The Optimal Geographic Distribution of Managed Competition Subsidies

Keaton Miller*†
Univ. of Oregon

Amil Petrin
Univ. of Minnesota, NBER

Robert Town
Univ. of Texas-Austin, NBER

Michael Chernew
Harvard Univ., NBER

January 20, 2023

Abstract

When markets fail to provide socially optimal outcomes, governments often intervene through 'managed competition' where firms compete for per-consumer subsidies. We introduce a framework for determining the optimal market-level subsidy schedule that features heterogeneity in consumer preferences and inertia, and firms with heterogeneous costs that can set prices and product characteristics in response to changes in the subsidy. We apply it to the Medicare Advantage program, which offers Medicare recipients private insurance that replaces Traditional Medicare. We calculate counterfactual equilibria as a function of the subsidies by estimating policy functions for product characteristics from the data and solving for Nash equilibria in prices. The consumer-welfare-maximizing budget-neutral schedule increases per-beneficiary annual consumer welfare by 113% over the current policy (an aggregate increase of $5.2 billion per year) and is well-approximated with a linear rule using market-level observables.

JEL Codes: I13, I18, D43, D61, L10.

Keywords: Managed Competition, Industrial Organization, Optimal Subsidies, Medicare Advantage.

*Miller (corresponding author): keatonm@uoregon.edu; Petrin: petrin@umn.edu; Town: robert.town@austin.utexas.edu; Chernew: chernew@hcp.med.harvard.edu

†The authors would like to thank Mark Colas, Francesco Decarolis, Clare Evans, David Evans, Gautam Gowrisankaran, Ben Hansen, Kendall Houghton, Van Kolpin, Scott Petty, Jeremy Piger, Christopher Snyder, Glen Waddell, Wes Wilson, anonymous referees, and seminar and conference participants at the University of Wisconsin-Madison, University of Oregon, the University of Minnesota, UCLA, Harvard/MIT, Northwestern University, University of Pennsylvania, University of Maryland, New York University, University of Texas - Austin, the 2019 NBER Winter IO and Insurance Meetings, the International Industrial Organization Conference, North American and European Meetings of the Econometric Society, and the Asian Workshop on Econometrics and Health Economics for helpful comments and feedback. All errors are our own. We are grateful for support from the National Institute on Aging under grant P01 AG032952.

1 Introduction

When the private market's provision of a good or service deviates from the socially optimal outcome, welfare may be improved through government intervention. In some cases, an obvious intervention is the government provision of the good or service. However, there is a long-standing concern that government production programs can be inefficient, as government bureaucracies may lack incentives to efficiently design and deliver goods (McKean and Minasian, 1966). This concern has led to an alternative approach whereby governments can either regulate private firms or procure goods directly from those firms. These approaches raise a host of strategic and informational issues that make efficient implementation challenging (Laffont and Tirole, 1993).

An alternative strategy is for a government to provide subsidies to consumers who purchase goods from competing firms with the idea that profit motives and market pressures will push firms to provide the optimal quantity, variety, and quality at a price nearing marginal cost. This "managed competition" scheme is employed by the US government to provide health insurance where market failure is a long-standing concern (Arrow, 1963, Rothschild and Stiglitz, 1976). For example, under the Medicare Advantage program (MA) we study, Medicare beneficiaries can forgo Traditional Medicare (TM) fee-for-service benefits and enroll in a health plan offered by private insurers. Insurers design plans that are meaningfully differentiated. Plans differ, for example, in their financial characteristics (e.g. deductibles and copays), benefits offerings (e.g. dental coverage), provider network structure, and patient administrative burden. Insurers assume the financial and logistical responsibility for an enrollee's care and, in turn, receive a risk-adjusted per-capita payment from the government based on a county-specific "benchmark rate" that varies considerably across counties. Similar approaches are used in Medicare Part D and the insurance markets created by the Affordable Care Act (Gruber, 2017). Elements of this approach appear in education, where public, charter, and private primary and secondary schools compete on program offerings, education quality, and productive efficiency (Poterba, 1996, Hoxby, 2000), as well as housing policy, where construction is influenced by differences in tax credits across geographies (Baum-Snow and Marion, 2009).

In this paper, we develop an approach for calculating the optimal subsidy schedule across heterogeneous markets in managed competition settings. We take the government's budget constraint as exogenous—solving for the optimal budget amount is a much more challenging problem although we provide evidence on the efficiency of the current budget level. In our setup, firms choose prices and other product characteristics in response to the subsidy set by the government and other competitive conditions. Allowing firms to set non-price characteristics is important as the welfare impact of strategic interactions and policy interventions can differ in complex ways from settings in which prices are the only strategic variable (Spence, 1975, 1976). Consumers are heterogeneous and choose plans based on observable plan characteristics and unobservable (to us) plan-specific quality. To our knowledge, we are the first to study the optimal geographic distribution of the subsidies in a differentiated products environment in which firms can adjust both price and non-price characteristics in response to changes in the subsidy. More broadly, the approach we develop to solve for the counterfactual oligopoly equilibrium with many endogeneous product characteristics is applicable to other settings.

To formalize the problem, consider a government which seeks to maximize consumer welfare by allocating a fixed exogenous budget $\bar{B}$ across M markets denoted by m. Each market contains a measure of consumers. Under managed competition, the government chooses a schedule of market-level subsidies $\{B_m\}$ (i.e. the benchmark rates). Let $CS_{mi}(B_m)$ be the welfare for consumer i in market m and $GovExp_{mi}(B_m)$ be the government spending on that consumer as a function of the subsidy. The problem[1] we study (which we refer to as the "optimal subsidy schedule problem") is

$$\max_{\{B_m\}} \sum_{m=1}^{M} \int_i CS_{mi}(B_m)di \quad \text{s.t.} \quad \sum_{m=1}^{M} \int_i GovExp_{mi}(B_m)di = \bar{B}. \tag{1}$$

The solution to this problem, which we refer to as the "optimal subsidy schedule" or simply the "optimal policy", depends on the derivatives of the welfare and spending functions

[1] We do not include the cost of public funds in Equation (1) since as the budget is fixed reallocating subsidies does not change the cost of public funds. We also do not include firm profits as we view the government's normative objective as maximizing the direct well-being of its citizenry. Including firm profits in the problem is straightforward.

with respect to the benchmark, which, in turn, depend on equilibrium interactions between firms and consumers. The consumer welfare generated by a dollar increase in government spending — the government spending to consumer welfare pass-through rate — in any given market depends on the demand elasticity, the marginal valuation of the product characteristics, firms' cost functions, and the nature of competition. These objects likely vary across markets, and therefore the optimal subsidy should also vary across markets. Furthermore, if consumers in different markets have heterogeneous preferences over products and characteristics, the way in which firms change the design of non-price characteristics will also change consumer welfare and the optimal subsidy schedule.[2] In practice, however, the subsidy schedule may be determined by summary measures which do not take supply and demand factors fully into account. For example, MA benchmark rates are set as a function of average risk-adjusted county-level TM costs, which may differ from private firms' costs and may be unrelated to demand conditions. Similarly, many charter schools receive government funding based on the per-pupil cost of public schools in the area (Hoxby, 2000). Political dynamics can also affect subsidy rates (Adrion, 2020).

To solve for the optimal subsidy schedule, we must predict market outcomes under candidate subsidy schedules across all markets. The traditional approach to computing counterfactual equilibria is to search for a fixed-point in firms' best-response functions for prices and product characteristics (e.g. Fan, 2013, Wollmann, 2018). This approach is impractical in our setting due to the large number of markets, the often-large number of plans in each market, and the complexity of MA plans – plans choose premiums and 10 different product characteristics. Importantly, many plans are offered at a $0 premium which means modeling non-premium plan features is essential. We introduce a new counterfactual approach which we believe to be of independent methodological interest. We estimate policy functions for product characteristics from the data, use those estimated functions to predict characteristics under counterfactual benchmarks and then solve the firms' first-order conditions for premiums taking those characteristics as given. We provide conditions under

[2]In Appendix A, we introduce a simple conceptual model in which increases in a government subsidy leads a monopolist to both decrease prices and increase plan quality. Furthermore, we note that interactions between all of these effects are important. While a planner would seek to move resources to areas where consumers are more elastic *ceteris paribus*, if those areas also feature firms which do not pass-through subsidies to benefits at a high rate, it may be optimal to decrease the subsidies in those areas.

which the actions generated by our approach converge to equilibrium play as the estimated policy functions converge and Monte Carlo evidence that this approach well-approximates the equilibria calculated by explicitly solving the firms' best response functions.

We apply our approach to calculate the optimal county-level MA subsidy schedule. The first step is to estimate Medicare beneficiaries' preferences over MA plans. Using detailed individual panel data on consumer demographics, choice sets, realized choices, and aggregate market-level plan shares for the years 2008-2017, we estimate a flexible demand system. Our demographic variables include a self-reported health status, age, race, educational attainment, and income which allow us to capture plan preferences that vary with these variables. The panel nature of our data allows us to estimate switching costs, which are relevant due to the prevalence of narrow provider networks.

Our model implies that premiums and plan characteristics are endogenous which invalidates many standard instruments. We instrument for prices by assuming that some costs are correlated across plans offered by different insurers in the same county-year market. We leverage detailed plan cost data submitted to the government as part of the regulatory process and construct, for each plan, the mean costs of that plan's competitors for required coverage (i.e. the items and services covered by TM). We construct instruments for plan characteristics from the panel nature of the plan data by assuming that county-level year-over-year changes in the benchmark are approximately random and uncorrelated (which we test and verify as likely to hold).

Our estimates imply beneficiaries are on average price-sensitive with mean implied plan premium elasticities of -6.29. Higher income beneficiaries are less premium sensitive than lower income beneficiaries. MA plans are more attractive to younger Medicare recipients, non-Whites, and those with lower educational attainment. The estimated average switching cost between MA insurers is \$285, which is comparable to the enrollment-weighted average annual plan premium of \$415. The average consumer values an average dental, vision, and hearing coverage package at \$604 per year.

The second step in calculating the optimal benchmarks is to estimate plan marginal cost functions. We invert firms' implied first-order conditions for premiums and product characteristics to infer marginal costs accounting for regulatory requirements and the potential for plan-specific differences in the cost of providing a given benefit. The marginal cost function estimates align with independent utilization information. Our estimates imply that in 2017, MA generated a total of \$4.62 billion in consumer surplus as measured by aggregating individual-level compensating variation and \$3.10 billion in variable profit with \$126 billion in total government payments to MA plans and \$314 billion in spending on TM in our data.[3]

The final step prior to calculating the optimal benchmark is to estimate the plan characteristic policy functions. Our policy function estimates appear sensible – increasing the benchmark increases plan benefit provision and reduces patient out-of-pocket expenses. We combine the demand estimates, the plan marginal cost function, the policy functions and the implied premium setting first-order conditions together and use a multi-start search to solve Equation (1).

We find the optimal MA benchmarks are meaningfully different from the current benchmarks – the average absolute difference between the optimal benchmark and the 2017 benchmark schedules is \$486 or 4.95%. We find that the optimal benchmarks would have increased aggregate consumer welfare to \$9.84 billion per year through a combination of increasing the total share of MA from 29.8% to 43.2% and increasing the mean compensating variation for MA enrollees from \$368.90 to \$542.98. These gains are not evenly distributed – relative to the 2017 schedule, the optimal policy reduces benchmarks (and therefore consumer welfare) for 35.8% of beneficiaries. We explore other social welfare functions and find benchmark schedules that increase the aggregate consumer welfare and reduce the variance relative to the 2017 policy. We test our approximate counterfactual equilibria by calculating the profitability of deviations from our predicted choices and find that firms' first-order conditions are reasonably satisfied.

