1, \Phi_i = 0) \cdot \int_{\alpha_{i,nc}} \frac{exp(\alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum_{k \in K} \sum_{\substack{j' \in J \\ j' \neq j}} exp(\alpha - \alpha_{i,nc} p_{kj'mt} + \gamma x_{kj'mt} + \xi_{kj'mt})} d\alpha_{i,nc}$$

$$+ Pr(I_i = 0, \Phi_i = 1) \cdot \int_{\alpha_{i,c}, \beta_{i,c}} \frac{exp(\alpha - \alpha_{i,c} p_{kjmt} + \beta_{i,c}\mu_0 + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum_{k \in K} \sum_{\substack{j' \in J \\ j' \neq j}} exp(\alpha - \alpha_{i,c} p_{kj'mt} + \beta_{i,c}\mu_0 + \gamma x_{kj'mt} + \xi_{kj'mt})} d(\alpha_{i,c}, \beta_{i,c})$$

$$+ Pr(I_i = 0, \Phi_i = 0) \cdot \int_{\alpha_{i,nc}} \frac{exp(\alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum_{k \in K} \sum_{\substack{j' \in J \\ j' \neq j}} exp(\alpha - \alpha_{i,nc} p_{lj'mt} + \gamma x_{kj'mt} + \xi_{kj'mt})} d\alpha_{i,nc}, \tag{15}$$

where $\theta_{post} = (\alpha, \overline{\alpha}_c, \sigma_{\alpha,c}, \overline{\beta}_c, \sigma_{\beta,c}, \rho, \overline{\alpha}_{nc}, \sigma_{\alpha,nc}, \gamma, \phi, \iota, \mu_0, w)$ the set of parameters to be estimated. The expected consumer surplus for an individual i given prices and other plan attributes in the post-SRS period is given by,

$$CS_i(\overline{x}_{kjmt}, r_{kt}, \theta_i, \epsilon_{ikjmt}) =$$

$$Pr(I_i = 1, \Phi_i = 1) \cdot \frac{1}{\alpha_{i,c}} log\left(1 + \sum_{k \in K} \sum_{j \in J} exp(\alpha - \alpha_{i,c} p_{kjmt} + \beta_{i,c}(wr_{kt} + (1-w)\mu_0) + \gamma x_{kjmt} + \xi_{kjmt})\right)$$

$$+ Pr(I_i = 1, \Phi_i = 0) \cdot \frac{1}{\alpha_{i,nc}} log\left(1 + \sum_{k \in K} \sum_{j \in J} exp(\alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt})\right)$$

$$+ Pr(I_i = 0, \Phi_i = 1) \cdot \frac{1}{\alpha_{i,c}} log\left(1 + \sum_{k \in K} \sum_{j \in J} exp(\alpha - \alpha_{i,c} p_{kjmt} + \beta_{i,c}\mu_0 + \gamma x_{kjmt} + \xi_{kjmt})\right)$$

$$+ Pr(I_i = 0, \Phi_i = 0) \cdot \frac{1}{\alpha_{i,nc}} log\left(1 + \sum_{k \in K} \sum_{j \in J} exp(\alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt})\right). \tag{16}$$

5.3 Identification

In this subsection, I discuss on how I identify the main parameters of my model. There are four sets of parameters that I estimate in my full demand and supply model; three coming from the demand side and one from the supply side. I postpone the discussion of the supply-side parameters until a later section where I introduce my supply side model, and for now, I focus on the three sets of the demand-side parameters. First, I discuss how I identify the different consumer types in the model. Second, I discuss on how I identify the parameters that refer to the consumer preference heterogeneity, and third, I discuss on how I identify the parameters that refer to the consumer mean preferences.

Consumer types: An empirical challenge in this study is to separately identify the different consumer types I control for in my model. Specifically, it is impossible to separately identify the "*unaware*" from the "*non-carer*" consumer types as their behavior makes them seem observationally equal in the available individual choice data. Consequently, it is impossible to identify the joint distribution that governs the combination of the relevant consumer types ("*aware/ carers*", "*aware/ non-carers*", "*unaware/ carers*", "*unaware/ non-carers*").

To better understand this challenge and for ease of exposition, let's focus on the post-SRS period and let's assume that the star ratings are perfectly informative, which implies that $w = 1$. Let's also assume that consumer preferences for quality, q_{kt} are governed by a simple random coefficient that is defined as follows,

$$\beta_i = \begin{cases} \overline{\beta}_c & \text{if } \phi_i = 1 \\ 0 & \text{if } \phi_i = 0 \end{cases} . \tag{17}$$

Then, the utility specification takes the following form,

$$E[u_{ikjmt}|\overline{x}_{kjmt}, r_{kt}, \theta_i, \epsilon_{ikjmt}, \iota_i] = \begin{cases} \alpha - \alpha_i p_{kjmt} + \beta_i r_{kt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 1 \\ \alpha - \alpha_i p_{kjmt} + \beta_i \mu_0 + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 0 \end{cases}$$

$$= \begin{cases} \alpha - \alpha_{i,c} p_{kjmt} + \overline{\beta}_c r_{kt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 1, \phi_i = 1 \\ \alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 1, \phi_i = 0 \\ \alpha - \alpha_{i,c} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 0, \phi_i = 1 \\ \alpha - \alpha_{i,nc} p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \iota_i = 0, \phi_i = 0 \end{cases} . \tag{18}$$

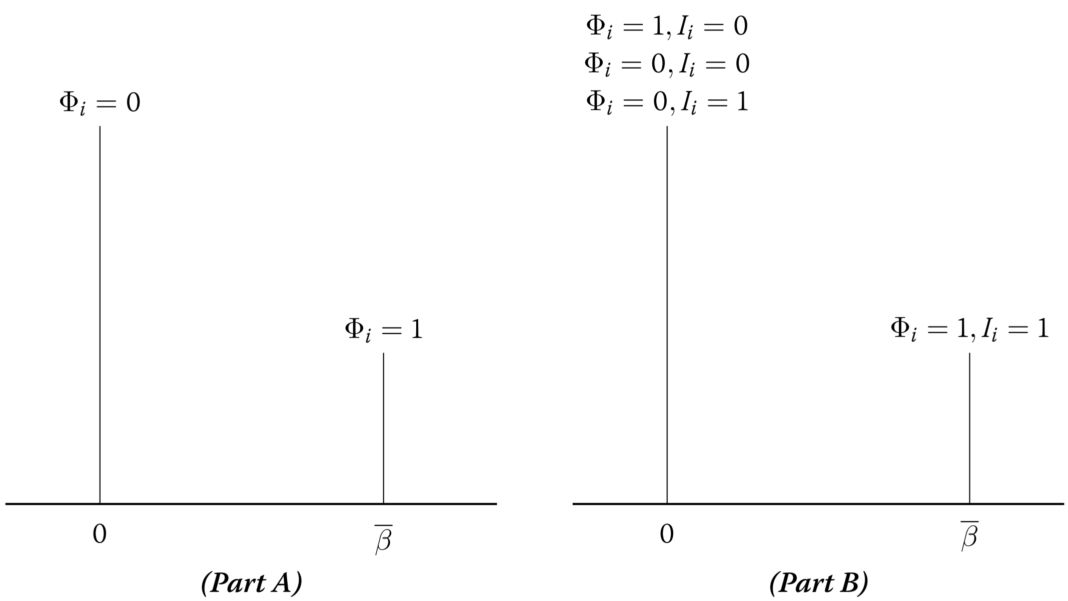
As is obvious from the second part of eq. 18 there is a fraction of consumers–consumers corresponds to the types for which $\iota_i = 0, \phi_i = 0$–who care about star ratings, however, they appear as they do not. Figure 1 helps us to articulate the identification challenge more clearly. Each part of the figure shows fractions of consumers who are massed at 0 and $\overline{\beta}_c$ representing the fractions of consumers who do not respond and who respond to star ratings, respectively. Under the assumption that there is a random fraction of consumers who do not care for plan quality, one could conclude that the fraction of consumers who do not respond to star ratings represent the "*non-carer*" types, whereas the ones who respond to star ratings represents the "*carer*" types as shown in *Part A* of the figure. However, since unawareness of SRS makes the "*unaware*" types observationally equal to the "*non-carer*" types, there are actually three sets of consumer types– the "*aware/ non-carers*", the "*unaware/ carers*", and the "*unware/ non-carers*"– that are hidden behind the proportion of people who do not respond to star ratings as shown in *Part B* of the figure.

To overcome this challenge I designed and ran a survey targeting Medicare population in the US and I separately identified the different existing consumer types as well as the preferences for quality that govern each consumer type. In section 5.4 I discuss in detail how I designed and executed the survey as well as the strategy I followed to separately identify the different consumer types and their corresponding preferences.

Consumer preference heterogeneity: The identification of the parameters that represent the consumer preference heterogeneity is mainly coming from individuals' insurance plan choices from the individual level data the MCBS provide. The variation that identifies these parameters is mainly the different characteristics of the insurance plans each individual with specific characteristics chooses.

Consumer mean preferences: A well-known identification problem in the literature is that plan characteristics can be endogenous to ξ_{kjmt}. In a standard model that builds upon Berry, Levinshon, and Pakes (1995) the variable that is treated as endogenous is usually the price of a product, whereas other product attributes are treated as exogenous. In my supply model, that I will discuss in section 6, I assume that insurers make decisions on both prices, p_{kjmt}, and star ratings, r_{kt}, and take all the other plan characteristics, x_{kjmt}, as exogenous. Assuming that x_{kjmt} are exogenous, I am able to identify the mean consumer preferences, γ, and I am also able to use functions of x_{kjmt}'s to instrument for the prices as in Berry, Levinshon, and Pakes (1995) to identify $(\overline{\alpha}_c, \overline{\alpha}_{nc})$. An extra set of instruments that I am able to use is the natural cost shifter that arise from the bonuses that come from the QBP program. Lastly, to identify $\overline{\beta}_c$, I also need to instrument for star ratings, r_{kt}. A set of instruments that have been introduced by Fan (2013) is the demographics in the market of an insurer's competitors.

FIGURE 1. **Identification - coefficient on r_{kt}**



5.4 Survey

In this section, I describe the tools I used in the survey to separately identify the different consumer types and to recover the joint distribution that arises from their combination so that I estimate my main demand model. There are two sets of consumer types in the model; the first refers to the extent consumers are aware of the SRS–the "*aware*" and the "*unaware*" types–and the second refers to the extent consumers care about the part of plan quality, reflected by the star ratings, the SRS was concerned with–the "*carer*" and the "*non-*

carer" types. Hence, I build two sets of questions in the survey; the first acquiring information on consumer information level and the second eliciting information on consumer preferences for the star ratings. The first set of questions was simply asking respondents whether they were aware of the SRS. The second set of questions was composed of a conjoint analysis, also known as Hypothetical Choice Experiment (HCE) in the literature, eliciting respondents' stated preferences for star ratings.