Changes in product characteristics are responsible for 35.4% of the total change in consumer welfare. We show that an analysis that does not take changes in product characteristics into account generates a meaningfully different policy from our endogenous-characteristic specification. Relative to the 2017 benchmark policy, the optimal policy creates winners and losers—the average compensating variation for MA enrollees in markets that receive a higher

[3]We exclude individuals who are 'dual-eligible' for Medicare and Medicaid and plans designed for those individuals from our analysis; see Section 4.

(lower) benchmark increases (decreases) by \$207 (\$18).

Finally, we show that the derivatives of the welfare and spending functions are related to market-level observables, and that a linear rule using these observables is able to obtain over 95% of the gains from the optimal policy with a 0.32% increase in spending.

To understand the intuition behind our results, it is helpful to note that our estimates imply that MA subsidies are, at the margin, too generous from a social surplus perspective. We find that an across-the-board \$1 increase in benchmarks from the 2017 policy increases MA expenditures by \$182 million, decreases TM expenditures by \$173 million, and increases aggregate consumer welfare by \$7.84 million. This, plus the large variation in benchmarks, implies that the marginal utility of a dollar of benchmark (mediated through the competitive interactions of plans) is likely relatively low in current high payment areas whereas in lower payment areas the marginal utility of an extra dollar of benchmark is higher. Roughly, reallocating subsidies from high benchmark areas to low can increase welfare because of the differential marginal value. The oversimplification in the above discussion is that we allow the marginal value of an increase in benchmark (to enrollees) and the plans' marginal cost of providing benefits to vary across markets so that our optimal policy may in fact reallocate benchmark away from some low benchmark counties if their marginal utility from an increase in the benchmark is low.

We describe our approach to computing counterfactuals in Section 2. We discuss the institutional details of the MA program in Section 3 and detail our data on Medicare beneficiaries and MA plans in Section 4. We present a model of demand for MA in Section 5 and discuss the supply side in Section 6. We describe our estimation procedure in Section 7 and present estimates in Section 8. We implement our counterfactual approach and present our results in Section 9. We conclude in Section 10.

1.1 Literature review

We build upon an extensive MA literature; see McGuire et al. (2011) for a review. Our work is most related to Town and Liu (2003), Lustig (2010), Aizawa and Kim (2018) and Curto et al. (2021). Town and Liu (2003) estimate a nested logit demand system for MA plans and calculate that the program generated \$113 in consumer surplus and \$244 in profits per Medicare beneficiary in 2000 with significant geographic variation. Curto et al. (2021) estimate a related model using more recent data and find that the program generated approximately $2,600 in per-capita annual surplus, with the majority captured by insurers. They also estimate that average MA plan costs are 12% lower than TM costs, though in 47% of counties MA does not have a cost advantage over TM. We innovate with respect to these papers by adding rich micro-level data on demographics and plan product characteristics, allowing for endogenous premium and non-premium characteristics, and consider counterfactual subsidy policies that affect plan design decisions across both premium and many non-premium plan characteristics (e.g. different copays and benefit offerings).[4] These modeling differences appear to matter as our premium elasticity estimates are much larger in magnitude and our estimate of the consumer surplus created by the MA program is much smaller relative to Curto et al. (2021). Ultimately, Curto et al. (2021) (Table 3) find that an across-the-board increase in the benchmark increases welfare, while we find the opposite. Aizawa and Kim (2018) estimate a demand model that is similar to ours in order to explore the role of advertising in equilibrium selection. Our demand model innovates upon theirs by adding additional heterogeneity in switching costs.

There is also a literature examining the rate at which MA benchmark increases are passed through to consumers. Using an unanticipated change in the benchmark in 2000, Cabral et al. (2018) estimate a pass-through rate of 54%, while Duggan et al. (2016) use variation in the benchmark across urban and rural counties and estimate a smaller pass-through. Song et al. (2013) calculate a pass-through from benchmarks to plan bids, which are a measure of premiums and the actuarial value of benefits, of 53%. We expand upon this literature by considering firms' plan design decisions in response to benchmark changes and measuring consumer valuations of those plans.

Our work is also related to research on optimal subsidy structures in health insurance contexts. Tebaldi (2017), Jaffe and Shepard (2017) and Einav et al. (2018) examine the optimality of different subsidy and/or risk-adjustment strategies in different ACA insurance

[4]Perhaps most importantly, Curto et al. (2021) assume that the effect of premiums and plan characteristics on utility and costs can be captured by a one dimensional, endogenous measure of plan generosity – the plan bid. In contrast, we allow premiums and multiple plan characteristics to enter both the utility and cost functions flexibly.

exchanges. Ericson and Starc (2015) examine the implications of age-based premium regulation in an ACA-like insurance exchange. Bundorf et al. (2012) study health-status-linked premiums for employer-sponsored plans.[5] Decarolis et al. (2020) examine the optimality of using vouchers versus the current subsidy strategy in Medicare Part D and find that the two systems generate similar welfare. We innovate by examining the impact of subsidies on non-premium plan characteristics and calculating counterfactual outcomes.

Finally, our counterfactual approach builds on past efforts to use policy function estimation for counterfactual analysis. Goolsbee and Petrin (2004) study competition between pay television systems and use estimated functions for premiums and product characteristics to calculate the welfare gains caused by the introduction of satellite TV. Sweeting (2007) uses a similar approach to study radio station repositioning. Benkard et al. (2018) estimate strategic entry and exit behavior in the airline industry and simulate industry outcomes under counterfactual merger scenarios. We extend these efforts by combining our estimated policy functions with our demand model to solve for an equilibrium in premiums and calculate the welfare effects of policy changes.

2 An Approximation Approach To Counterfactual Analysis

Below we detail our approach to calculating counterfactuals when traditional approaches to computing an equilibrium cannot be implemented. Traditional counterfactual techniques generally solve a potentially high-dimensional, non-linear fixed-point problem. We instead use policy function estimation with (optionally) a reduced-dimension fixed-point problem. The intuition behind our approach is simple. We are interested in understanding the way in

[5] We do not model selection incentives. The importance of selection in these markets and the role of risk adjusted subsidies is well-known – see Geruso and Layton (2017) for a review. There is evidence that MA plans were able to successfully cream-skim lower cost enrollees prior to the period we study (Brown et al., 2014). However, before our sample period CMS implemented improvements in its risk-adjustment algorithm and other policies aimed at reducing selection incentives. The evidence indicates that these changes were largely successful and significantly reduced the incentive for plans to cream skim healthy enrollees (Newhouse et al., 2012, 2013, 2015, 2014, 2019a,b). We test for the presence of residual selection after risk-adjustment in our data and, consistent with the literature, fail to find meaningful selection in our data. While we do not model selection, we believe that with the appropriate cost data and some ancillary assumptions, our approach can be extended to model selection.

which agents in a game change their behavior in response to some variation in the payoff-relevant information available to them. If $X(\mathcal{I})$ is the vector of actions taken by agents when the information set is $\mathcal{I}$, we seek to predict the actions that would occur if the agents had instead faced some counterfactual information set $\mathcal{I}'$ which does not appear in the data. Suppose we observe data $\{X_m, \mathcal{I}_m\}_{m=1}^M$ generated by an equilibrium process which allows us to consistently estimate policy functions, potentially with existing approaches (Goolsbee and Petrin, 2004, Bajari et al., 2007). An estimate of the counterfactual actions can be formed by simply evaluating the estimated policy functions at the counterfactual information set. More concretely, our data consists of a panel observations on market environments and firm product characteristic decisions for many counties over several years. One member of $\mathcal{I}$ is the benchmark subsidy rate. We construct $\mathcal{I}'$ by changing this subsidy rate for each county (potentially to a rate which has never been observed in that county) holding all other elements of the market environment fixed.

This logic can be extended. Suppose now that either some elements of $\mathcal{I}'$ lie outside of the domain of the estimated policy function, or that we do not trust the quality of our policy function estimates for some elements of X (e.g. the estimates are noisy or the necessary assumptions for consistent estimates do not hold). We proceed by partitioning the agents' action vectors into subvectors of partial actions for which an estimator is available and those for which one is not. We estimate the available policy functions as above and search for an equilibrium in a restricted game where agents take the estimated action components as given using a traditional fixed-point approach. We show that as the number of observations grows and our estimate of the policy functions improves, this 'augmented' procedure converges to equilibrium play. The estimated partial actions converge to equilibrium play and the fixed-point algorithm fills in the rest of the equilibrium actions.

The remainder of this section establishes that convergence result. The main additional detail we must consider is the possibility that the counterfactual information set $\mathcal{I}'$ is itself estimated (or partially estimated) from the data. In our application, there are unobservable demand and cost characteristics in $\mathcal{I}$ that must be estimated after parameters are estimated. We present a Monte Carlo analysis of our technique in Section D. Readers uninterested in these details may wish to skip to Section 9 where we discuss the implementation of this method on our data.

2.1 The game

Consider a finite game with F players denoted $f = 1, \cdots, F$. Each player observes some payoff relevant information $\mathcal{I}_f$ which is a member of a compact finite-dimensional subset $\mathcal{S}_f$ of the real numbers. Let $X_f \in \mathcal{X}_f$ be the action taken by player f, where $\mathcal{X}_f$ is also a compact finite-dimensional subset of the real numbers. Let X_{-f} be a vector capturing the actions for players other than f. Payoffs for f are a function of their own action, the actions of others, and their knowledge: $\pi_f(X_f; X_{-f}, \mathcal{I}_f)$.

We define aggregate objects by dropping the f subscripts: $\mathcal{I}$ is the set of all of the $\mathcal{I}_f$, and is a member of $\mathcal{S}$, the set of all the $\mathcal{S}_f$. $\mathcal{S}$ is compact and finite-dimensional by construction. $\mathcal{X}$ is the product of the $\mathcal{X}_f$; $X \in \mathcal{X}$ denotes a vector capturing actions for each player, and so on. Given these definitions, we may write $\pi_f(X, \mathcal{I}_f)$ (or $\pi_f(X, \mathcal{I})$) for an individual player's payoff or $\pi(X, \mathcal{I})$ for a vector of payoffs.

A Nash equilibrium in this game given information $\mathcal{I}$ is a vector of actions $X^* = \{X_f^*\}_{f=1}^F$ such that $\pi_f(X_f^*; X_{-f}^*, \mathcal{I}_f) \geq \pi_f(X_f'; X_{-f}^*, \mathcal{I}_f)$ for all f and X_f'. We focus on pure strategy equilibrium and assume that the game is well-behaved in the following sense:

Assumption WB—Well Behaved: $\pi_f(X, \mathcal{I}_f)$ is smooth over $\mathcal{X}$ and $\mathcal{S}_f$ for all f. π admits a single equilibrium X^* for each information set $\mathcal{I}$. $\pi_f(X_f; X_{-f}^*, \mathcal{I}_f)$ is globally concave and has a single maximum at X_f^* for all f, which is interior.

This assumption implies that the equilibrium can be characterized by the conditions

$$\frac{\partial \pi_f}{\partial X_{fl}^*}(X_f^*; X_{-f}^*, \mathcal{I}_f) = 0 \text{ for all } f, l \tag{2}$$

where X_{fl}^* is the l-th element of X_f^*. Note that these conditions imply that $\frac{\partial^2 \pi_f}{\partial \mathcal{I}_f \partial X_f}$ is symmetric and $\left(\frac{\partial^2 \pi_f}{\partial X_f \partial X_f'}\right)^{-1}$ exists in some neighborhood around any equilibrium X^* for all f.

By the implicit function theorem, within that neighborhood

$$\frac{\partial^2 \pi_f}{\partial X_f \partial X_f'} \cdot \frac{\partial X_f}{\partial \mathcal{I}_f} + \frac{\partial^2 \pi_f}{\partial \mathcal{I}_f \partial X_f} = 0$$

$$\Rightarrow \frac{\partial X_f}{\partial \mathcal{I}_f} = -\left(\frac{\partial^2 \pi_f}{\partial X_f \partial X_f'}\right)^{-1} \left(\frac{\partial^2 \pi_f}{\partial \mathcal{I}_f \partial X_f}\right).$$

Thus, in some neighborhood around X^*, there is a one-to-one relationship from $\mathcal{I}_f$ to X_f for all f—in other words, given Assumption WB, the policy functions $X_f(\mathcal{I}_f)$ exist.[6]

2.2 Approximating equilibria with augmented policy function estimation

Suppose we collect data on players' actions and information sets $\{X_m, \tilde{\mathcal{I}}_m\}_{m=1}^M$ where $\tilde{\mathcal{I}}_m$ is some subvector of a full information set $\mathcal{I}_m$. This subvector notation is necessary as in practice researchers may not observe all of the elements of $\mathcal{I}$. As mentioned above, the information set in our application includes unobservable (to researchers) demand and cost characteristics. We are interested in approximating the Nash equilibrium outcomes for some information set $\mathcal{I}'$ that we do not observe. We begin by assuming the data we observe represent equilibrium outcomes.

Assumption DGP—Data Generating Process: The data consist of observations of actions and partial information sets $\{X_m, \tilde{\mathcal{I}}_m\}$. For each observation $(X_m, \tilde{\mathcal{I}}_m)$, there is a unique full information set $\mathcal{I}_m$ such that X_m is an equilibrium for $\mathcal{I}_m$.

The strength of this assumption depends on the underlying payoff functions. In applied work, as in our application, it is common to write models with unobservables that can 'rationalize' the data precisely (Rust, 1987, Berry et al., 1995). Estimates of these un-

[6] Without the single equilibrium component of Assumption WB this argument implies that policy correspondences exist. In this way, Assumption WB plays the role here that Assumption ES (equilibrium selection) plays for Bajari et al. (2007). In general, however, π_f may reach its maximum at many points in which case the policy correspondences are only guaranteed to be upper-hemicontinous. For example, a monopolist firm choosing prices and characteristics for multiple products may face a payoff function which is exchangeable in its arguments–i.e. the order of the products in the firm's payoff function does not matter, only the relative prices and characteristics. In this case, the assumption may be satisfied by imposing an ordering on the products e.g. that the products enter the payoff function in lexicographic order with respect to prices and product characteristics.

observables may then be calculated as 'residuals' from the parameter estimation process; if the parameters are estimated consistently, these estimates of the unobservables are also consistent. Thus, Assumption DGP may be satisfied by the assumptions underlying the estimation procedure. Note that under Assumption WB, X_m is the unique equilibrium for $\mathcal{I}_m$.

To characterize our estimation requirements, we suppose the possibility of a split in the game's action vector such that, given partial actions for each player, it is feasible to solve for the equilibrium in the reduced game played on the restricted action space. Define the *partition* of X_f as $X_f = (X_{1f}, X_{2f})$ where $X_{1f} \in \mathcal{X}_{1f}$ and $X_{2f} \in \mathcal{X}_{2f}$. Let $\mathcal{X}_1$ and $\mathcal{X}_2$ be the product of the $\mathcal{X}_{1f}$ and $\mathcal{X}_{2f}$, respectively. Note that the partition is over the elements of agents' actions and may differ across agents. We assume that consistent estimators of the full information set and the policy function for the X_1 are available.

Assumption CE—Consistent Estimation: Given data $\{X_m, \tilde{\mathcal{I}}_m\}_{m=1}^M$, there exists a continuous estimator of a full information set for each observation $\hat{\mathcal{I}}_m(\{X_m, \tilde{\mathcal{I}}_m\}_{m=1}^M) \equiv \hat{\mathcal{I}}_{Mm}$ which is consistent: $\text{plim}_{M\to\infty}\hat{\mathcal{I}}_{Mm} = \mathcal{I}_m$. Furthermore, there exists some continuous estimator $\hat{X}_1(\mathcal{I}; \{X_m, \tilde{\mathcal{I}}_m\}_{m=1}^M) \equiv \hat{X}_{1n}(\mathcal{I})$ which is consistent: for some information set $\mathcal{I}'$, $\text{plim}_{M\to\infty}\hat{X}_{1n}(\mathcal{I}') = X_1^*(\mathcal{I}')$ where $X_1^*(\mathcal{I}')$ is the partial action associated with the equilibrium strategy $X^*(\mathcal{I}')$.

Note that since $\hat{X}_{1n}$ is continuous, for some random sequence of information sets $\{\mathcal{I}_n'\}$ such that $\text{plim}_{M\to\infty}\mathcal{I}_n' = \bar{\mathcal{I}}'$, we have $\text{plim}_{M\to\infty}\hat{X}_{1M}(\mathcal{I}_n') = X_1^*(\bar{\mathcal{I}}')$. Per the above discussion, a consistent estimator of the information set may be available through the parameter estimation process. If WB is assumed, a policy function estimator may be obtained by approximating the policy function with basis functions over the (compact) space $\mathcal{S}$ (e.g. Chebyshev polynomials) that increase in degree more slowly than the data used to estimate their parameters, as long as the $\mathcal{I}'$ (or $\{\mathcal{I}_n'\}$) are in the basis space formed by the $\mathcal{I}_i$. Nonparametric techniques may also be used (Bajari et al., 2007). In the absence of WB, the strength of this assumption again depends on the underlying payoff functions; even if policy functions (as opposed to correspondences) exist, they may be discontinuous in which case establishing the consistency of an estimator may be challenging.

Given a set of partial actions X_1 for all agents and a partial action X_{1f} for a specific agent, we write X_{-1f} to denote the vector of partial actions for players other than f and analogously X_{-2f}. We define the *partitioned payoff function* $\pi^p_f(X_{2f}, X_{-2f}; X_1, \mathcal{I}_f) \equiv \pi_f(X_f; X_{-f}, \mathcal{I}_f)$ where X_f, X_{-f} on the right-hand side are formed by stacking the relevant members of X_1 and X_2. Under Assumption WB, π^p_f is smooth in its arguments and concave over a compact domain, so $\max_{X_{2f}} \pi^p_f(X_{2f}; \cdot)$ is well-defined and the first-order conditions for maximization of π^p_f are identical to elements of the first-order conditions for maximization of π_f:

$$\frac{\partial \pi^p_f}{\partial X_{2fl}} = \frac{\partial \pi_f}{\partial X_{2fl}} \text{ for all } f, l.$$

We define the *partitioned game* as the game with actions X_{2f} and payoffs π^p_f with the set of partial actions X_1 and information set $\mathcal{I}$ taken as given. Suppose there exists some vector-valued algorithm G which calculates an equilibrium in the partitioned game given X_1 and $\mathcal{I}$. We refer to G as a *partitioned solver* and denote the output of the algorithm as $X^G_2(X_1, \mathcal{I}) \equiv G(X_1, \mathcal{I})$. Since π^p_f is smooth and G solves the partitioned game, $\frac{\partial \pi^p_f}{\partial X^G_{2fl}} = \frac{\partial \pi_f}{\partial X^G_{2fl}} = 0$ for all f, l when evaluated at $X_1, \mathcal{I}$. We now have the ingredients necessary for our convergence result which we frame in terms of a sequence of counterfactual information sets as in practice elements of the counterfactual information set may be estimated.

Proposition 1. *Suppose Assumptions WB, DGP, and CE hold. Let G be a partitioned solver. Let $\{\mathcal{I}'_n\}$ be a random sequence of information sets with $plim_{M\to\infty}\mathcal{I}'_n = \bar{\mathcal{I}}'$ and $\bar{\mathcal{I}}' \in \mathcal{S}$. Let $\hat{X}(\mathcal{I}'_n)$ be the vector of actions formed by stacking $\hat{X}_1(\mathcal{I}'_n)$ with $X^G_2(\hat{X}_1(\mathcal{I}'_n), \mathcal{I}'_n)$. There exists some vector $\bar{X}$ such that $plim_{M\to\infty}\hat{X}(\mathcal{I}'_n) = \bar{X}$. $\bar{X}$ is a Nash equilibrium under information set $\bar{\mathcal{I}}'$.*

Proof. We first demonstrate the existence of $\bar{X}$. Note that under DGP and CE, the data $\{X_m, \tilde{\mathcal{I}}_m\}$ can be used to construct an estimate of the full information set for each observation. Let $\{X_m, \hat{\mathcal{I}}_{Mm}\}$ be the data used to construct $\hat{X}_1$ per CE. Since $\hat{I}_{Mm}$ is consistent and $\hat{X}_1$ is continuous, $\text{plim}_{M\to\infty}\hat{X}_{1M}(\mathcal{I}'_M)$ exists and is equal to $X^*_1(\bar{\mathcal{I}}')$. Since G is a partitioned solver, $X^G_2(\hat{X}_1(\mathcal{I}'_M), \mathcal{I}'_M) \equiv X^G_{2M}$ exists for all M. Under WB the policy functions are continuous in the information sets and therefore $\text{plim}_{M\to\infty}X^G_{2M}$ exists and is equal to $X^G_2(X^*_1(\bar{I}'), \bar{I}')$. $\bar{X}$ therefore exists.

Under WB, $\frac{\partial \pi_f^p}{\partial X_{2fl}^G} = \frac{\partial \pi_f}{\partial X_{2fl}^G} = 0$ for all f, l when evaluated at $X_1^*(\bar{\mathcal{I}}'), X_2^G(X_1^*(\bar{I}'), \bar{I}'), \bar{\mathcal{I}}'$. Furthermore, since $X_1^*(\bar{\mathcal{I}}')$ is the partial action associated with the equilibrium strategy $X^*(\bar{\mathcal{I}}')$ the first-order conditions with respect to X_1 are satisfied. Thus, the first-order conditions (2) are satisfied, and under WB, this implies $\bar{X}$ is the equilibrium strategy when the information set is $\bar{\mathcal{I}}'$. □

As there is no requirement on the dimension of the policy function estimator $\hat{X}_1$ or the partitioned solver G, this proposition encompasses both the traditional approach (where only G is used) and an approach in which only policy functions are used. Furthermore we note that in finite samples the counterfactual information vector $\mathcal{I}_n'$ may differ substantially from $\bar{\mathcal{I}}'$ and the action vector generated from our procedure $(\hat{X}_1(\mathcal{I}_n'), X_2^G(\hat{X}_1(\mathcal{I}_n'), \mathcal{I}_n'))$ may not satisfy the first-order conditions with equality. However, Assumption WB suggests a test. After computing approximate counterfactual actions, one may measure the extent to which the first-order conditions are satisfied. We do so for the approximated actions at our counterfactual policy in Table F.1. In Section D we conduct Monte Carlo exercises to explore the implications of these two points by measuring the performance of our approach when all elements of the action vector are approximated with policy functions and when a partitioned solver G is use. The results suggest that a G can significantly improve the approximation.