5.4.1 Information on Information

To elicit information on the level of consumer awareness of the SRS I built a set of questions in which I firstly identified respondents' current insurance choice (TM versus MA), I briefly described them the SRS and its scope, and then I directly asked them whether they knew anything about the SRS. In case respondents answered that they were enrolled in an MA plan and that they were also aware of the SRS, I also asked them how many star ratings their current insurance plan had received. Figure 2 in the appendix presents the main question I asked to elicit consumer awareness of the SRS.[63]

5.4.2 Information on Preferences–Conjoint Analysis

To elicit information on consumer preferences for star ratings, I used the tool of a conjoint analysis following Ben-Akiva, McFadden, and Train (2019). Conjoint analyses have been widely used in marketing research to measure consumer preferences and to forecast demand for components of a prospective product or service. The idea behind a conjoint analysis is that respondents are invited to make a series of choices having to make a trade-off between hypothetical product options.

In the current set-up, assuming that consumers prefer lower prices and higher star ratings, I generated a number of menus (choice sets) that randomly presented a trade-off between prices and star ratings. Table 11 in the appendix presents all the possible values of the attributes presented to respondents in the conjoint experiments. Respondents were presented at random 4 different menus, each providing two options among which respondents had to choose their most preferred. Figure 3 in the appendix shows an example question of a menu presented to respondents. Giving respondents a set of menus to choose one plan every time generated a panel of series of choices for each respondent that was crucial for the identification of the "*carer*" versus the "*non-carer*" consumer types. Specifically, the proportion of the "*non-carer*" consumer types was identified by the proportion of respondents who always chose the cheapest

[63]For the full questionnaire of the survey, please see website-TBD

option across all the different menus they were given.

Identification of the rest of the coefficients was attained in a similar fashion as in standard random coefficient models. An advantage of this experimental design was the random assignment of the different menus to different respondents as well as the random assignment of prices and star ratings to different menus in combination with the assumption that the two options in every menu were identical in every other dimension. The combination of those aspects secured the necessary orthogonality condition needed for identification of the parameters without need for use of additional instrumental variables.

In what follows, I present the model that governed the estimation of respondents' stated preferences for star ratings and prices. Each respondent $n = 1, ..., N$ receives a set of $m = 1, ..., M$ menus, each offering $j = 1, ..., J_m$ alternatives.[64] Each option, j, per menu, m, is characterized by two plan characteristics; prices, p_{jm}, and star ratings, r_{jm}.[65] Let Φ_n a discrete random variable that determines whether a respondent cares about star ratings and follows the Bernoulli distribution, $\Phi_n \sim Bern(\phi)$. The utility a respondent n receives from an option j in menu m is given by,

$$u_{njm} = \alpha_n^s p_{jm} + \beta_n^s r_{jm} + \epsilon_{njm}, \tag{19}$$

where α_n^s, β_n^s are random coefficients that represent the respondents' stated preferences for prices and star ratings, respectively, and are distributed as follows,

$$\theta_n^s = \begin{pmatrix} \alpha_n^s \\ \beta_n^s \end{pmatrix} = \begin{cases} \begin{pmatrix} \alpha_{n,c}^s \\ \beta_{n,c}^s \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{\alpha}_c^s \\ \overline{\beta}_c^s \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha_c^s} & \rho \\ \rho & \sigma_{\beta_c^s} \end{pmatrix} \right) & \text{w/ prob. } \phi \\ \\ \begin{pmatrix} \alpha_{n,nc}^s \\ \beta_{n,nc}^s \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{\alpha}_{nc}^s \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha_{nc}^s} & 0 \\ 0 & 0 \end{pmatrix} \right) & \text{w/ prob. } 1 - \phi \end{cases}, \tag{20}$$

where ϕ is the probability that a respondent does not care for star ratings. Lastly, I assume that ϵ_{njm} follows the Type I Extreme Value distribution. With this set of assumptions, the choice probability of a respondent

[64]In this experiment $N = 415$, $M = 4$, and $J_m = 2 \; \forall \; m \in M$.

[65]In reality, consumers face a richer set of plan characteristics when they are called to choose which insurance plan they will enroll in. However, I departed from this pattern and I gave respondents a simplifying version of the menus they face in reality to keep the experiment easy and the length of the survey short.

n choosing plan j in menu m is given by:

$$P_{njm}(p_{jm}, r_{jm}|\Phi_n = \phi_n; \theta^s) = \begin{cases} \frac{exp(\alpha^s_{n,c} p_{jm} + \beta^s_{n,c} r_{jm})}{\sum_{k=1}^{J_m} exp(\alpha^s_{n,c} p_{km} + \beta^s_{n,c} r_{km})} & \text{if } \phi_n = 1 \\ \frac{exp(\alpha^s_{n,nc} p_{jm})}{\sum_{k=1}^{J_m} exp(\alpha^s_{n,nc} p_{km})} & \text{if } \phi_n = 0 \end{cases}, \tag{21}$$

where $\theta^s = (\overline{\alpha}^s_c, \sigma_{\alpha^s_c}, \overline{\beta}^s_c, \sigma_{\beta^s_c}, \rho, \overline{\alpha}^s_{nc}, \sigma_{\alpha^s_{nc}}, \phi)$ the parameters to be estimated.

Define the vector $d_n \equiv (d_1, .., d_m, .., d_M)$ to be the portfolio of choices for a respondent n, where $d_m \equiv (d_{1m}, .., d_{jm}, ..., d_{Jm})$, is a vector that shows which plan the corresponding respondent, j chose in the corresponding menu, m she chose with,

$$d_{jm} = \begin{cases} 1 & \text{if } n \text{ chooses } j \text{ in menu } m \\ 0 & \text{o/w} \end{cases}. \tag{22}$$

Then, the probability that a respondent n chooses portfolio $d = (d_1, ..., d_m, ..., d_M)$ is given by,

$$P_{id}(d|r, p; \theta^s) = \sum_{\phi_n \in \{0,1\}} P(\Phi_n = \phi_n) \cdot \int_{\alpha^s_n, \beta^s_n} \prod_{m=1}^{M} \prod_{j=1}^{J_m} P_{njm}(p_{jm}, r_{jm}|\Phi_n = \phi_n; \theta^s)^{d_{jm}} dF(\alpha^s_n, \beta^s_n), \tag{23}$$

where $r \equiv (r_1, .., r_m, .., r_M)$; $r_m = (r_{1m}, .., r_{jm}, .., r_{Jm})$, and $p \equiv (p_1, .., p_m, .., p_M)$; $p_m = (p_{1m}, ..., p_{Jm})$.

5.5 Estimation

In this section, I describe the procedure I followed to estimate the parameters of my model combining data based on actual market behavior (i.e. the Revealed Preferences (RP) observations in the MCBS and CMS data sources) with data coming from the choice experiments (i.e. the Stated Preferences (SP) observations in my own survey). Studies that combine stated and revealed preference data are usually found in the behavioral, marketing and transport economics literature. The combination of these two types of data sets requires the estimation of an extra parameter that converts the SP coefficients into RP coefficients to account for potential differences between the two data sets and possible biases that could arise in the conjoint analysis.[66]. Hence, I follow two main steps. First, I estimate the SP coefficients that arise in the

[66]The literature on SP methods contains extensive discussions of the potential differences RP and SP data could have because SP data are not based on actual market behavior. Mainly, the context and the format of the hypothetical setting in a conjoint analysis could affect responses that would lead to biased SP data that the econometrician should take into account in her analysis. For a detailed discussion on the topic see Ben-Akiva, Bradley, Morikawa, Benjamin, Novak, Oppewal, and Rao (1994)

conjoint model and second, I estimate the conversion parameter that converts the SP coefficients into the RP coefficients.

I estimate the SP parameters, $\theta^s = (\overline{\alpha}^s_c, \sigma_{\alpha^s_c}, \overline{\beta}^s_c, \sigma_{\beta^s_c}, \rho, \overline{\alpha}^s_{nc}, \sigma_{\alpha^s_{nc}}, \phi)$, using the method of Simulated Maximum Likelihood (SML). The corresponding likelihood function for each respondent, n, is defined as follows,

$$L_n = \phi \cdot \frac{1}{S} \sum_s \prod_{m=1}^{M} \prod_{j=0}^{J_m} \left[\frac{exp(\tilde{\alpha}^s_{n,c} p_{jm} + \tilde{\beta}^s_{n,c} r_{jm})}{\sum_{k=1}^{J_m} exp(\tilde{\alpha}^s_{n,c} p_{km} + \tilde{\beta}^s_{n,c} r_{km})} \right]^{d_{jm}} + (1-\phi) \cdot \frac{1}{S} \sum_s \prod_{m=1}^{M} \prod_{j=0}^{J_m} \left[\frac{exp(\tilde{\alpha}^s_{n,nc} p_{jm})}{\sum_{k=1}^{J_m} exp(\tilde{\alpha}^s_{n,nc} p_{km})} \right]^{d_{jm}}, \tag{24}$$

where $\tilde{\alpha}^s_{n,c}, \tilde{\beta}^s_{n,c}, \tilde{\alpha}_{n,nc}$ are random simulated draws from the corresponding distribution in eq. 20.

I make the assumption that the RP random coefficients are linearly related with the SP random coefficients in the following way,

$$\theta_i = \kappa \cdot \theta^s_n.^{67} \tag{25}$$

In this way, I convert the util levels I recover from the estimation of the conjoint experiments into util levels that I will recover from the estimation of the main demand model. I estimate the conversion parameter and the rest parameters of the demand model in two stages following Goolsbee and Petrin (2004), who estimate the mean utility levels, δ_{kjmt}, and the parameters that capture the individual-level variation with a maximum likelihood approach in the first stage and the parameters common to individuals–those parameters defining δ_{kjmt}– with an Instrumental Variables (IV) approach in the second stage. In an analogous way, I estimate the mean utility levels and a conversion parameter for the individual-level variation parameters in the first stage[68] and imposing this conversion parameter to be the same for the mean preferences for prices and star ratings, I estimate the remaining parameters of the model using an IV approach

[67]Eq. 25 implies that

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \kappa \cdot \begin{pmatrix} \alpha^s_n \\ \beta^s_n \end{pmatrix} = \begin{cases} \kappa \cdot \begin{pmatrix} \overline{\alpha}^s_c + \sigma_{\alpha^s_c} \nu_{n1} + \rho \nu_{n2} \\ \overline{\beta}^s_c + \sigma_{\beta^s_c} \nu_{n2} \end{pmatrix} & \text{w/ prob. } \phi \\ \kappa \cdot (\overline{\alpha}^s_{nc} + \sigma_{\alpha^s_{nc}} \nu_{n3}) & \text{w/ prob. } 1-\phi \end{cases}, \tag{26}$$

where $\nu_n = (\nu_{n1}, \nu_{n2}, \nu_{n3})' \sim N(0, I)$.