Finally, we note that frequently the objects of interest in counterfactual analyses are not players' actions themselves, but rather functions of those actions. For example, to solve the optimal subsidy problem, we must compute consumer welfare and government expenditures, which are functions of firm actions. As long as the relevant functions are continuous, our method produces consistent estimates of their values at the counterfactual information sets.

3 The Medicare Advantage Program

Enacted in 1965, Traditional Medicare (TM) provides health insurance to seniors (age 65 or older) through its Part A (hospital) and Part B (physician and outpatient) programs. Under TM, Medicare pays service providers according to a pre-set fee-for-service (FFS) reimbursement schedule while beneficiaries pay applicable copays and/or coinsurance. Eligibility has since expanded to include those eligible for federal disability benefits and end-stage renal disease (ESRD) patients.

In 1982, in response to the increasing federal budget burden of Medicare, Congress authorized Medicare administrators to engage in a series of "Part C" trials based on the ideas of Enthoven (1978). During these trials the government handed over management of the medical care of select groups of Medicare enrollees to private insurers in exchange for a payment that did not vary with the realized medical expenditures of each individual. To the extent that the rise in cost was driven by principal-agent problems, this mechanism was seen as a way to ensure that providers bore more of the financial risk of medical decisions (Smith et al., 1997). This program was brought to the entire country in 1997 under the name Medicare+Choice.

Medicare+Choice initially struggled to attract plans and nationwide enrollment hovered near 5 million – less than 10% of those eligible. Critics blamed low subsidy rates and the fact that flat payments incentivized firms to cream-skim relatively healthy individuals from the risk pool (Brown et al., 2014). The Medicare Prescription Drug, Improvement, and Modernization Act of 2003 aimed to remove this incentive by risk-adjusting payments. Under the new system, firms submit demographic and diagnostic data about enrollees to the Centers for Medicare and Medicaid Services (CMS) at the time of enrollment. CMS assigns each enrollee a score based on its FFS expenditures on similar individuals in TM; a score of 1.0 indicates average risk. Payments to firms are then adjusted according to these risk scores. Proponents argued that this mechanism would compensate firms for taking on risk without reimbursing specific procedures thus maintaining the profit motive which would (in theory) lead to cost reductions. The program was renamed Medicare Advantage (MA). As of 2022, roughly half of Medicare beneficiaries were enrolled in MA plans.[7]

MA enrollees forgo the TM program and receive covered medical benefits exclusively through their MA plans. MA enrollees pay the Medicare Part B premium and may also pay a plan premium. Insurers compete along the dimensions of benefit design, premiums, and provider networks, and often heavily market their plans (Aizawa and Kim, 2018). Plans generally offer a set of 'in-network' providers which enrollees may utilize with lower cost-

[7] See McGuire et al. (2011) for a comprehensive history of the Medicare Advantage program.

sharing than 'out-of-network' providers. MA plans generally provide a more generous benefit package than TM, such as including dental, vision, and/or hearing coverage (DVH). Many plans include a drug benefit. Plans may also offer a reduction in the Part B premium.

The enrollee-specific subsidy from CMS to insurers is based on a "benchmark" rate for each county, which varies across geographies and over time and is not influenced by MA firms (Newhouse et al., 2012). CMS calculates the benchmark schedule each year using the average risk-adjusted per-capita FFS Medicare spending within the county. Counties are ranked by average spending and placed into quartiles. The benchmark for counties in the top quartile is set to 95% of their FFS spending. The benchmark for the second quartile is 100% of FFS spending, the third quartile benchmark is 107.5% of spending, and the bottom quartile benchmark is 115% of spending. A floor that varies by urban/rural status applies.

Each year, after benchmarks are published by CMS, insurers submit detailed proposals to offer MA plans. These 'bids' include benefit and cost-sharing designs, and detailed information about the insurer's expected revenues and expenses for both TM-covered and non-TM-covered services. Ultimately, the 'bid amount' represents the insurer's offer to provide all services covered by TM to a person of average risk in the plan's coverage area in exchange for a particular level of revenue. The bid amount must be related to the plan's projected costs and may be above or below the benchmark rate. Insurers that bid above the benchmark must charge premiums to enrollees. Firms that bid below the benchmark receive a portion of the difference as a 'rebate' that must be passed on to consumers through decreases in cost-sharing (e.g. reductions in copays) or by offerings of services not covered under TM (e.g. dental). Supplemental benefits may also be paid for by an additional premium. MA plans that offer a prescription drug benefit submit a separate bid which maps in a similar way to a Part D premium.

The rebate payment varies across carriers and over time based on the CMS 'star rating' measure of insurer quality. Payments in our data vary from between 50% and 75% of the difference between the benchmark and the bid. Under current policy, insurers with at least four stars (out of five) also receive a 5% bonus to the benchmark rate. The star rating itself is a summary of multiple measures of past service quality which change throughout our study period, such as the fraction of plan members receiving influenza vaccinations, the

30-day hospital readmittance rate, and enrollee assessments of care quality.

Beneficiaries can enroll in plans during an fixed Open Enrollment period in the fall prior to the plan year. Beneficiaries may also enroll in MA when they become newly Medicare eligible and after certain life events. These rules are designed to reduce adverse selection.[8] After enrollment, firms collect and transmit risk-adjustment information to CMS.

To summarize, the payment from CMS to insurers for an enrollee i living in county m enrolled in plan j in year t based on a benchmark B_{mt} and a plan bid b_{jt} can be calculated with

$$Payment_{ijt} = \begin{cases} B_{mt} \times \phi_{jt} \times Risk_{it} & \text{if } b_{jt} \geq B_{mt}\phi_{jt} \\ (b_{jt} + \lambda_{jt} \times (B_{mt} \times \phi_{jt} - b_{jt})) \times Risk_{it} & \text{if } b_{jt} < B_{mt}\phi_{jt} \end{cases} \tag{3}$$

where ϕ_{jt} captures any bonus to the benchmark rate and λ_{jt} is the rebate percentage. We denote the market-level (i.e. county-year level) benchmark with B_m and denote risk-neutral (i.e. $Risk = 1.0$) plan-specific benchmarks with $B_{jt} \equiv B_{mt} \times \phi_{jt}$.[9]

MA is a significant component of the federal budget. In 2017, payments to plans in our data were \$122 billion and TM spending on the individuals in our data totaled \$317 billion. MA market structure is relatively concentrated. The top five firms nationwide, Aetna, Blue Cross Blue Shield, Humana, Kaiser Permanente, and UnitedHealth Group, account for 65% of total enrollment. The average beneficiary has access to 10 plan options with 64% of beneficiaries having access to 5 or more plans. 25% of beneficiaries in our 2015 data have access to 3 or fewer plans. The average bid is 90% of TM costs (MedPAC, 2017).

Figure 1 illustrates the 2017 policy and the resulting market outcomes with county-level maps of the US.[10] The left map illustrates the ratio of the 2017 benchmark rate to the average FFS spending in 2017. The right map illustrates the total MA share in each county. As consumer surplus is related to the total MA share, these graphs offer a simple assessment of the current government policy. If private costs are tightly linked to the government's costs

[8]Since 2014, enrollees have been allowed to switch to a "5 star" plan at anytime. As only 1% of enrollees switch plans mid-year in our data, we treat any mid-year switchers as choosing the plan in which they spent the most time.

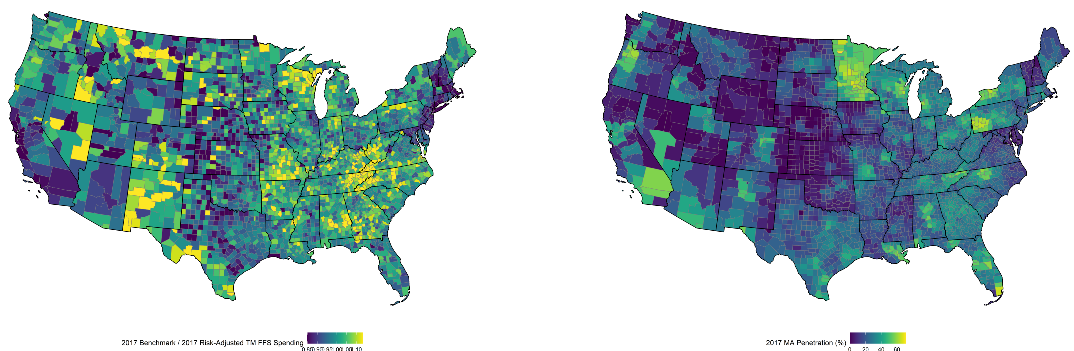
[9]'Regional PPOs'—certain plans offered in one or more entire states—set premiums and benefits as other plans do, but face a slightly different payment system. For computational tractability, we assume Regional PPO plans operate identically to other plans. As Regional PPO plans have a total market share of 1.0% in our data, our results are not likely to be affected by this assumption.

[10]Appendix Figures G.1 presents another view of the benchmark distribution.

and differences in those costs were the only source of heterogeneity across markets, then we would expect those areas which had larger benchmarks relative to FFS spending to have greater enrollment. Instead, we see significant deviations from this pattern. Some areas with high relative benchmarks, such as much of New Mexico, do not have particularly high enrollment, while other areas with high enrollment, such as Minnesota and southwestern Pennsylvania, do not have particularly high relative benchmarks. This suggests that there may be gains by redistributing government funds across counties.

Figure 1: 2017 Medicare Advantage Benchmarks Relative to Traditional Medicare Spending, and Market Penetration, by County

(a) Benchmark / FFS spending (b) MA Penetration



Notes: Map (a) illustrates the ratio of the 2017 benchmark rate to the 2017 risk-adjusted TM (FFS) spending in each county. To show detail, the data are windsorized at the 5th and 95th percentiles. Map (b) illustrates the county-level MA penetration rate in 2017, defined as the total number of people enrolled in any MA plan divided by the number of Medicare beneficiaries in the county. All data from CMS.

4 Data

For our analysis, we combine detailed, county-level, administrative data on plan characteristics and enrollment from CMS with micro-level data on consumer choices and beneficiary characteristics from the Medicare Current Beneficiary Survey (MCBS). We describe these data below.

4.1 Medicare Advantage plans

We collect data on all plans offered from 2008 to 2017 from public CMS files. For each plan, we collect county-level enrollment, premiums, the Part B premium reduction, in-network copayment rates for primary care visits and 7-day hospital stays, the star rating, and indicators for basic and expanded drug coverage (as defined by CMS), and dental, vision, and hearing coverage of any type.[11] We also collect benchmark rates. We do not observe the bids directly. Rather, we observe plan-level risk-adjusted payments which, when combined with the above data and Equation (3), allow us to uniquely identify a bid for each plan-county. We combine the enrollment counts with CMS eligibility data to form product shares at the plan-county-year level. Finally, CMS releases detailed costs estimates submitted by firms during the bid process after a five year delay. We obtain these costs for all plans from 2008-2015. While this cost information is extra-ordinarily detailed, we focus on plans' reported risk-adjusted cost of providing TM-equivalent coverage.

We focus on the market for individual insurance described in Section 3, and drop plans sponsored by employers and plans designed for individuals who are "dual-eligible" for Medicare and Medicaid, as plans in these categories operate under a different payment system and benefit structure. Due to CMS restrictions, we drop plan-county observations with ten or fewer enrollees. For consistency, we drop plans outside our micro-data sample area.