[68]Remember that I have recovered from the first step of the estimation the entire distribution of the random coefficients for prices and star ratings using SP choice data. Hence, in the first stage of the Goolsbee and Petrin (2004) approach I only need to recover the conversion parameter along with the mean utility levels.

in the second stage. More precisely, the estimation proceeds as follows,

1. Let $\overline{\theta} = (\alpha, \overline{\alpha}^s_c, \overline{\beta}^s_c, \overline{\alpha}^s_{nc}, \kappa, \gamma, w, \mu_0)$ to be the set of parameters that determine δ_{kjmt}. For a given candidate value $\tilde{\overline{\theta}}$,

   (a) use the Berry (1994) inversion to find the unique set of product fixed effects $\delta_{jkmt}(\tilde{\overline{\theta}})$–the mean utility levels–that match the predicted market shares to the observed market shares, and also

   (b) estimate the conversion parameter, κ, using the method of SML.

   The estimation amounts to a mixture of logits and the likelihood function for a consumer i is given by,

$$L_i = \prod_{t=1}^{T} \prod_{i=1}^{N} \prod_{j=1}^{J_m} \left(P^{pre}_{ikjmt}\right)^{y^{pre}_{ikjmt}} \cdot \left(P^{post}_{ikjmt}\right)^{y^{post}_{ikjmt}}, \tag{27}$$

   where

$$P^{pre}_{ikjmt} = \sum_{\phi_i \in \{0,1\}} P(\Phi_i = \phi_i) \cdot P_{ikjmt}(\overline{x}_{kjmt} | \Phi_i = \phi_i; \theta^r_{i,pre}) \tag{28}$$

   the probability that a consumer i will choose a plan j offered by an insurer k in market m before the introduction of the SRS, and

$$P^{post}_{ikjmt} = \sum_{\phi_i, \iota_i \in \{0,1\}} P(\Phi_i = \phi, I_i = \iota_i) \cdot P_{ikjmt}(\overline{x}_{kjmt}, r_{kt} | \Phi_i = \phi_i, I_i = \iota_i; \theta^r_{i,post}) \tag{29}$$

   the probability that a consumer i will choose a plan j offered by an insurer k in market m after the introduction of the SRS.

2. Impose the estimated $\hat{\kappa}$ to be the conversion parameter of the parameters reflecting the consumer mean preferences for prices and star ratings and use an IV approach to estimate the following equation:[69]

$$\hat{\delta}_{kjmt} = \alpha + \hat{\kappa}(-\hat{\alpha}^s_i p_{kjmt} + \hat{\beta}^s_i r_{kt}) + \gamma x_{kjmt} + \xi_{kjmt} \tag{30}$$

[69] The precise equation is adjusted depending both on the period–pre and post-SRS–and on the consumer types. See section tbd of the appendix for the exact forms it takes.

6 The Model - Supply

In this section, I introduce a supply model in which insurance providers compete in more than one dimensions–prices and star ratings–endogenizing the incentives they receive from the QBP program as well as the different consumer types that exist in the market.[70] Using the demand elasticities I recover from the demand estimation, I estimate the supply model and I recover the fixed costs insurers incur when they decide which level of quality, namely which level of star rating, they will invest in and the marginal cost they incur when they offer their services. I use these estimates in a later section to simulate prices and star ratings under various counterfactuals.

6.1 Assumptions/ Primitives

I assume that competition occurs at a county level as is standard in this literature.[71] In each market, m, there is a total number, K_m, of insurers indexed by $k = 1, 2, ..., K_m$, each one offering a set, J_k, of plans indexed by $j = 1, 2, ..., J_k$.[72]

To be in alliance with the real environment of the MA market, I assume that there is a level of uncertainty that governs insurers' quality investment decision. This uncertainty results from the fact that CMS changes periodically the algorithm it uses when it calculates the star ratings. Although some of these changes are pre-announced, some are not and thus the resulting quality level after insurers make their investment decision cannot be predicted. Hence, the final level of quality, q_k, that arises after insurers make their decision is as follows

$$q_k = \tilde{q}_k + \eta_k, \tag{31}$$

where $\tilde{q}_k$ is the level of quality insurers choose and η_k is a random component due to the unpredicted government changes. I assume that η_k is normally distributed, $\eta_k \sim N(\mu_\eta, \sigma_\eta^2)$. It is important to note here that although insurers target an average summary rate of quality, $\tilde{q}_k$, consumers observe only the final

[70] By allowing insurers to compete in more than one dimensions, I allow my model to resemble Fan (2013) and the more recent Miller, Petrin, Town, and Chernew (2019).

[71] See Town and Liu (2003), Curto, Einav, Levin, and Bhattacharya (2018), Miller, Petrin, Town, and Chernew (2019).

[72] For ease of notation I do not use subscripts for time, t.

level of star rating that results based on the following formula,

$$r_k = g(\tilde{q}_k, \eta_k) = \begin{cases} 2 & \text{if } 1.75 \leq \tilde{q}_k + \eta_k < 2.25 \\ \ldots & \\ 5 & \text{if } \tilde{q}_k + \eta_k \geq 4.75 \end{cases} \tag{32}$$

Lastly, I assume that insurers compete simultaneously and statically in prices, p_{kj}, and quality, q_k, in two stages. In the first stage they choose quality, $\tilde{q}_k$, and in the second stage, after star ratings, r_k, are realized, they choose prices, p_{kj}, for each plan they offer.[73]

6.2 Profit Function

Let an insurer k to incur some fixed cost, fc_k, associated with the level of quality she chooses. I assume that the fixed cost of each insurer is independent from the number of plans she offers. This assumption implies that increasing the quality of services she offers, an insurer incurs higher cost, but the cost she incurs per enrollee she serves does not change. I also assume that insurers are heterogeneous with respect to fixed costs. This assumption reflects the fact that larger firms may incur higher costs comparing to smaller firms, but they may achieve higher economies of scale as the number of plans they sell increases. With this set of assumptions, I model the fixed cost of an insurer, fc_k, as follows,

$$fc_k(\tilde{q}_k, \nu_k; \tilde{\nu}) = (\nu + \nu_k)\tilde{q}_k, \tag{33}$$

where $\tilde{q}_k$ represents the quality of services each insurer chooses to offer, ν represents an average level of cost insurers incur depending on the quality level they choose, and ν_k represents firm heterogeneity with respect to fixed cost, which I assume to be normally distributed, $\nu_k \sim N(0, \sigma_\nu^2)$. Lastly, $\tilde{\nu} = (\nu, \sigma_\nu)$ represents the sets of parameters to be estimated. Then, the profit function that is relevant to an insurer's first-stage decision is determined by the difference between her variable profits and her fixed costs and is

[73] Notice that the level of quality is chosen at the insurer-level, whereas the level of price is chosen at the plan-level. This depicts the fact that star ratings are assigned (from the CMS) at the insurer-level. Hence, when insurers decide on the quality they will target they make their decision at this same level. However, when they price the plans they offer they price them at the plan level depending on the set of attributes each plan offers.

given by,

$$\pi_k^I = \int\limits_{\nu_k} \int\limits_{\eta_k, \eta_{-k}} \left(\pi_k^{II}(p^*(\tilde{q}_k, \tilde{q}_{-k})) - fc(\tilde{q}_k, \nu_k; \tilde{\nu})\right) dF(\eta_k, \eta_{-k}) dF(\nu_k), \tag{34}$$

where $\pi_k^{II}(p^*(\tilde{q}_k, \tilde{q}_{-k}))$ represents the variable profits from offering insurance services and $p^*(\tilde{q}_k, \tilde{q}_{-k})$ represents the equilibrium price arising in the second stage.

The variable profits, π_k^{II}, of the insurer are determined by the price, p_j, she sets, the marginal cost, c_j, she incurs, and the post-Risk Adjustment (RA) subsidy, B_j, she receives for each plan she offer, and they are given by,

$$\pi_k^{II} = \sum_{j \in J_k} \left[(p_j - c_j + B_j) \int_i s_{ij}(p, r, x; \theta) di \right]. \tag{35}$$

I model the marginal cost, c_j, to be log-linear in observed characteristics, $\tilde{x}_j = (q_j, x_j)$, and in some unobserved (to the econometrician) component, ω_j, as follows,

$$ln(c_j) = \tau \tilde{x}_j + \omega_j. \tag{36}$$

The subsidy each plan receives per enrollee it serves is a function of the arising star rating and is given by,

$$B_j = B(q_k, \tilde{B}_j) = \begin{cases} \tilde{B}_j & \text{if } q_k < 3.75 \\ 1.05\tilde{B}_j & \text{if } q_k \geq 3.75 \end{cases}, \text{[74]} \tag{37}$$

where $\tilde{B}_j$ represents the subsidy each plan receives from the government for each enrollee she serves after RA.[75]

[74] $B(q_k, \tilde{B}_j)$ is adjusted accordingly for $t = 2012 - 2014$. During this period it takes the following form:

$$B(q_k, \tilde{B}_j) = \begin{cases} \tilde{B}_j & \text{if } q_k < 2.75 \\ 1.03\tilde{B}_j & \text{if } 2.75 \leq q_k < 3.25 \\ 1.035\tilde{B}_j & \text{if } 3.25 \leq q_k < 3.75 \\ 1.04\tilde{B}_j & \text{if } 3.75 \leq q_k < 4.25 \\ 1.045\tilde{B}_j & \text{if } 4.25 \leq q_k < 4.75 \\ 1.05\tilde{B}_j & \text{if } q_k \geq 4.75 \end{cases} \tag{38}$$

.

[75] Before the QBP program this scheme did not exist and B_j was simply $\tilde{B}_j$.