Table 1 presents the mean plan characteristics of our 64,542 plan-county-year observations by benchmark quartiles calculated at the market (county-year) level. In the cross section, as benchmarks increase, observable plan benefits generally improve. The fraction of plans offered with zero premium increases from .336 in the first quartile to .486 in the fourth. However, these patterns are not always monotonic: the average deductible increases from \$63.08 in the first quartile to \$99.55 in the third quartile before decreasing to \$87.76 in the fourth quartile. These patterns reflect the fact that benchmarks are set as a function of average TM costs in previous years. While the costs faced by private insurers are surely correlated with average TM costs, there are likely meaningful cost differences which, when combined with heterogeneous demand responses, implies that the benchmark alone is an

[11] Two percent of plans use coinsurance, which we convert to copayments using the Medicare Physician Fee Schedule and the American Hospital Association Annual Survey.

insufficient statistic for understanding the benefit generosity behavior of firms.[12] In fact, the wedge between the benchmark and mean reported plan TM costs increases in the benchmark suggesting plans' costs differ meaningfully from the government's.

Table 1: Mean plan characteristics by market-level benchmark quartile

	Benchmark quartile			
Variable	**1st**	**2st**	**3rd**	**4th**
Annual premium ($)	636	657	562	501
Enrollment-weighted premium ($)	474	493	444	336
Fraction of plans with zero premium	.336	.361	.419	.486
Annual Part B premium reduction ($)	9.62	14.64	35.42	42.77
Deductible ($)	63.08	91.53	99.55	87.76
Star rating	3.11	3.19	2.45	2.24
Copays				
Primary care ($)	12.90	12.60	14.49	13.99
Specialist visit ($)	37.36	35.36	32.32	28.57
Hospital stay ($)	1,351	1,288	1,118	1,003
Supplemental coverage indicators				
Basic prescription drug	.791	.691	.771	.766
Enhanced prescription drug	.604	.829	.687	.680
Dental	.686	.698	.586	.639
Vision	.929	.919	.902	.891
Hearing	.620	.684	.718	.744
Annual benchmark ($)	9,469	10,121	10,724	11,939
Plan TM cost ($)	9,198	9,682	9,975	10,582
Enrollment	458	813	746	837
Market-level MA share	.206	.220	.236	.239
Observations	10,750	13,331	17,569	22,892

Notes: This table summarizes our MA plan data across quartiles of the base benchmark defined at the market (county-year) level. An observation is a plan-county-year. Reported figures are unweighted means unless noted. All costs are in 2017 dollars. The annual premium as defined here is the supplemental MA premium – all TM and MA enrollees pay the Part B premium. The star rating ranges from zero to five. Prescription drug coverage indicators are additive. "Plan TM cost" is the plans' costs of covering required services as disclosed during the bidding process and is limited to 2008-2015.

[12]There are several potential explanations for lack of a monotonic relationship between the benchmark and benefit provision. For example, a change in the benchmark rate could be a signal of a change in the risk distribution of consumers in the market, which could lead plans to try to cream-skim by changing their product characteristics (Decarolis and Guglielmo, 2017). The menu of product features may not be fully salient to consumers (Curto et al., 2021), thus increasing benefits may increase plan costs but yield small increases in enrollment. Our model can account for all of these possibilities (at least to some degree) with the exception of adverse selection. However, as mentioned above, the evidence suggests that the current implementation of the risk-adjustment system effectively reduces incentives for plans to cream-skim (Newhouse et al., 2015).

4.2 Medicare beneficiaries

We access data on individual Medicare beneficiaries from the 2008-2017 Medicare Current Beneficiary Survey (MCBS), a rolling-panel survey produced by CMS and Westat. Participants are interviewed repeatedly over three years, and responses are linked to CMS data to ensure accuracy. We observe demographics including income, age, sex, race, education, and county of residence. Respondents self-report their health status, choosing from Excellent, Very Good, Good, Fair, and Poor. We also observe MA plan choices. In some years, the MCBS does not report the plan choice directly and instead reports the insurer choice, along with information about plan premiums and features which we match to plan data.

The MCBS samples Medicare beneficiaries using a multi-level clustered procedure. While we do not observe beneficiaries in every county, within each geography included in the data there is considerable variation in demographics and plan enrollment. The MCBS provides sampling weights which we use to transform our results into a nationally representative form.[13]

The typical set of Medicare beneficiaries studied in the literature includes age 65-plus retirees without outside insurance (Curto et al., 2021). However, the MCBS and CMS data include others who are eligible to purchase MA plans including those with employer-provided insurance, those whose original Medicare eligibility was not age-related, those with ESRD, and those who are not full-year Part A/B enrollees. As these individuals purchase MA plans, we cannot exclude them without violating our assumption that the MCBS draws from CMS enrollment files. We instead create 'administrative' indicator variables. We exclude any individuals who were also eligible for Medicaid during the year and those with missing address information. After applying these exclusions, the sum of the MCBS sample weights differs from the total MA-eligible population in the CMS data by less than 2%.[14]

Medicare beneficiaries also have access to non-MA insurance options, and variation in the price of those options may make MA plans more or less attractive. We focus on Medicare supplemental insurance (a.k.a. Medigap) which pays for medical costs not covered by TM.

[13] In 2017 CMS updated the weighting methodology to ensure that the MCBS matched average MA enrollment. For consistency across years we reweight the pre-2017 data to match current methodology.

[14] According to our CMS data, in 2017 the total number of Medicare beneficiaries not also eligible for Medicaid was 42.7 million. The total MCBS weight for 2017 is 43.4 million.

For example, TM covers 80% of the cost of physician visits, and a Medigap plan may pay for the rest. Medigap plan designs are standardized by CMS and indexed by letters. For each person, we obtain the rate for Medigap Plan C offered by United Healthcare that year from Weiss Ratings. Plan C covers most of the coinsurance and deductibles that TM enrollees are responsible for and is the most popular Medigap plan.[15]

Summary statistics on our 78,812 individual-year observations covering 3,851 county-year markets and 42,261 unique individuals are reported in Table 2. The mean age of individuals in our data is 73. Slightly more than half of our observations are of females. Over 90% of individuals are coded by CMS as White. Over 75% self-report "Good" or better health. 25% report having college degrees and 16% did not graduate high school. 25% receive employer-sponsored insurance, and 14% are Medicare-eligible for non age-related reasons. The second set of columns splits the data by MA enrollment. On average, MA enrollees have lower income, are less likely to be White, and have lower educational attainment.

The third set of columns of Table 2 illustrates the panel nature of our data and focuses on panel observations for which the individual was enrolled in TM in the previous year – 27,297 observations total. We split the data into those who switched from TM to MA and those who remained in TM. Those who switched are generally similar to the larger group of MA enrollees, though switchers are slightly healthier on average.

Finally, we supplement these data with market-level average demographics from the Area Health Resource File published by the Health Resources and Services Administration. For each market, we collect the median household income, the percent of those 65-and-older in deep poverty, the unemployment rate, the population density, and the number of doctors, Medicare-certified hospitals, skilled nursing facilities, and hospice facilities. Summary statistics for our markets in 2017 by benchmark quartile are reported in Table G.3.

[15] Massachusetts, Minnesota, and Wisconsin have alternative plan definitions; in those states we use the rate for the plan closest to Plan C. Additionally, United Healthcare did not offer plans in New York during our study period. For individuals in New York, we averaged the Plan C rates offered by all other insurers.

Table 2: Medicare beneficiary micro-data summary statistics

Variable	All observations		By MA enrollment		TM → MA switch	
	Mean	Std. dev.	MA	TM	Yes	No
MA enrollment indicator	.280	.449	1	0	1	0
Income ($)	52,684	76,202	43,117	56,415	43,858	53,172
Age	73.3	9.84	73.5	73.2	72.8	74.8
Medigap price ($)	2,722	674	2,810	2,687	2,654	2,742
Demographic indicators						
Female	.536	.499	.548	.531	.526	.536
Black	.081	.273	.095	.075	.094	.069
Hispanic	.010	.101	.017	.008	.018	.006
Education indicators						
Bachelor's degree or higher	.250	.433	.188	.275	.216	.260
Attended college	.307	.461	.315	.304	.294	.296
Graduated high school	.285	.451	.305	.277	.299	.288
Health status indicators						
Excellent	.177	.381	.175	.177	.174	.165
Very Good	.313	.464	.315	.312	.319	.313
Good	.304	.460	.307	.303	.319	.314
Fair	.151	.358	.155	.149	.149	.153
Poor	.056	.230	.049	.059	.049	.055
Administrative indicators						
Employer-provided insurance	.254	.435	.008	.350	.000	.352
Non-aged eligibility	.144	.351	.147	.143	.193	.141
ESRD	.007	.081	.004	.008	.002	.007
Full-year Part A/B enrollee	.905	.293	.977	.877	.962	.901
Observations	78,812		22,108	56,704	1,345	25,958

Notes: This table summarizes our data on Medicare beneficiaries from the Medicare Current Beneficiary Survey (MCBS). An observation is a person-year. The Medigap price is the United Healthcare premium for Medigap Plan C (see text for details). Demographic categories are defined by CMS administrative data. Education indicators are mutually exclusive. The first set of two columns reports means and standard deviations for all observations in the microdata. The third and fourth columns split the observations into those enrolled in MA and those enrolled in TM. The last two columns split the observations by switching behavior conditional on observing past-year TM enrollment. Income and Medigap price are in 2017 dollars. All figures reported here are weighted according to MCBS sampling weights.

5 Demand

We model the demand for MA plans by extending the discrete choice demand setting of Goolsbee and Petrin (2004) allowing for switching costs between Traditional Medicare to Medicare Advantage and for the possibility of switching costs between MA plans. Agents consist of consumers (Medicare beneficiaries) i and insurers/firms f which operate in markets (counties) m. Periods (years) are denoted by t. Consumers are described by a vector of demographic characteristics observable to the econometrician z_{it} and unobservable characteristics $\nu_{it}, \{\epsilon_{ijt}\}$. Each insurer offers plans $j \in \mathcal{J}_{ft}$. Plans are unique to individual markets and are described by a premium p_{jt} and a vector of characteristics x_{jt} that includes supplemental benefits and cost-sharing rules, and a characteristic ξ_{jt} which is observed by consumers but not by the econometrician. Insurers have time- and market-plan-invariant vertical quality v_f. Therefore, ξ_{jt} represents plan-market- and time-specific deviations from that quality, including plan-specific benefits and network breadth.

Let y_{it} be Medicare beneficiary i's income and h_{it} be a vector of indicators corresponding to i's health status. Consumers enter period t enrolled in plan k_{it}. We define three switching cost indicators W_{nijt}. Let W_{1ijt} equal one if k_{it} is the outside good – we call this the *Medicare-to-MA* indicator. Let W_{2ijt} – the *MA Interfirm* indicator – be one if k_{it} is offered by a different firm than j. Finally, let W_{3ijt} – the *MA Intrafirm* indicator – be one if k_{it} and j are different plans offered by the same insurer.