6.2.1 Necessary Equilibrium Conditions

The necessary equilibrium conditions for prices and star ratings that result from the insurers' optimization problem arise solving in a backwards induction. The optimality condition that results from the second stage of the insurer's problem is given by,

$$\frac{\partial \pi_k^{II}}{\partial p_j} = 0 \Rightarrow \int_i s_{ij}(p,r;\theta)di + \sum_{j \in J_k} \left(p_j - c_j + B_j\right) \frac{\partial \left(\int_i s_{ij}(p,r;\theta)di\right)}{\partial p_j} = 0. \tag{39}$$

The optimality condition that arises from the first stage of the insurer's problem is given by,

$$\frac{\partial \pi_k^{I}}{\partial \tilde{q}_j} = 0 \Rightarrow \int_{\eta_k, \eta_{-k}} \left(\frac{\partial \pi_k^{II}}{\partial p_k^*} \frac{\partial p_k^*}{\partial \tilde{q}_k} + \sum_{-k} \frac{\partial \pi_k^{II}}{\partial p_{-k}^*} \frac{\partial p_{-k}^*}{\partial \tilde{q}_k} - \frac{\partial fc_j}{\partial \tilde{q}_k} \right) dF(\eta) = 0. \tag{40}$$

6.3 Identification

A challenge of the supply model is the difficulty to separately identify the variable σ_ν that reflects the firm cost heterogeneity from the variable σ_η that reflects the firm's uncertainty with respect to the level of quality that will result in equilibrium. Given all the competitive aspects of the model, identification of firms' cost heterogeneity, σ_ν, is coming from the different observed equilibrium quality levels that would not arise if insurers were homogenous with respect to fixed cost; if insurers incurred the same fixed cost then the same level of quality would arise in equilibrium. Identification of insurers' uncertainty with respect to the level of quality that will result in equilibrium, σ_η, is coming from the dispersion that is observed around the thresholds insurers have to reach in order to get a certain level of star rating and that would not arise if they were not uncertain; if insurers were certain about the algorithm used in the calculation of the star ratings they would target the threshold necessary for a level of star rating perfectly.

Lastly, I use the optimality conditions of the insurers' problem to identify their cost structure. The total set of parameters to be estimated is $(\tau, \mu_\eta, \sigma_\eta, \nu, \sigma_\nu)$. As it is standard in the literature, from the optimality condition given in eq. 39 I invert out the marginal cost of the insurers, which I utilize in eq. 36 to estimate τ. As there are concerns in the demand side that ξ_j is correlated with p_j, there are similar concerns that ξ_j will also be correlated with insurers' costs. However, after estimating the demand parameters, I calculate $\hat{\xi}_j$ for each plan and include it in $\tilde{x}_j$. Then, I estimate τ assuming that any unobservable components of cost are uncorrelated with $\hat{\xi}_j$ and my observables.

7 Results

In this section, I present the results that arise from the estimation of the conjoint experiment, the main demand model, and the supply model. Demand side results are still preliminary and due to possible changes in a later version of the draft. Also, supply side estimates will be determined in a later version of the draft.

7.1 Survey Results

The first parameter that I recover from the survey is the proportion of the Medicare beneficiaries that were aware of the SRS. Table 2 shows the results. Overall, 80.4% of the entire Medicare population appears to be "*unaware*" of the SRS. This result is also similar to a relevant result found by HealthMine in August, 2018.[76] From the TM surveyed beneficiaries only 13.3% responded that they were "*aware*" of the SRS, while from the MA surveyed population 31.8% responded that they were "*aware*" of the SRS. The fact that the MA population appears to be more informed about the SRS is not surprising if we take into account the fact that the SRS refers to MA plans and that the high levels of inertia in the market prevent the TM enrollees to get informed about the characteristics of the MA plans. Overall, the results imply that regardless the efforts and the sources the government puts in the collection of the necessary information to construct the star ratings and in the construction of the star ratings per se, consumers are not aware of these quality measures.

Table 10 of the appendix also shows how the two different consumer types in both the TM and MA populations differed based on their demographic information. The results are similar for both populations. Respondents who reported that they were "*unaware*" of the SRS were slightly older, slightly less educated, and less healthier. Also more females reported to be "*unaware*". The data set does not provide enough variation to compare the population in race and ethnicity.

[76]For their results see https://www.prnewswire.com/news-releases/only-22-of-medicare-advantage-members-are-familiar-with-star-ratings-of-those-stars-helped-half-choose-a-plan-healthmine-survey-300690119.html.

TABLE 2. ***"Aware"/ "Unaware"* of the SRS consumer types**

Consumer types	**Aware**	**Unaware**
Population		
Traditional Medicare	0.133	0.867
Medicare Advantage	0.318	0.682
Overall	0.196	0.804

Notes: This table shows the consumer types regarding Medicare beneficiaries' level of awareness of the SRS.

Given respondents' types–"*aware*", "*unaware*"–I estimate the distribution of their preferences for prices and star ratings from the conjoint experiments. In this way, I recover the marginal preferences for prices and star ratings of the "*aware*" and the "*unaware*" consumer types separately. Table 3 shows the estimates that arise across different specifications. The first two columns show the estimates that arise under the assumption that "*carer*" and "*non-carer*" types have the same preferences for prices across "*aware*" and "*unaware*" types and the last two columns show the estimates that arise under the assumption that "*carer*" and "*non-carer*" types have different preferences for prices across "*aware*" and "*unaware*" types. All coefficients of specification (1) with the exception of $\sigma^s_{\beta,c}$ for the "*unaware*" types are statistically significant. The results of this specification indicate that the "*aware*" types care more about star ratings comparing to the "*unaware*" types. This is reflected by the parameter ϕ that is almost 92% for the "*aware*" types, while it is 71.3% for the "*unaware*" types. Overall, the "*aware*" types value a star rating for \$24.171 per month, while the "*unaware*" types value a star rating for \$19.833 per month. This difference is not surprising since we would expect that "*aware*" consumers to be more conscious when they make the plan choice decision comparing to "*unaware*" consumers. Estimates in specification (2) are mostly statistically significant with the exception of the coefficients $\sigma^s_{\beta,c}$ for both the "*aware*" and the "*unaware*" types and the price coefficients for the "*aware*" types. This is happening because in this specification I have increased the set of parameters that I estimate and at the same time the number of "*aware*" types if fairly low. I solve this problem by sending the survey in a few more respondents. Updated estimates will be posted in a later version of the draft. Overall, the results are similar to the ones of specification (1) with the "*aware*" types to value a star rating for \$26.204 per month and the "*unaware*" types to value a star rating for \$20.269 per month.

In the appendix I also present more tables for different specifications with respect to the distribution of respondents' preferences for prices and star ratings. Specifically, table 12 shows estimates for the case in which the random coefficients for prices and star ratings are assumed to be statistically independent and table 13 shows estimates for the case in which the preferences for prices do not vary at the individual level and there is only one random coefficient–that one for star ratings. Overall, the results are similar to the ones presented in this section with the advantage that the lower the number of the parameters that are estimated the more the predictive power of the parameters becomes.

Table 3. **Conjoint Experiment Estimates**

	(1)		(2)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}_c^s$	-0.1074	-0.1058	-0.0840	-0.0993
	(0.0221)	(0.0137)	(0.0150)	(0.0145)
$\sigma_{\alpha,c}^s$	0.0734	0.0501	0.0610	0.0494
	(0.0168)	(0.0179)	(0.0434)	(0.0180)
$\overline{\beta}_c^s$	2.806	2.9427	2.3720	2.7407
	(0.5367)	(0.3006)	(0.3460)	(0.3419)
$\sigma_{\beta,c}^s$	1.2681	0.6355	0.3633	0.4642
	(0.7076)	(0.6939)	(0.5198)	(0.7047)
ρ	0.0583	0.0593	0.0500	0.0505
	(0.0347)	(0.0250)	(0.0223)	(0.0270)
ϕ	0.9252	0.7131	0.9280	0.7344
	(0.0672)	(0.0491)	(0.0611)	(0.0552)
$\overline{\alpha}_{nc}^s$			-0.0840	-0.1261
			(2.8500)	(0.0445)
$\sigma_{\alpha,nc}^s$			0.0100	0.0590
			(1.4608)	(0.0420)
Observations	112	303	112	303
Assumptions				
$\overline{\alpha}_c^s = \overline{\alpha}_{nc}^s$, $\sigma_{\alpha,c}^s = \sigma_{\alpha,nc}^s$	YES		NO	
Monthly star value given $\phi_i = 1$	\$26.126	\$27.813	\$28.238	\$27.600
Overall monthly star value	\$24.171	\$19.833	\$26.204	\$20.269

Notes: This table presents the estimates of the respondents' distribution of preferences for prices and star ratings across different specifications. Standard errors are reported in parentheses.

7.2 Demand Side Results

Table 4 shows preliminary demand estimates that arise after I combine the SP and RP choice data. The conversion parameter, κ is 0.909 that implies that the coefficients of the RP choice data are very close to the coefficients of the SP choice data. The rest of the parameters are as close to the ones that usually arise in the literature. An updated table that will take more parameters into account will be posted in a later version of the draft.

TABLE 4. **Demand Estimates after combining SP and RP choice data**

Variable	Estimates	Std. Errors
1st stage		
κ	0.909	(TBD)
2nd stage		
Out-of-pocket	-.008	(.000)
Coverage indicators		
Prescription drugs	.511	(.057)
Vision	.250	(.042)
Dental	-.206	(.042)
Copays		
Primary doctor	.019	(.002)
Specialist	.014	(.002)
Coinsurance		
Primary doctor	.098	(.002)
Specialist	-.091	(.005)
Plan types		
HMO indicator	.858	(.043)
PPO indicator	.313	(.061)

Notes: This table presents the estimates that I recover after I combine SP and RP choice data. Standard errors are reported in parentheses.

7.3 Supply Side Results

TBD

8 Counterfactul Analysis

In this section, I conduct counterfactual analyses to understand the impact of the demand versus the supply side policies that were implemented in the MA market. I investigate four different scenarios that are policy relevant. First, given the monthly value the "*aware*" consumers give to a star rating, I investigate the equilibrium outcomes that arise under the assumption that consumers are fully aware of the SRS. This scenario has immediate implications on the actions the government can take to increase the level of consumer awareness. Second, I investigate the equilibrium outcomes that arise under the assumption that the star rating signals are perfectly informative to consumers for the quality of plan services. This scenario has immediate implications on the steps necessary for making the star ratings better signals of quality. Third, I investigate the different equilibrium outcomes that arise under different bonus schemes. This scenario has immediate implications on the government expenditures on the implemented financial scheme incentives. Lastly, I investigate what is the optimal combination of the two implemented policies–quality disclosure on the demand side and financial incentives on the supply side–from the perspective of the social planner.