Let u_{ijmt} denote the consumer's utility from enrolling in plan j. Dropping the market subscripts, the choice specific utility for MA plans is given by:

$$\begin{aligned} u_{ijt} = \left(\alpha_0 + \alpha_1 y_{it} + \alpha_2 y_{it}^2\right) p_{jt} + \sum_{n=1}^{3} \beta_{wn} W_{nijt} + \sum_{n=1}^{3} \sum_{h} \beta_{whn} W_{nijt} h_{it} \\ + \beta_z z_{it} + \beta_x x_{jt} + \xi_{jt} + v_f + \beta_\nu \nu_{it} + \epsilon_{ijt}. \end{aligned} \tag{4}$$

The α parameters capture income-varying premium sensitivity. β_{wn} and β_{whn} capture health-dependent switching costs. β_z captures heterogeneous tastes for MA plans by demographics, and β_x captures mean tastes for plan characteristics x_{jt}.[16] ν_{it} is an unobservable (to the

[16]In contrast to this specification, Curto et al. (2021) assume that the effect of plan premiums and characteristics on utility is completely captured by the plan's bid relative to the benchmark.

econometrician) preference that consumer i has for MA which is assumed to be drawn independently from a standard normal distribution—β_ν determines the variance of this random coefficient. We have explored specifications in which ν_i is fixed over time and found similar results. ϵ_{ijt} represents the idiosyncratic taste of consumer i for plan j which is assumed to be drawn independently from the Type-I extreme value distribution.

Consumers have access to an outside good, the price of which may vary with demographics $p_{0t}(z_i)$. The utility of the outside good is

$$u_{i0t} = (\beta_{o0} + \beta_{o1} y_{it} + \beta_{o2} y_{it}^2) p_{0t}(z_{it}) + \epsilon_{i0t}. \tag{5}$$

We normalize by subtracting Equation (5) from each u_{ijt}.

We include switching costs due to the consistent finding of inertia in plan enrollment (Nosal, 2012, Aizawa and Kim, 2018).[17] Enrollees in MA face restrictive provider networks that vary across plans. In addition, Medicare beneficiaries are automatically re-enrolled in their previous plan if they take no action during their open enrollment period—it is virtually costless to re-enroll. Similar to Handel (2013), we model these costs directly in utility.

Following Berry et al. (1995), it is useful to rewrite u_{ijt} into a product-level mean

$$\delta_{jt} = \alpha_0 p_{jt} + \beta_x x_{jt} + v_f + \xi_{jt} \tag{6}$$

and an individual-specific deviation from that mean

$$\begin{aligned} \mu'_{ijt} = \left(\alpha_0 + \alpha_1 y_{it} + \alpha_2 y_{it}^2\right) p_{jt} + \sum_{n=1}^{3} \beta_{wn} W_{nijt} + \sum_{n=1}^{3} \sum_{h} \beta_{whn} W_{nijt} h_{it} + \beta_z z_{it} + \beta_\nu \nu_{it} \\ - (\beta_{o0} + \beta_{o1} y_{it} + \beta_{o2} y_{it}^2) p_o(z_{it}) + \epsilon_{ijt}. \end{aligned} \tag{7}$$

Let $\mu_{ijt} = \mu'_{ijt} - \epsilon_{ijt}$. Given our distributional assumptions on ϵ_{ijt} and ν_{it}, the probability

[17] Like Aizawa and Kim (2018), we do not model consumers as forward-looking for several reasons. First, such analysis is computationally intensive. Second, it likely requires assuming that individuals choose according to a model of neoclassical preferences with a discount factor close to one. However, recent work has shown that in related settings that model does not explain Medicare beneficiary behavior well (e.g. Dalton et al., 2018). Third, our estimation approach captures the inertia that is salient for our counterfactual analysis.

that consumer i chooses plan j (i.e. the share function) is

$$s_{ijt} \equiv \Pr(i \text{ chooses } j) = \int_{\nu} \frac{\exp(\delta_{jt} + \mu_{ijt}(\nu))}{1 + \sum_{k \in \mathcal{J}_m} \exp(\delta_{kt} + \mu_{ikt}(\nu))} d\nu, \tag{8}$$

and the total share of plan j is

$$s_{jt} = \int_{z_{it}} s_{ijt}(z_{it}) dz_{it}. \tag{9}$$

We define consumer welfare in terms of compensating variation: the amount that, if the choice of MA plans was removed, consumer i would have to receive as income in order to achieve the same level of expected utility before idiosyncratic shocks are realized (Hicks, 1945, Diamond and McFadden, 1974, Nevo, 2000). Let $\alpha_{it} = \alpha_0 + \alpha_1 y_{it} + \alpha_2 y_{it}^2$. The expected consumer welfare for beneficiary i is

$$CS_{it} = E[\max_j u_{ijt}]/\alpha_{it} = \frac{1}{\alpha_{it}} \ln \left(1 + \sum_j \exp(\delta_{jt} + \mu_{ijt}) \right). \tag{10}$$

This is the definition of consumer welfare we employ in the optimal subsidy schedule problem. However, we note that while the government seeks to maximize the sum of this compensating variation across all markets (as it does not observe ϵ_{ijt}), consumers ultimately only accrue welfare from MA if they enroll in an MA plan. We therefore calculate mean compensating variation *for MA enrollees* via

$$\overline{CS}_t^{cond} = \frac{\int_i CS_{it}}{\int_i s_{it}}, \tag{11}$$

where s_{it} is the probability that consumer i chooses any MA plan in period t. Following the literature (see, e.g. Petrin, 2002, Town and Liu, 2003), we report the mean compensating variation both per Medicare beneficiary and per MA enrollee, as well as the aggregate consumer welfare $CS_t = \int_i CS_{it}$. This formulation of consumer surplus assumes that our parameterization of demand holds for inter- and infra-marginal consumers (McFadden, 1974). Though we do not observe switches for every consumer, we do observe switches by consumers in each demographic category. As μ_{ij} includes switching costs, our estimates of consumer surplus are net of those costs and in this sense are short run.

6 Supply

Our model of insurer behavior largely follows the typical multi-product firm approach in the literature (see, e.g. Berry et al., 1995, Petrin, 2002), albeit with two notable differences. First, firms choose prices and product characteristics simultaneously. Second, as detailed in Section 3, firms submit a 'bid' b_{jt} to CMS for each plan they offer, which maps into revenue from the government through subsidies and (potentially) from consumers through premiums as a function of the plan's characteristics. In Appendix B we show that under certain assumptions the CMS rules imply that the mapping is unique and thus we can write the firm's problem in the traditional way in terms of prices and product characteristics with the addition of a subsidy that depends on characteristics.

Let x_{jt} and ξ_{jt} be the product characteristics as defined above, and let p_{-jt}, x_{-jt} and ξ_{-jt} be the set of prices and product characteristics for all plans other than j. Let $c_{jt}(x, \xi)$ be the per-enrollee marginal cost incurred by the firm. Let $sub_{jt} = sub(x_{jt}; B_{jt}, \lambda_{ft})$ be the function that maps product characteristics and the benchmark B_{jt} $(= B_{mt}\phi_{ft})$ to the subsidy received by the firm where λ_{ft} is the firm's rebate percentage and ϕ_{ft} is the firm's benchmark bonus, taken to be exogenous.[18] Let N_m be the number of Medicare beneficiaries in market m. Plan-level profit is

$$\begin{aligned}\pi_{jt}(p_{jt}, x_{jt}, \xi_{jt}; p_{-jt}, x_{-jt}, \xi_{-jt}) = (p_{jt} + sub(x_{jt}; B_{jt}, \lambda_{ft}) - c_{jt}(x_{jt}, \xi_{jt}))N_m \times \\ s_{jt}(p_{jt}, x_{jt}, \xi_{jt}; p_{-jt}, x_{-jt}, \xi_{-jt}) \quad (12)\end{aligned}$$

and the firm's profit is

$$\max_{\{p_{jt}, x_{jt}, \xi_{jt}\}_{j \in \mathcal{J}_f}} \sum_{j \in \mathcal{J}_f} \pi_{jt}(p_{jt}, x_{jt}, \xi_{jt}; \cdot). \quad (13)$$

These equations map into the primitives of the counterfactual framework introduced in Section 2. The agent's action X_f is the vector $\{p_{jt}, x_{jt}, \xi_{jt}\}_{j \in \mathcal{J}_f}$. The agent's payoff function π_f is the profit function $\sum_{j \in \mathcal{J}_f} \pi_{jt}(\cdot)$. The agent's information set $\mathcal{I}_f$ consists of

[18] CMS uses past values of performance measures (two years before the plan year) to calculate the star rating (and thus λ and ϕ), and changes the characteristics used from year to year. Insurers therefore likely find it difficult to manipulate specific characteristics to obtain higher rebate percentages. We thank an anonymous referee for clarifying this point.

the parameters necessary to compute sub_{jt}, c_{jt}, and s_{jt}, including information describing the distribution of consumers in the market. Before turning to the cost model, we note that we do not model plan entry and exit due to computational constraints.[19]

6.1 Costs

We assume the marginal cost of offering a plan is constant in the number of enrollees. We further assume that risk-adjustment is effective so that c_{jt} does not vary with enrollees' health status. We model the logarithm of the marginal cost function as

$$\ln(c_{jt}) = \gamma_f + \gamma_r + \gamma_m + \gamma_t + \gamma_{x,jt} x_{jt} + \gamma_{\xi,jt}\xi + \omega_{jt}, \tag{14}$$

where γ_f is a firm-specific cost, γ_r is a star-rating-specific cost, γ_m is a market-specific cost, and γ_t is a time-varying cost. These parameters are fixed effects to be estimated. $\{\gamma_{x,jt}, \gamma_{\xi,jt}\}$ are the plan-varying costs of providing x and ξ, respectively. ω_{jt} is an unobservable (to the econometrician) plan-level cost.[20] We note that since this cost function, the share function, and the subsidy function derived in Appendix B are smooth, the firm's profit function is smooth as well and therefore Assumption SP of Proposition 1 is satisfied.

The solution to Equation (13) is partially characterized by the first-order conditions

$$\{p_{jt}\}: \quad s_{jt} + \sum_{k \in J_f} (p_{kt} + sub(x_{kt}; \cdot) - c_{kt}) \frac{\partial s_{kt}}{\partial p_{jt}} = 0 \tag{15}$$

Following Berry et al. (1995), we define a $J \times J$ matrix Δ_t whose (j, k) elements are given

[19] A natural question in response to this assumption is: What do we lose by holding the set of firms and plans constant? In our data, the total share of plans which enter and exit is less than 1%. As our focus is on consumer welfare, and in this model of demand consumer welfare is approximately proportional to share—see Equation (10)—we conclude that entry and exit do not generate large changes to consumer welfare over the range of benchmarks we see in the data. This issue would still be concerning if the optimal benchmark schedule involved a large number of benchmarks outside the range of benchmarks we see in the data. Fortunately, it does not (see Figure 3). We therefore conclude that entry and exit is not a first order concern in our context.

[20] In contrast, Curto et al. (2021) assume that on the margin plan costs have a one-to-one relationship with plan bids.

by

$$\Delta_{jkt} = \begin{cases} -\frac{\partial s_{kt}}{\partial p_{jt}}, & \text{if } \{j,k\} \in \mathcal{J}_f \\ 0, & \text{otherwise.} \end{cases}$$

The first-order conditions (15) can then be solved in vector form to obtain cost levels:

$$\mathbf{c}_t = \mathbf{p}_t + \mathbf{sub}_t - \Delta_t^{-1} * \mathbf{s}_t. \tag{16}$$

Plan-characteristic level costs $\gamma_{x,jt}$ can be recovered from the first-order conditions with respect to product characteristics.

$$\begin{aligned} \{x_{jt}\}: \quad & \left(\frac{\partial sub_{jt}}{\partial x_{jt}} - \frac{\partial c_{jt}}{\partial x_{jt}}\right) s_{jt} + \sum_{k \in J_f} (p_{kt} + sub(x_{kt}; \cdot) - c_{kt}) \frac{\partial s_{kt}}{\partial p_{jt}} = 0 \\ \Rightarrow \gamma_{x,jt} = & \frac{\partial sub_{jt}}{\partial x_{jt}} c_{jt}^{-1} + \sum_{k \in J_f} (p_{kt} + sub(x_{jt}; \cdot) - c_{kt}) \frac{\partial s_{kt}}{\partial x_{jt}} (s_{jt} c_{jt})^{-1}, \end{aligned} \tag{17}$$

where the second line uses $\frac{\partial c_{jt}}{\partial x_{jt}} = \gamma_{x,jt} c_{jt}$. Similar logic applies to the recovery of $\gamma_{\xi,jt}$.