8.1 Analysis under full information assumption

TBD

8.2 Analysis under perfectly informative star ratings

TBD

8.3 Searching for the optimal combination of demand and supply policies.

TBD

9 Conclusion

In an effort to help consumers make more informed choice decisions, policy makers in a number of contexts have implemented quality disclosure policies in hopes that such policies will also enhance quality competition on the supply side of the corresponding market. In cases where this type of demand side policies fail to inform consumers effectively, policy makers turn their attention in the supply side of the market providing suppliers various financial incentives to improve quality and overall welfare. The ambiguous welfare effects that arise from the combination of those two types of policies make the disentanglement of their relative effects important from both policy and economic perspective.

In this study, I investigate the welfare effects of the SRS quality disclosure policy (2008-2015) and the QBP financial incentive policy (2012-2015) that have been implemented on the demand and the supply side of the Medicare Advantage market, respectively. The study of those combined policies is of particular interest, since early evidence showed that consumers did not respond to quality disclosure, while insurers made significant quality investments that were subsidized by the government.

In my analysis, I combine unique SP choice data coming from an electronic survey I conduct with RP micro-level choice data coming from the MCBS and I build and estimate a full demand and supply model that resembles the MA market. On the demand side, I allow consumers to learn about quality in a Bayesian fashion after the SRS policy becomes in effect, while also allowing for different consumer types; types "*aware*" or "*unaware*" of the SRS and types who care or do not care about the quality as revealed by the SRS. On the supply side, I allow insurance providers to compete in both prices and quality, while also endogenizing the different existing consumer types and the financial incentives they receive.

The combination of RP and SP choice data is very advantageous providing more precise demand estimates. Survey results show that almost 80% of the Medicare population is "*unaware*" of the SRS, while the estimates show that an "*aware*" beneficiary values a star rating for $24 on average per month, while an "*unaware*" beneficiary values a star rating for $19 on average per month. This result is of particular interest since it allows me to conduct unique welfare analysis in which I can investigate scenarios in which consumers are perfectly informed of the SRS.[77]

The strategy I follow to identify consumer awareness can be adopted in any context in which information in any form is provided to consumers. My demand framework that allows for different consumer

[77]A more detailed analysis will be covered in a later draft of the paper.

types with different distributions of preferences and my estimation strategy that combines both SP and RP choice data can be adopted in any discrete choice model to provide more accurate estimates.

Interesting avenues for future research would be to allow for insurance providers to engage in dynamic behavior when investing in quality. Incorporating dynamic behavior in the supply side would resemble the market more realistically. So far the literature has not addressed that. Allowing for consumers to engage in dynamic behavior in this context would also be very interesting, since that would effectively take into account the existing inertia that is prevalent in the market. Nosal (2012) estimates such a model in a more general context of the MA market. It would be interesting to build a demand model in which consumers would be dynamic utility maximizers, while also taking into account the different existing consumer types to analyze the SRS policy more precisely. Lastly, another interesting avenue for future research would be the investigation of incentive schemes that are based on relative as opposed to absolute performance evaluation.

References

J. Abaluck and J. Gruber. Choice inconsistencies among the elderly: Evidence from plan choice in the medicare part d program. *American Economic Review*, 101(4):1180–1210, June 2011.

J. Abaluck and J. Gruber. Evolving choice inconsistencies in choice of prescription drug insurance. *American Economic Review*, 106(8):2145–2184, August 2016.

D. A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44(3):1007–1040, August 2003.

N. Aizawa and Y.S. Kim. Advertising and risk selection in health insurance markets. *The American Economic Review*, 108(3):828–867, March 2018.

K.J. Arrow. Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5):941–973, December 1963.

N.D. Beaulieu. Quality information and consumer health plan choices. *Journal of Health Economics*, 21:43–63, 2002.

M. Ben-Akiva, M. Bradley, T. Morikawa, J. Benjamin, T. Novak, H. Oppewal, and V. Rao. Combining revealed and stated preferences data. *Marketing Letters*, 5(4):335–349, October 1994.

M. Ben-Akiva, D. McFadden, and K. Train. Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis. *Foundations and Trends in Econometrics*, 10(1-2):1–144, 2019.

S. Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2): 242–262, Summer 1994.

S. Berry and P. Jia. Tracing the woes: An empirical analysis of the airline industry. *American Economic Journal: Microeconomics*, 2(3):1–43, August 2010.

S. Berry, J. Levinshon, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, July 1995.

P. Bordalo, N. Gennaioli, and A. Shleifer. Salience theory of choice under risk. *The Quarterly Journal of Economics*, 127(3):1243–1285, August 2012.

P. Bordalo, N. Gennaioli, and A. Shleifer. Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843, October 2013.

Zach Y. Brown. An empirical model of price transparency and markups in health care. *Working paper*, 2018.

Z.Y. Brown and J. Jeon. Endogenous information acquisition and insurance choice. *Working paper*, November 2019.

D. Brownstone, D.S. Bunch, and K. Train. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research*, Part B(34):315–338, 2000.

M. Cabral, M. Geruso, and N. Mahoney. Do larger health insurance subsidies benefit patients or producers? evidence from medicare advantage. *American Economic Review*, 108(8):2048–2087, August 2018.

A. Caplin and M. Dean. Revealed preference, rational inattention, and costly information acquisition. *The American Economic Review*, 105(7):2183–2203, July 2015.

F. Carlsson. Design of stated preference surveys: Is there more to learn from behavioral economics. *Environ Resource Econ*, 46:167–177, 2010.

M. Catalano, B. Lo Casto, and M. Migliore. Car sharing demand estimation and urban transport demand modelling using stated preference techniques. *European Transport*, (40):33–50, 2008.

E. Cherchi and D.A. Hensher. Workshop synhesis: Stated preference surveys and experimental design, an audit of the journey so far and future research perspectives. *Transportation Research Propedia*, 11:154–164, 2015.

M. Chernew, G. Gowrisankaran, and D.P. Scanlon. Learning and the value of information: Evidence from health plan report cards. *Journal of Econometrics*, 144(1):156–174, May 2008.

A.T. Ching. A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review*, 51(4):1175–1207, November 2010.

V. Curto, L. Einav, J. Levin, and J. Bhattacharya. Can health insurance competition work? evidence from medicare advantage. *Journal of Political Economy (R and R)*, 2018.

L. Dafny and D. Dranove. Do report cards tell consumers anything they don't already know? the case of medicare hmos. *RAND Journal of Economics*, 39(3):790–821, Autumn 2008.

C.M. Dalton, G. Gowrisankaran, and R. Town. Salience, myopia, and complex dynamic incentives: Evidence from medicare part d. *Technical Report, National Bureau of Economic Research*, 2018.

M. Darden and I.M. McCarthy. The star treatment: Estimating the impact of star ratings on medicare advantage enrollments. *Journal of Human Resources*, 50(4):980–1008, Fall 2015.

A. Deaton. Price elasticities from survey data: Extensions and indonesian results. *Journal of Econometrics*, 44(3): 281–309, 1990.

F. Decarolis and A. Guglielmo. Insurer's response to selection risk: Evidence from medicare enrollment reforms. *Journal of Health Economics*, 56:383–396, 2017.

F. Decarolis, A. Guglielmo, and C. Luscombe. Open enrollment periods and plan choices. *Working paper*, 2017.

M. Duggan, A. Starc, and B. Vabson. Who benefits when the government pays more? pass-through in the medicare advantage program. *Journal of Public Economics*, 141:50–67, 2016.

Abe Dunn. The value of coverage in the medicare advantage insurance market. *Journal of Health Economics*, 29(6): 839–855, December 2010.

D. Earnhart. Combining revealed and stated data to examine housing decisions using discrete choice analysis. *Journal of Urban Economics*, 51(1):143–169, January 2002.

T. Erden, M.P. Keane, and B. Sun. A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*, 27(6):1111–1125, November 2008.

K.M.M. Ericson. Consumer inertia and firm pricing in the medicare part d prescription drug insurance exchange. *American Economic Journal: Economic Policy*, 6(1):38–64, February 2014.

Y. Fan. Ownership consolidation and product characteristics: A study of the us daily newspaper market. *The American Economic Review*, 103(5):1598–1628, August 2013.

M. Fioretti and H. Wang. Quality, quality payments, and risk selection in private medicare. *Working paper*, 2018.

A. Goolsbee and A. Petrin. The consumer gains from direct broadcast satellite and the competition with cable tv. *Econometrica*, 72(2):351–281, March 2004.

P.E. Green, A.M. Krieger, and Y. Wind. Thirty years of conjoint analysis: Reflections and prospects. *Interfaces (Marketing Engineering)*, 31(3):56–73, 2001.

M. Grennan and R. Town. Regulating innovation with uncertain quality: Information, risk, and access in medical devices. *NBER Working Paper*, 2019.

B. Handel 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*, 32(1):155–178, Winter 2018.

B. Handel, J.T. Kolstad, and J. Spinnewijn. Information frictions and adverse selection: Policy interventions in health insurance markets. *Review of Economics and Statistics, forthcoming*, 2018.

B.R. Handel. Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, 103(7):2643–2682, December 2013.

B.R. Handel and J.T. Kolstad. Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *The American Economic Review*, 105(8):2449–2500, August 2015.

F. Heiss, D. Mcfadden, and J. Winter. *Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans*, chapter Reseearch Findings in the Economics of Aging, pages 413–481. The University of Chicago Press.

K. Ho, J. Hogan, and F.S. Morton. The impact of consumer inattention on insurer pricing in the medicare part d program. *RAND Journal of Economics*, 48(4):877–905, Winter 2017.

G.Z. Jin and P. Leslie. The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451, May 2003.

G.Z. Jin and A.T. Sorensen. Information and consumer choice: The value of publicized health plan ratings. *Journal of Health Economics*, 25(2):248–275, March 2006.

J.D. Ketcham, C. Lucarelli, E.J. Miravete, and M.C. Roebuck. Sinking, swimming, or learning to swim in medicare part d. *American Economic Review*, 102(6):2639–2673, October 2012.

J.D. Ketcham, C. Lucarelli, and C.A. Powers. Paying attention or paying too much in medicare part d. *American Economic Review*, 105(1):204–233, January 2015.

J.T. Kolstad and M.E. Chernew. Quality and consumer decision making in the market for health insurance and health care services. *Medical Care Research and Review*, 66(1):28S–52S, February 2009.

E.P. Kroes and R.J. Sheldon. Stated preference methods: An introduction. *Journal of Transport Economics and Policy*, 22(1):11–25, January 1998.

T.J. Layton and A.M. Ryan. Higher incentive payments in medicare advantage's pay-for-performance program did not improve quality but did increase plan offerings. *Health Services Research*, 50(6), December 2015.

G. Lewis and G. Zervas. The welfare impact of consumer reviews: A case study of the hotel industry. *Working paper*, 2016.

J. Lustig. Measuring welfare losses from adverse selection and imperfect competition in privatized medicare. *Manuscript*, 2010.

T. Madeira. The cost of removing deadlines: Evidence from medicare part d. *JMP*, January 2015.

J. J. McCall. Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126, February 1970.

I.M. McCarthy and M. Darden. Supply-side responses to public quality ratings. *American Journal of Health Economics*, 3(2):140–164, Spring 2017.

D. McFadden. *Handbook of Choice Modelling*, chapter The New Science of Pleasure: Consumer Choice Behavior and the Measurement of Well-Being. Edward Elgar, 2013.

D. McFadden, A.C. Bemmaor, F.G. Caro, J. Dominitz, B.-H. Jun, A. Lewbel, R.L. Matzkin, F. Molinary, N. Schwartz, R.J. Willis, and J.K. Winter. Statistical analysis of choice experiments and surveys. *Marketing Letters*, 16(3/4): 189–196, 2005.

T.G. McGuire, J.P. Newhouse, and A.D. Sinaiko. An economic history of medicare part c. *The Milbank Quarterly*, 89(2):289–332, June 2011.

K. Miller. Do private medicare firms have lower costs? *mimeo*, 2015.

K.S. Miller, A. Petrin, R. Town, and M. Chernew. Optimal managed competition subsidies. *NBER Working Paper*, (25616), February 2019.

O.A. Morgan, J.C. Whitehead, W.L. Huth, and G.S. Martin. A split-sample revealed and stated preference demand model to examine homogenous subgroup consumer behavior responses to information and food safety technology treatments. *Working paper*, 2013.

K. Nosal. Estimating switching costs in the market for medicare advantage plans. *Working paper*, 2012.

D.J. Phaneuf, L.O. Taylor, and J.B. Braden. Combining revealed and stated preference data to estimate preferences for residential amenities: A gmm approach. *Land Economics*, 89(1):30–52, February 2013.

M. Polyakova. Regulation of insurance with adverse selection and switching costs: Evidence from medicare part d. *American Economic Journal: Applied Economics*, 8(3):165–95, July 2016.

R.O. Reid, P. Deb, B.L. Howell, and W.H. Shrank. Association between medicare advantage plan star ratings and enrollment. *The Journal of the American Medical Association*, 309(3):267–274, 2013.

M. Shum. Does advertising overcome brand loyalty? evidence from the breakfast-cereals market. *Journal of Economics and Management Strategy*, 13(2):241–272, 2004.

J. So. Adverse selection, product variety, and welfare. evidence from medicare advantage. *jmp*, November 2019.

Z. Song, M.B. Landrum, and M.E. Chernew. Competitive bidding in medicare advantage: Effect of benchmark changes on plan bids. *Journal of Health Economics*, 32(6):1301–1312, 2013.

G. J. Stigler. The economics of information. *The Journal of Political Economy*, 69(3):213–225, June 1961.

L. Sun. Welfare impacts of supply-side regulation in medicare advantage market. *Working paper*, 2017.

R. Town and S. Liu. The welfare impact of medicare hmos. *The RAND Journal of Economics*, 34(4):719–736, Winter 2003.

K. Train and W.W. Wilson. Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research*, Part B(42):191–203, 2008.

G.J. Wedig and M. Tai-Seale. The effect of report cards on consumer choice in health insurance market. *Journal of Health Economics*, 21:1031–1048, 2002.

Y. Wu. Supply response to consumer inertia: Strategic pricing in medicare part d. *Working paper*, January 2016.

L. Yang, C.F. Choudhury, and M. Ben-Akiva. Stated preference survey for new smart transport models and services: Design, pilot study and new revision. *Working paper*, 2009.

J. Yeo and D.P. Miller. Estimating switching costs with market share data: An application to medicare part d. *International Journal of Industrial Organization*, 61:459–501, November 2018.

A Appendix

A.1 Tables

Table 5. **Individual Performance Measures**

Domain	Measure ID	Measure Name	Weights
Staying Healthy	C01	Colorectal Cancer Screening	1
	C02	Cardiovascular Care - Cholesterol Screening	1
	C03	Diabetes Care - Cholesterol Screening	1
	C04	Annual Flu Vaccine	1
	C05	Improving or Maintaining Physical Health	3
	C06	Improving or Maintaining Mental Health	3
	C07	Monitoring Physical Activity	1
	C08	Adult BMI Assessment	1
Managing Chronic Conditions	C09	Special Needs Plan Care Management	1
	C10	Care for Older Adults - Medication Review	1
	C11	Care for Older Adults - Functional Status Assessment	1
	C12	Care for Older Adults - Pain Assessment	1
	C13	Osteoporosis Management in Women who had a Fracture	1
	C14	Diabetes Care - Eye Exam	1
	C15	Diabetes Care - Kidney Disease Monitoring	1
	C16	Diabetes care - Blood Sugar Controlled	3
	C17	Diabetes care - Cholesterol Controlled	3
	C18	Controlling Blood Pressure	3
	C19	Rheumatoid Arthritis Management	1
	C20	Improving Bladder Control	1
	C21	Reducing the Risk of Falling	1
	C22	Plan All - Cause Readmissions	3
Member Experience with the Health Plan	C23	Getting Needed Care	1.5
	C24	Getting Appointments and Care Quickly	1.5
	C25	Customer Service	1.5
	C26	Rating of Health care Quality	1.5
	C27	Rating of Health Plan	1.5
	C28	Care Coordination	1.5
Member Complaints and Changes in the Health Plan's Performance	C29	Complaints about the Health Plan	1.5
	C30	Members Choosing to Leave the Plan	1.5
	C31	Health Plan Quality Improvement	5
Health Plan Customer Service	C32	Plan Makes Timely Decisions about Appeal	1.5
	C33	Reviewing Appeals Decisions	1.5
Drug Plan Customer Service	D01	Appeals Auto - Forward	1.5
	D02	Appeals Upheld	1.5
Member Complaints and Changes in the Health Plan's Performance	D03	Complaints about the Drug Plan	1.5
	D04	Members Choosing to Leave the Plan	1.5
	D05	Drug Plan Quality Improvement	5
Member Experience with the Drug Plan	D06	Rating of Drug Plan	1.5
	D07	Getting Needed Prescription Drugs	1.5
Drug Safety and Accuracy of Drug Pricing	D08	MPF Price Accuracy	1
	D09	High Risk Medication	3
	D10	Diabetes Treatment	3
	D11	Medication Adherence for Diabetes Medications	3
	D12	Medication Adherence for Hypertension	3
	D13	Medication Adherence for Cholesterol (Statins)	3

Notes: This table describes all the domains and the individual metrics corresponding to them, along with the weights each metric is assigned in the calculation of summary Part C, D and overall MA-PD rates.

Table 6. **Medicare Beneficiary Individual-Level Summary Statistics**

	All Observations		By MA enrollment	
Variable	mean	sd	MA	TM
MA enrollment indicator	0.346	0.475	1	0
Demographics				
Income ($)	33982	79408	32484	34239
Age	72.4	10.9	73.13	72.3
Female	0.556	0.496	0.561	0.554
Black	0.079	0.270	0.087	0.075
Hispanic	0.019	0.137	0.027	0.014
Education				
Bachelor's degree	0.161	0.367	0.148	0.166
Attended college	0.204	0.403	0.211	0.199
High School	0.343	0.474	0.341	0.343
Health Status				
Excellent	0.157	0.364	0.162	0.155
Very good	0.282	0.450	0.297	0.274
Good	0.307	0.461	0.309	0.305
Fair	0.178	0.383	0.176	0.180
Poor	0.072	0.259	0.054	0.082
Observations	53267		17620	35647

Notes: This table displays the summary statistics of the Medicare sampled population resulted after the necessary data cleaning. An observation is defined as a person-year. All statistics reported are weighted by the sampling weights provided by the MCBS. The demographic categories are defined as in the CMS administrative data. The first two columns report means and standard deviations across all observations. The last two columns split the sample by MA enrollment.

Table 7. **Individual Summary Statistics by Star Rating**

Overall Star	Mean	Std. Dev.
2	0.001	0.035
2.5	0.059	0.236
3	0.276	0.447
3.5	0.332	0.471
4	0.136	0.343
4.5	0.178	0.383
5	0.015	0.121
Observations	3013	

Notes: This table displays the proportion of MA beneficiaries that chose a plan with a certain level of star rating. The number of observations reduces to 3013 because the SRS begins after 2008 (while the data set covers the period 2006-2016) and because star ratings are not available for all MA plans in the data set.

Table 8. **Plan Characteristics by Star Rating**

	Star Ratings					
Variable	**2.5**	**3**	**3.5**	**4**	**4.5**	**5**
Part C premium ($/month)	19.38	14.70	27.46	36.63	48.40	33.52
Part D premium ($/month)	14.67	14.84	19.10	24.26	29.33	23.36
OOPC ($/month)	320.47	302.64	326.08	342.53	378.51	376.73
Supplemental Coverage						
Prescription drugs	.78	.86	.84	.80	.76	.79
Dental	.54	.50	.35	.38	.39	.28
Vision	.82	.82	.84	.84	.87	.74
Hearing	.70	.65	.65	.68	.73	.56
Plan types						
HMO	.42	.55	.55	.63	.72	.26
PPO	.25	.28	.31	.29	.25	.32
PFFS	.32	.15	.13	.07	.01	.41
Observations	1707	3874	3731	1625	1247	1678

Notes: This table shows summary statistics on plan characteristics by each different level of Star Rating. An observation is a year-contract-plan. Information on plans receiving less than 2.5 star ratings is not reported because of limited availability of such plans in the data.