6.2 Government spending

Evaluating candidate solutions to the optimal subsidy schedule problem requires calculating the total government expenditure on the Medicare program, which consists of the sum of the MA payments given by $sub(\cdot)$ and spending on TM. Let TM_{mt} be the average risk-adjusted TM spending in the market. As we do not observe individual-specific risk scores, we calculate MA and TM spending using the average risk level in the market. That is, we set $Risk_{it} = Risk_{mt}$ for all i when calculating payments to firms. Thus, dropping time subscripts,

$$GovExp_i(B_m) = \sum_j s_{ij} sub_j(B_m; \cdot) Risk_m + \left(1 - \sum_j s_{ij}\right) TM_m Risk_m, \tag{18}$$

where TM_m is the average per-enrollee TM spending in the market. In other words, as consumers switch in to or out of MA, $GovExp$ includes their costs across both programs. We treat TM_m as exogenous due to the risk adjustment system—i.e. we assume MA enrollment does not change within-county risk-adjusted TM spending, though we test this assumption after estimating the parameters of firms' marginal cost function. However, as benchmarks change and beneficiaries move between TM and MA in a county, the across-county average TM and MA risk scores and average costs change as well.

7 Estimation

We estimate the parameters of the demand system following the two-stage approach of Goolsbee and Petrin (2004). First, we estimate parameters which capture individual-level variation in MA preferences—those parameters that define μ'_{ij}— via maximum likelihood and recover mean plan-level utilities δ_j. We then estimate the parameters of Equation (6) with an instrumental variables approach.

Let $\theta_I = \{\alpha_1, \alpha_2, \beta_w, \beta_{wh}, \beta_z, \beta_\nu, \beta_0\}$ be the set of parameters which determine μ'_{ij}. For a candidate value $\tilde{\theta}_I$ we use the Berry (1994) inversion with the Berry et al. (1995) contraction mapping to compute the unique set of product mean utilities $\delta_j(\tilde{\theta}_I)$ that match predicted shares to the aggregate county-level market shares observed in the CMS data. Let C_{ij} be an indicator variable that is equal to one if person i chose product j. The likelihood function is

$$L_{it}(C_{ijt}; \tilde{\theta}_I, \delta(\tilde{\theta})_I) = \prod_j s_{ijt}^{C_{ijt}}, \tag{19}$$

where s_{ijt} is given by Equation (8).[21] In the first stage of our estimation procedure, we apply the MCBS sample weights w_{it} and maximize the weighted log likelihood function

$$l(C; \tilde{\theta}) = \sum_i ln(L_{it})w_{it}. \tag{20}$$

At the estimate $\hat{\theta}_I$ we store the unique $\hat{\delta}_j(\hat{\theta}_I)$ and regress it on observable product characteristics to estimate the parameters of Equation (6).

We estimate policy functions with a first-order approximation. That is, for product j

[21] In practice, we form an empirical analog of the share function by numerically integrating over draws of the ν_{it} distribution. Furthermore, due to the rolling panel design of the MCBS a fraction of our observations have no past enrollment data with which to form the W_{nijt} variables. For these observations we draw from the distribution implied by shares of the plans offered in the previous period.

and characteristic x_l, we write

$$x_{ljt} = \beta_{f,l} \times B_{jt} + \beta_{f,z} \times \tilde{I}_{ft} + \epsilon_{ljt}, \tag{21}$$

where $\beta_{f,l}$ is the firm-level first-order approximation of the effect of the change in the benchmark, $\tilde{I}_{ft}$ is an approximation of the information set $\mathcal{I}_{ft}$, and ϵ_{ljt} captures measurement error, approximation error, and other factors that influence product characteristics such as plan-product-characteristic-level cost and demand expectation shocks.

7.1 Instruments

Since ξ_{jt} is chosen by firms and observed by consumers but not observed by us, it is likely to be correlated with the plan premium and other product characteristics. To identify the α_0 and β_x coefficients, we must therefore find instruments for premiums and plan characteristics. First, we note that not all product characteristics are likely to be endogenous. For example, for each plan, the basic drug coverage indicator remains constant over time and so it is plausibly exogenous. Furthermore, the star rating is set at the insurer level reflecting health outcomes with a two year lag and therefore is also plausibly exogenous.

We construct one instrument from the observation that our cost function includes a geographic component; costs are therefore correlated across plans in a given market. Our data includes detailed information on insurers' cost projections for TM-covered services submitted during the bidding process. These projections must be a) related to the plan's past realized costs, and b) certified by a professional actuary. For each plan, we calculate the average total cost of TM-covered services across competitors weighted by their conditional shares. This instrument is excluded from the demand system if competitors' TM-covered service costs c^{TM}_{-jt} are uncorrelated with ξ_{jt} i.e. $E[c^{TM}_{-jt}\xi_{jt}] = 0$. Since these costs are private information—these data are not released by CMS until five years after the plan year has concluded—it is not likely that firms choose ξ_{jt} based on the costs of particular competitors.

The panel nature of our data points to additional potential instruments. First, we note that observable and unobservable plan characteristics as well as premiums are likely functions of the benchmark and hence correlated with each other thereby invalidating BLP-type instruments in our setting. However, if the costs of plan characteristic provision are correlated over time and if county benchmark updates are independent, using lagged values of the residuals from a regression of plan characteristics on benchmarks and other time-invariant state variables should be valid instruments. Intuitively, the residuals proxy for plan characteristic costs as the common impact of the benchmark will have been removed from the insurer's choice of plan characteristics. If updates to the benchmark are independent and shocks to ξ are uncorrelated with benefit provision costs, then ξ will be orthogonal to these lagged plan characteristic residuals.[22]

We examine the validity of these instruments first by testing the independence of benchmark updates. We estimate that a \$1 increase in $(B_{mt-1} - B_{mt-2})$ is associated with a \$0.03 decrease in $(B_{mt} - B_{mt-1})$ (t-statistic = 0.17 when clustering by year) and conclude benchmark updates are approximately independent. Next, we examine the assumption of persistence of the plan characteristic residuals. We find that the correlation coefficient between the contemporaneous and lagged residuals ranges from 0.6659 (enhanced drug coverage) to 0.8754 (hospital copay). First-stage F-statistics testing the explanatory strength of instruments in accounting for plan characteristic variation range from 298 (annual premium) to 3,396 (hospital copay). Taken together, the evidence suggests that the necessary conditions for our lagged residual instrumenting strategy seem to hold.[23]

While these results suggest that benchmarks updates are exogenous from ξ, they nonetheless may be correlated with plan-product-characteristic-level *costs* of insurers making them endogenous in the policy functions. Therefore, we need instruments to consistently estimate Equation (21) and satisfy Assumption CPE. We take advantage of the difference in the payment floors coming from county-level differences in urban/rural status and leverage the identification strategy of Duggan et al. (2016). These benchmark differences are driven by small population differences across counties that map into CMS's definition of urban and

[22]More formally, for characteristic x_l and time-invariant 'state' variables a_j we can write $x_{ljt} = g_{f,l}(B_{mt}, a_j) + u_{l,jt}$ and $\xi_{jt} = g_{f,\xi}(B_{mt}, a_j) + u_{\xi,jt}$. Here, the u terms are mean-zero random variables. Define $o_{mt} \equiv B_{mt} - B_{mt-1}$. Suppose o_{mt} is a random variable distributed independently across time and with respect to u. Further suppose $E[u_{l,jt-1} u_{\xi,jt} | B_{mt}, a_j] = 0$; time-varying information relevant to the choice of x_l in period $t-1$ is not relevant to the choice of ξ in period t after conditioning on the benchmark and a_j. Under these assumptions, $E[u_{l,jt-1}\xi_{jt}] = 0$. We thus instrument for x_{ljt} with $\widehat{u_{l,jt-1}}$.

[23]This is similar to our policy function regression. The key difference is that the $\tilde{Z}_f$ of Equation (21) may include time-varying components such as demographics and competitors' past choices.

rural that are very likely orthogonal to plan characteristic costs. We obtain the Rural-Urban Continuum Code from the Area Health Resources File and instrument benchmarks with rural-urban category identifiers interacted with year fixed effects.[24] We take Assumption PRS as given, though we have tested for multiple equilibria by re-estimating our policy functions by Census Region and by benchmark quartile.

8 Results

In this section we first describe our estimates of consumer preferences and MA plan costs function parameters. We then discuss of our estimates of the firm-varying product characteristic policy functions.

8.1 Demand

We report maximum likelihood estimates of individual-specific parameters in Table 3. The estimates imply that higher income consumers are less price-sensitive than lower income consumers. The highest switching costs are incurred by consumers switching from TM to MA. Inter-firm switches are less costly and intra-firm switches are cheaper still. These results suggest that the primary component of switching costs is the disutility of changing providers. We interact the switching costs with indicators for self-reported health status, with 'Poor' as the excluded group. The point estimates indicate that healthier individuals face lower costs of switching, consistent with the provider-changing hypothesis, though the size of the standard errors prevent us from making clear inferences between adjacent health statuses.

Our demographic estimates imply that younger beneficiaries, non-Whites, and those with less education have stronger preferences for MA plans. These results align with other findings that MA enrollment of Black and Hispanic beneficiaries has grown faster than enrollment of White beneficiaries (Meyers et al., 2021). Our administrative indicators enter with appropriate signs and reasonable magnitudes. For example, those with employer-provided insurance or ESRD are extremely unlikely to choose an MA plan. Finally, our MA sector random

[24] We thank anonymous referees for suggesting this strategy.

coefficient enters significantly with a magnitude roughly equal to the inter-insurer switching cost suggesting that idiosyncratic preferences for MA are important.

Table 4 reports estimates of the mean utility parameters. The first column presents OLS estimates assuming prices and characteristics are exogenous. The second column reports IV estimates when the premium is instrumented with our cost instrument. The third column reports the results when prices and product characteristics are both treated as endogenous and instrumented with our full set of instruments. Consistent with OLS estimates on price being biased towards zero, the IV premium coefficients are larger in magnitude than the OLS coefficient. Furthermore, in general the coefficients on product characteristics are larger in magnitude when they are treated as endogenous, though the estimates are noisier.

We focus on specification (3). The parameter estimates in this specification are quite sensible. For the plans with a positive premium the average plan elasticity is -6.29. The average semi-elasticity of increasing premiums by $1 is -.073, similar to estimates from the literature. For example, using an earlier sample period, Aizawa and Kim (2018) estimate an average MA semi-elasticity of -.075. Combining these estimates with Table 3, the median consumer is willing to pay roughly $425 for prescription drug coverage, $300 for hearing coverage, and $190 for a reduction of $1,000 in the copay for a hospital stay. The cost incurred by an median-income individual switching from TM to MA is approximately $650, similar to the mean annual premium in our data, while the same individual switching between plans within an MA insurer incurs a cost of only $140.