TABLE 9. **Comparison of Demographics between MCBS and Qualtrics' Samples**

	MCBS		Own Survey	
Variable	**MA**	**TM**	**MA**	**TM**
MA enrollment indicator	1	0	1	0
Demographics				
Income ($)	32484	34239	25963	28520
Age	73.13	72.3	71.75	71.96
Female	.561	.554	.482	.509
Black	.087	.075	.012	.028
Hispanic	.027	.014	.006	0
Education				
Bachelor's degree	.148	.166	.317	.355
Attended college	.211	.199		
High School	.341	.343	.304	.346
Health Status				
Excellent	.162	.155	.097	.132
Very good	.297	.274	.475	.339
Good	.309	.305	.310	.386
Fair	.176	.180	.113	.113
Poor	.054	.082	.003	.018
Observations	17620	35647	309	106

Notes: This table compares demographics between the sampled population in the MCBS and Qualtrics' data. MA observations are more than TM observations in my survey's sample comparing to the relevant proportion in the MCBS sample because of the sample stratification I imposed. Also, information for the "college attainment" variable is not reported for the Qualtrics sample as the relevant variable did not exist in the corresponding question I constructed.

Table 10. **Awareness demographic variation by MA/ TM population**

Variable	**Medicare Advantage** Aware	**Medicare Advantage** Unaware	**Traditional Medicare** Aware	**Traditional Medicare** Unaware
Demographics				
Age	71.591	71.834	70	72.260
Female	0.448	0.497	0.357	0.532
Black	0.030	0.004	0	0.032
Hispanic	0	0.009	0	0
Education				
High School	0.244	0.331	0.214	0.358
Bachelor's degree	0.397	0.279	0.357	0.347
Graduate degree	0.193	0.227	0.357	0.163
Health Status				
Excellent	0.132	0.080	0.357	0.097
Very good	0.479	0.473	0.357	0.336
Good	0.306	0.312	0.214	0.413
Fair	0.081	0.127	0.071	0.119
Poor	0	0.004	0	0.021
Plan Characteristics				
Monthly premium	61.765	90.734		
Star rating	4.483			
Observations	98	211	14	92

Notes: This table shows how the different consumer types on both the TM and MA populations varied regarding their demographic information.

Table 11. **Attributes for Conjoint Experiments**

Attribute	**Levels**
Prices	0, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 37, 39, 41, 42, 43, 44, 46, 49, 50, 51, 52, 53, 54, 56, 57, 60, 61, 62, 63, 64, 66, 67, 69, 70, 71, 72, 73, 75, 78, 80, 81, 82, 87, 88, 89, 99, 101
Star Ratings	2, 2.5, 3, 3.5, 4, 4.5, 5

Notes: This table presents all the possible values of the attributes presented to respondents in the conjoint experiments.

TABLE 12. **More Conjoint Experiment Estimates**

	(1)		(2)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}^s_c$	-0.0931	-0.1047	-0.0843	-0.0971
	(0.0148)	(0.0121)	(0.0149)	(0.0137)
$\sigma^s_{\alpha,c}$	0.0688	0.0592	0.0611	0.0567
	(0.0142)	(0.0105)	(0.0133)	(0.0105)
$\overline{\beta}^s_c$	2.5017	2.8749	2.3728	2.6426
	(0.3814)	(0.2892)	(0.3395)	(0.3324)
$\sigma^s_{\beta,c}$	0.4771	0.6368	0.3633	0.5645
	(0.8294)	(0.5018)	(0.7543)	(0.5105)
ϕ	0.9402	0.7366	0.9287	0.7619
	(0.0764)	(0.0609)	(0.0649)	(0.0674)
$\overline{\alpha}^s_{nc}$			-0.1552	-0.1444
			(2.6373)	(0.0716)
$\sigma^s_{\alpha,nc}$			0.8743	0.0783
			(9.5293)	(0.0640)
Observations	112	303	112	303
Assumptions				
$\overline{\alpha}^s_c = \overline{\alpha}^s_{nc}$, $\sigma^s_{\alpha,c} = \sigma^s_{\alpha,nc}$	YES		NO	
$\rho = 0$	YES		YES	
Monthly star value given $\phi_i = 1$	\$26.871	\$27.458	\$28.147	\$27.215
Overall monthly star value	\$25.264	\$20.225	\$26.140	\$20.735

Notes: This table presents the estimates of the respondents' distribution of preferences for prices and star ratings across different specifications. Standard errors are reported in parentheses.

TABLE 13. **More Conjoint Experiment Estimates**

	(1)		(2)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}^s_c$	-0.0808	-0.0864	-0.0687	-0.0845
	(0.0108)	(0.0081)	(0.0107)	(0.0086)
$\overline{\beta}^s_c$	2.4598	2.1322	2.3400	2.0869
	(0.5628)	(0.4423)	(0.3888)	(0.4005)
$\sigma^s_{\beta,c}$	1.6593	1.8239	1.1757	1.7580
	(0.4941)	(0.2831)	(0.3998)	(0.3300)
ϕ	0.9598	0.9598	0.8635	0.9130
	(0.1863)	(0.1383)	(0.1061)	(0.1314)
$\overline{\alpha}^s_{nc}$			-0.1637	-0.1193
			(0.1975)	(0.1105)
Observations	112	303	112	303
Assumptions				
$\overline{\alpha}^s_c = \overline{\alpha}^s_{nc}$, $\sigma^s_{\alpha,c} = \sigma^s_{\alpha,nc}$	YES		NO	
$\sigma^s_{\alpha,c} = 0$, $\sigma^s_{\alpha,nc} = 0$	YES		YES	
$\rho = 0$	YES		YES	
Monthly star value given $\phi_i = 1$	\$30.443	\$24.678	\$34.061	\$24.697
Overall monthly star value	\$29.219	\$23.686	\$29.411	\$22.548

Notes: This table presents the estimates of the respondents' distribution of preferences for prices and star ratings across different specifications. Standard errors are reported in parentheses.

A.2 Figures

To help beneficiaries find the insurance plan that best matches their needs, Medicare rates Medicare Advantage plans on a "star" scale from 1 to 5, with higher stars indicating higher quality.

Every year before the enrollment period begins, each plan is assigned an Overall Star Rating that indicates different levels of quality in terms of health outcomes of the people who enroll in the plan, the way plans help enrollees manage their chronic conditions, members' experiences with the plan, access to medical care, as well as customer service.

Do you remember seeing/ hearing/ reading about Overall Star Ratings for Medicare Advantage plans?

(Please select one option below)

Yes

No

FIGURE 2. **Identifying "*aware/ unaware*" Consumer Types**

If the two plans presented below were identical in every other way except for the following two characteristics, which one would you choose to enroll in?

Note: *There is no right or wrong answer. You should select the option that best reflects your personal preferences.*

(Please select one option below)

Plan 1	Plan 2
Monthly premium: $21	Monthly premium: $29
Overall Star Rating: 2	Overall Star Rating: 2.5

FIGURE 3. **HCE Example Question**

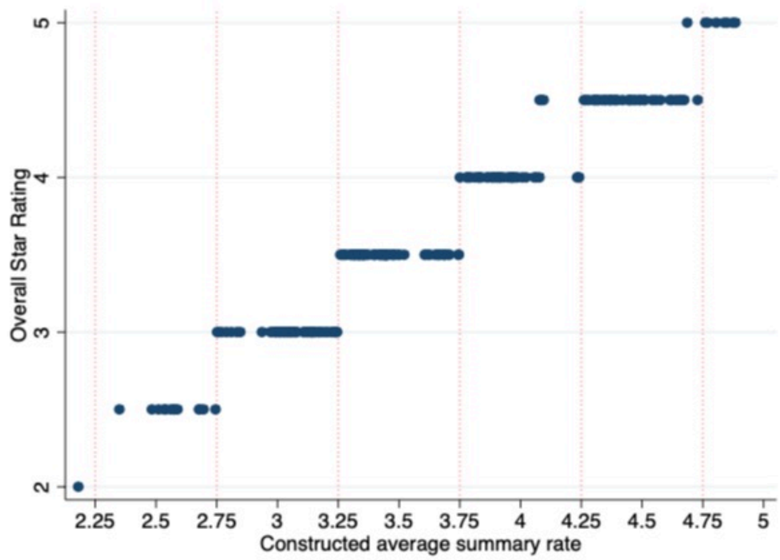


Figure 4. **Constructed Average Summary Rates**

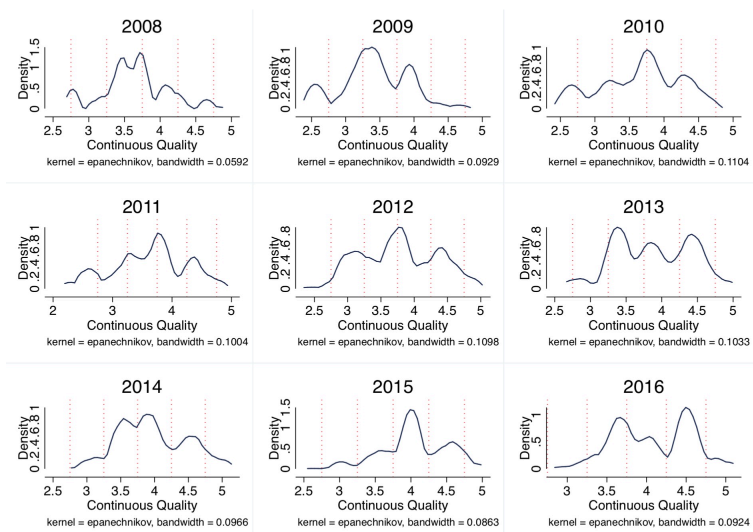


Figure 5. Evolution of overall continuous quality density over time (weighted by enrollment)

A.3 More on the Institutional Details

A.3.1 Star Rating System

This section describes the methodology followed by CMS for the calculation of star ratings.

- CMS classifies contracts into three types:
  1. MA-only (offers Part C benefits only)
  2. PDP (offers Part D benefits only)
  3. MA-PD (offers both Part C and D benefits)
- Each contract type is rated over a number of quality and performance measures (MA-only up to 32, PDP up to 15, and MA-PD up to 44[78]) that span five broad categories consistent with CMS's goals. These categories are the following:
  1. outcomes that refer to a beneficiary's health resulting from the provided care,
  2. intermediate outcomes that help move closer to true outcomes,
  3. patient experience that refers to a beneficiaries' perspective on the care they received,
  4. access that refers to any issues that may create obstacles in receiving the needed care, and
  5. process that refers to the method by which health care is provided.
- Every year, CMS reviews the measures constituting these categories, and depending on their reliability, potential data issues and other received feedback, it makes changes on the current measures, deletes and/ or adds redundant and/ or more appropriate measures, respectively.
- Star rates are reported in five different levels.[79] These are:
  1. base level that reflects individual measures comprised of numeric data (percentage scores),
  2. star level that reflect star rates calculated based on algorithms converting base level measure rates on a 5-star scale,

[78] 44 is the sum of 32 and 15 after subtracting 3 measures that overlap between MA-only and PDP contract types.

[79] This description is based on the latest version of the 2015 data available. Previous years' forms of the reported data were similar with the difference that before 2011 overall star rates are not reported.

3. domain level that reflects a second level into which each measured id grouped with similar measures (at a star level),[80]
4. summary Part C (Part D) level that reflects measures (at a star level) grouped together to form the Part C (Part D) summary for a contract, and
5. overall level that reflect Part C and Part D measures (at a star level) grouped together.

- The domain rate is the unweighted mean of the individual star ratings. To receive a domain rate, the contract must meet or exceed a minimum requires number of individual metrics.
- The summary Part C, Part D, and overall MA-PD star ratings are weighted averages of the individual star ratings. To receive a Part C, and/ or a Part D summery rate, a contract must meet a minimum required number of individual metrics. For the Part C and D summary rates, half stars are also assigned to allow more variation across contracts.
- Lastly, for MA-PD contract types to receive an overall rate, the contract must have stars assigned to both Part C and D summary rates, and the overall star rating is calculated using a weighted average of the Part C and D summary rates. For the overall star rates, half stars are also assigned to allow more variation across contracts.

A.3.2 Recovering Continuous Levels of Quality

To recover continuous levels of quality, I follow the instructions along with the necessary data provided by CMS, and I adjust it according to the year.[81] In what follows, I describe the main process.

- Summary Part C and D rates are calculated by taking a weighted average of the individual Part C and D metrics, respectively.
- Among all the metrics, the "improvement" metric plays a crucial role in the final assignment of rate.
- To reward consistently high performance contracts, CMS uses both the mean and the variance of individual metrics. Specifically, an integration factor is calculated and added to the mean score for the reward of high performing for a long period of time contracts.

[80]Totally, there are 9 domains comprised of up to 47 measures. MA-only contracts are measured on 5 domains, PDP contracts on 4 domains, and MA-PD contracts on all 9 domains.

[81]In a later version of the draft, I will also provide tables and graphs of the distributions of the continuous levels of quality as they arise when I keep the algorithm of the SRS constant over time and only changing the base-year algorithm.

- A few overlapping metrics between the Part C and D individual metrics are excluded for the calculation of the overall star ratings.
- In what follows, I describe the star ratings recovery process:

1. Classify contracts as “MA only”, “PDP”, and “MA-PD” if offering only Part C, only Part D, and both Part C and Part D benefits, respectively.
2. Generate the weights each individual metric receives.
3. Calculate the “MA-only” rate:
    (a) Calculate the Part C weighted **average** w/ and w/o the improvement metric.
    (b) Calculate the Part C weighted **variance** w/ and w/o the improvement metric.
    (c) Categorize both versions of weighted average as (i) below 65th pctile, or (ii) above 65th and below 85th pctile, or (iii) above 85th pctile.
    (d) Categorize both versions of weighted variance as (i) below 30th pctile, or (ii) above 30th and below 70th pctile, or (iii) above 70th pctile.
    (e) Develop the integration factor for both versions (w/ and w/o improvement) depending on the mean and variance categories generated above.
    (f) Add integration factor to the mean score for both versions.
    (g) If a contract has Part C weighted average without the improvement measure less or equal to 2, keep that level of quality. If a contract has Part C weighted average without the improvement measure greater or equal to 4, keep the maximum of the Part C weighted average with and without the improvement measure. For all other cases, keep the Part C weighted average with the improvement measure.
4. Calculate the “PDP” rated following the same steps as in “MA-only”.
5. Calculate the overall “MA PD” rate: follow the same steps as in “MA only” and “PDP” cases excluding the overlapping metrics.
6. Create Part C, D and overall scores rounding to the nearest half star.

A.4 Survey

In this section, I describe how I designed and executed the electronic survey I conducted in more detail. The full questionnaire of the survey will be posted in a later version of the draft. The survey consisted of a set of 5 sets of questions. The first set of questions filtered respondents so that they qualify for the target population and classified respondents as TM versus MA beneficiaries. The second set of questions investigated consumer awareness with respect to the SRS. The third set of questions conducted the conjoint experiments. The fourth set of questions investigated the general process consumers follow before they make their insurance plan choice. Finally, the fifth set of questions collected demographic information of the respondents.

A.4.1 Filtering respondents

In order to ensure that the survey respondents qualified the targeted Medicare population each respondent had to go through a series of questions. First, they had to report their age. Respondents below the age of 65 were excluded from the sample. Second, they were asked to report the state of their current residence to ensure there was enough geographic variation in the sample and to ensure that the respondents were indeed US residents. Respondents that reported they lived in Alaska were excluded from the sample, as this group of the population was also excluded in my main analysis. As long as they passed the first two steps of validation, they had to report what was the main source–TM versus MA–from which they received their health insurance coverage. To ensure that there would not be confusion regarding the type of coverage this question was referring to, I provided respondents example insurance cards explaining in detail what the differences between the corresponding options were and I asked them to choose the option that applied to them. In this question, I also included a third option, "Other", to account for the fact that there could be some respondents that although they would qualify for Medicare, they would receive their coverage from another source. Respondents who chose "Other" were excluded from the sample. The final step of screening was related to whether the insurance plans the respondents were enrolled in were employer sponsored. In such cases, respondents were also excluded from the sample.[82]

[82]Disabled population should also be excluded from the sample. However, the addition of such a question would complicate the screening process due to the institutional complications of the industry and the properties that qualify a beneficiary to be classified as "disabled". Moreover, it is common to avoid such questions when conducting surveys (especially at the beginning of the survey) as they can affect the answers respondents might provide and lead to selection. Hence, I did not include a question relevant to disability.

A.4.2 Eliciting information on information

I investigated consumer awareness with respect to the SRS with a main question in which after I described to respondents the SRS and its scope I asked them to report whether they had ever heard of it. To make sure that the information I provided about the SRS on this question would not affect respondents' answers, I followed the exact description the PlanFinder website gives. MA respondents that reported they were aware of they SRS, they were also asked to report the star rating of their current plan. Also, in this set of questions I investigated whether respondents were aware of the SRS when they firstly made their MA plan choice (in case they had ever chosen an MA plan).

A.4.3 Conjoint experiments

I used this set of questions to elicit respondents' preferences for star ratings relative to prices. Detailed description of the experiment is given in the main sections of the paper. A comment that is important to add, here, is that the experiment was conducted after ensuring that all types of respondents–aware and unaware–were equally informed about the SRS and the average quality of the plans they could find in the market. Hence, I repeated what the SRS and its scope were and I provided a chart that showed the percentage of MA plans that were receiving certain levels of star ratings in a local market.

A concern that might arise regarding the experiment is the trade-off between the realism of the options and the ability of respondents to report their preferences in a short time. If the characteristics provided are too numerous (generally more than 6 in a given choice) respondents tend to choose the ones that look simpler and more important to them without paying attention to the rest. In my conjoint analysis there are not such concerns, since I included only two attributes for each choice I provided to respondents.

A.4.4 Investigating the general process of a choice decision

To have enough information for a series of robustness checks that I wanted to include in my analysis, I included a set of questions that investigated the general process Medicare beneficiaries usually follow when they make their annual plan enrollment decision. First, I asked whether respondents are receiving any help when they try to decide which plan to enroll in and if they do so, I further asked them who and how they helped them giving them a series of options in each case. In this market it is well known that the aged population receives help from their relatives, friends or even independent agents/ brokers. Hence, their choices might not reflect what they actually know about the SRS and controlling for that is beneficial. Second, I asked them to rate the importance they place on different plan characteristics–premium, copay/ coinsurance, networks, star ratings, drug coverage, extra benefits (dental, vision, hearing)–on a 5-point scale from "Not Important" to "Very Important" when they make their enrollment decision. The purpose of this question would be to investigate how important star ratings are comparing to other plan characteristics and whether awareness of the SRS affects the importance of any of them. Third, I asked them an open-ended question in which I encouraged them to tell me whether there was anything else that they took into consideration when they usually decide the plan they will enroll in.[83] Lastly, I asked them to report how satisfied they were (on a scale from 1-100) with TM or MA depending on the source of their coverage.

A.4.5 Collecting demographic information

In the last set of questions, I collected demographic information of my sample. Specifically, I asked them to report their health status, their gender, their ethnicity, their marital status, their monthly spending, and their education level providing them various options. In most cases the options I provided were the same as the ones provided by the MCBS so that I would be able to compare the samples between my survey and the MCBS. Lastly, I asked them to report the zip-code of their current residence. In addition to the necessary geographic variation I was able to check by asking this question, I was also able to use it in combination with the reported state of residence as a validation test for the quality of the data.

A.4.6 Data quality

Respondents were guaranteed that any of the information they would provide would be kept confidential and would only be published at an aggregate level. With the exception of the demographic questions for which people tend to be sensitive, all respondents were forced to answer all questions. Although this increased the risk of respondent dropping out of the survey, it also guaranteed that the collection of the information would be consistent and the data set fairly balanced.

To test the quality of the my survey data I followed a series of tests. First, I checked that indeed all respondents matched the screening questions. Second, I checked whether the state of their current residence was in alliance with the zip-code they provided. Third, I excluded respondents who spent less than 1/3 of the average length the survey would take a respondent. Fourth, for a group of respondents that

[83]See website for their open-ended answers. - TBD

were at the low end of the distribution of the length they spent to take the survey, I checked the extent to which they were choosing consistently the same options in the conjoint experiments and/ or whether consistently chose the same level of importance in all the plan characteristics of the relevant question I provided. Lastly, I checked whether the answers they provided in the open-ended questions were indeed related to the topic and did not reveal any confusions.

Concerns regarding potential inattention with respect to the survey might arise, overall. To minimize such concerns, I set the option to advance to the next question not to appear for the first few seconds each question is available and I recorded how long it took each respondent to take the entire survey and to respond each separate question, as well. The average length of the survey was 8.4min. The average time respondents spent on the first question that explain them the context and the goal of the survey was 40sec. The average time they spent to answer the awareness question was 30sec. Lastly, the average time they spent on the conjoint experiment was 1-2min with most of the time spent on the first two questions.[84]

[84]The relevant tables will be posted in a later version of the draft. Also, in a later version of the paper I will report tables in which I estimate the model separately among respondents who took longer-than-average and shorter-than-average times to respond. A case in which the estimates are identical will suggest that more attention does not affect preferences.