8.2 Supply

Table 5 reports the implied estimates of marginal costs as well as the contribution of plan characteristics to the logarithm of marginal cost, Equation (14). We report the means and standard deviations of these estimated characteristic-level costs and the overall plan-level cost for the five largest firms nationally and group smaller firms into a sixth category.[25]

The implied marginal cost estimates appear quite reasonable. On average, we estimate

[25] Appendix Table G.1 reports estimates of Equation (14) using an approach closely mirroring Berry et al. (1995): inverting only the pricing first-order condition to recover marginal costs, then regressing those costs on product characteristics. The results are broadly similar to what we report here.

Table 3: Maximum likelihood estimates of individual-specific preferences

Variable	Coeff.	Std. Err.	Variable	Coeff.	Std. Err.
Price × Income	0.1766	0.0321	MA × Demographics		
Price × Income^2	-0.0028	0.0014	Excellent health	0.5421	0.2050
			Very good health	0.5849	0.1923
TM-to-MA switch ×			Good health	0.4503	0.1908
Constant	-4.1212	0.2132	Fair health	0.4346	0.2027
Excellent health	-0.0797	0.2184	Age	2.0606	0.1810
Very good health	-0.2134	0.2053	Age^2	-0.1442	0.0125
Good health	-0.2131	0.2052	Female indicator	-0.1123	0.0348
Fair health	-0.2577	0.2191	Black indicator	0.4814	0.0690
			Hispanic indicator	0.4215	0.1595
Inter-Insurer switch ×			Graduated high school	-0.1661	0.0562
Constant	-1.9003	0.1061	Attended college	-0.3464	0.0575
Excellent health	-0.0861	0.1215	College degree or higher	-0.8418	0.0669
Very good health	-0.0637	0.1148			
Good health	-0.1138	0.1143	Administrative indicators		
Fair health	-0.1353	0.1234	Employer-provided insurance	-5.8128	0.1792
			Non-aged eligibility	0.4633	0.0663
Intra-Insurer switch ×			ESRD diagnosis	-1.7561	0.2158
Constant	-0.8171	0.1334	Full year enrollment	2.8909	0.1090
Excellent health	-0.1770	0.1530			
Very good health	-0.1426	0.1449	Outside good (Medigap) price ×		
Good health	-0.1224	0.1446	Linear	0.3834	0.0581
Fair health	-0.1348	0.1527	Income	-0.1986	0.0135
			Income^2	0.0040	0.0005
Random Coefficient	1.6243	0.1266			
Weighted Log Likelihood			-73,967		
Observations			78,812		

Notes: The coefficients reported here are maximum-likelihood estimates of the parameters of Equation (7), the individual-specific components of utility. An observation is an individual-year. MA and outside good prices are measured in thousands of 2017 dollars. Income is measured in hundreds of thousands of 2017 dollars. The omitted group for the switching cost interactions is 'Poor' health. Age is measured in decades. Educational indicators are mutually exclusive. The Medigap price is the price of Plan C, see text for details. Standard errors are robust to heteroskedasticity.

Table 4: Estimates of mean preferences for plan characteristics

Variable	(1)	(2)	(3)
Annual Premium (per $1000)	-0.4005	-6.9249	-6.7368
	(0.0203)	(0.6562)	(1.0318)
Part B reduction (per $1000)	0.1368	0.6617	1.8192
	(0.0693)	(0.1020)	(0.2986)
Deductible (per $1000)	-0.3062	-1.3622	-1.8048
	(0.0386)	(0.1326)	(0.2860)
Copays			
Primary care visit	-0.0270	-0.0074	-0.0499
	(0.0018)	(0.0037)	(0.0096)
Specialist visit	0.0140	-0.0404	-0.0438
	(0.0016)	(0.0062)	(0.0130)
Hospital stay copay (per $1000)	0.0068	-0.8844	-1.2624
	(0.0280)	(0.1045)	(0.2440)
Supplemental coverage indicators			
Prescription drug	0.5928	3.6512	2.8598
	(0.0421)	(0.3155)	(0.3956)
Enhanced prescription drug	0.3558	0.1912	1.8271
	(0.0384)	(0.0740)	(0.2543)
Dental	0.0684	0.9969	1.3355
	(0.0313)	(0.1064)	(0.2608)
Vision	-0.0041	0.2470	0.6098
	(0.0450)	(0.0812)	(0.2799)
Hearing	-0.0290	1.1064	2.0865
	(0.0385)	(0.1323)	(0.3605)
CMS-assigned star rating	0.1498	0.2535	0.3355
	(0.0111)	(0.0230)	(0.0496)
Firm fixed effects?	Yes	Yes	Yes
Endogenous annual premium?	No	Yes	Yes
Endogenous product characteristics?	No	No	Yes
Observations	50,439	50,439	36,077
Mean implied elasticity (if < 0)	-0.3077	-6.4684	-6.2908
	(0.2100)	(4.4004)	(4.2793)
Mean ds_j/dp_j	-0.0035	-0.0745	-0.0725
	(0.0061)	(0.1257)	(0.1222)

Notes: The coefficients reported here are estimates of the components of the utility function which are common across individuals i.e. the parameters of Equation (6). An observation is a plan-county-year. Estimates in Column (1) are formed via OLS. In Column (2) we instrument for the annual premium using the share-weighted average of competitors' plans projected costs of TM-covered services. In Column (3) we consider all product characteristics as endogenous and add lagged residuals from regressions of the characteristics on the benchmark as instruments. See text for details. All dollar values are in 2017 dollars. Parentheses indicate robust standard errors in the top panel and standard deviations in the bottom panel.

MA costs to be \$868 per enrollee-month, 4.2% less than TM costs of \$906 per enrollee-month. After adjusting for inflation, the comparable average MA cost reported by Curto et al. (2021) is \$830 per enrollee-month. We estimate an average economic profit margin (profit divided by revenue) of 2.1% in 2017. Consistent with this estimate, MedPAC reports an average 2017 MA insurer accounting profit margin of 2.7% (MedPAC, 2019).

CMS reports the actual risk-adjusted mean per-capita TM expenditures in each market, which, if our estimation approach is consistent, are likely to be strongly correlated with our estimated marginal costs. We mimic the spirit of an exercise in Curto et al. (2021) and compare the share-weighted estimated MA cost for zero-premium plans to risk-adjusted county-level TM costs. The two cost measures are positively correlated with a coefficient of .703.

Our estimates of the impact of benefit provision on marginal cost also appear sensible. For example, our estimates imply a \$1 increase in the primary care copay decreases marginal costs by an average of \$5.12, implying that the MA population visits doctors an average of 5.12 times per year. This is close to the Centers for Disease Control and Prevention estimate of 4.98 doctor visits per year per individual age 65-or-older in 2016 (Ashman et al., 2019). Similarly, a \$1 increase in the hospital visit copay decreases marginal costs by \$0.129, which is comparable to the Kaiser Family Foundation estimate of 0.252 hospital visits per year per TM enrollee in 2015. This finding, when combined with our estimate of the consumer valuation of hospital copays, is consistent with previous findings of behavioral hazard in the use of care (Loewenstein et al., 2013, Baicker et al., 2015). Adding basic drug coverage to plans without drug coverage costs an average of \$438, whereas adding enhanced drug coverage costs an additional \$278. Dental coverage costs \$202 to provide. Taken together with the profit comparison above, we conclude that our estimates of the marginal cost function, which we obtain through inverting first-order conditions, are in-line with estimates made through other methods, including claims-based methods Angrist and Pischke (2010), Nevo and Whinston (2010).

Finally, our approach assumes that the MA risk-adjustment system is effective and that marginal costs (net of risk-adjustment) do not vary by the realized risks of the enrollees.[26]

[26]As discussed in the Introduction, this assumption is consistent with the recent literature that finds

If this assumption was violated and firms faced higher (lower) marginal costs for higher-risk enrollees *after risk adjustment*, our estimated marginal costs would be biased upward (downward), which could influence our counterfactual calculations. We test for this by estimating the relationship between the risk-adjusted bid and the realized risk of the plan. After aggregating to plan-year observations as projected risk is reported at that level, we estimate that a 1% increase in bids increases risk by 0.0047% (t-statistic = 0.77).[27] While this suggests that the risk-adjustment system may slightly under-compensate plans with higher-risk enrollees, selection with respect to the benchmark is likely second-order.

8.3 Policy functions

To estimate the parameters of Equation (21), the policy functions, for each product characteristic, we must first define $\tilde{I}_{ft}$, the approximation of the information set $\mathcal{I}_{ft}$. The profit function (12) suggests that systematic market-level cost shocks and changes in demographics likely influence product characteristics. We therefore include market-level average demographics, including the fraction of those 65-and-older who are White, Black, and Hispanic, the fraction of those 65-and-older in deep poverty, median household income, unemployment rate, and population density. We also include lagged market-level averages of all product characteristics in each regression.[28]

The results are reported in Table 6. As no plan in our panel changed their basic drug coverage indicator, we do not consider changes to that indicator in our counterfactual. In general, the signs of the coefficients line up with the prior that plans should improve benefits when the benchmark increases; the net effect of an increase in the benchmark for each plan

that risk adjustment and other policy measures have largely succeeded in significantly reducing MA plan's incentive to cream-skim healthy enrollees.

[27] We also re-estimate the regression of Table G.1 with the addition of the plan's realized risk as a covariate. We estimate that a 1% increase in average enrollee risk increases marginal costs by 0.014% (t-statistic = 3.91).

[28] We treat enhanced drug and DVH coverage indicators as continuous variables throughout this exercise for simplicity and consistency with our other characteristics. See Section C for details. Our results are robust to logit and probit specifications for these variables.

Table 5: Marginal cost parameter estimates

	γ_{xj}					
Variable	Aetna	BCBS	Humana	Kaiser	UHG	Other
Cost-sharing characteristics						
Part B reduction (per \$1000)	0.0166	0.0173	0.0173	0.0176	0.0174	0.0174
	(0.0017)	(0.0018)	(0.0017)	(0.0015)	(0.0018)	(0.0019)
Deductible (per \$1000)	-0.0165	-0.0172	-0.0172	-0.0174	-0.0173	-0.0172
	(0.0017)	(0.0018)	(0.0017)	(0.0014)	(0.0018)	(0.0019)
Primary care copay	-0.0005	-0.0005	-0.0005	-0.0005	-0.0005	-0.0005
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Specialist copay	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hospital copay (per \$1000)	-0.0115	-0.0120	-0.0120	-0.0122	-0.0121	-0.0121
	(0.0012)	(0.0013)	(0.0012)	(0.0010)	(0.0012)	(0.0013)
Supplemental coverage characteristics						
Prescription drug	0.0381	0.0399	0.0404	0.0393	0.0417	0.0399
	(0.0050)	(0.0046)	(0.0043)	(0.0042)	(0.0044)	(0.0048)
Enhanced prescription drug	0.0243	0.0255	0.0258	0.0251	0.0266	0.0255
	(0.0032)	(0.0029)	(0.0027)	(0.0027)	(0.0028)	(0.0031)
Dental	0.0178	0.0186	0.0189	0.0183	0.0195	0.0187
	(0.0023)	(0.0021)	(0.0020)	(0.0020)	(0.0021)	(0.0022)
Vision	0.0081	0.0085	0.0086	0.0084	0.0089	0.0085
	(0.0011)	(0.0010)	(0.0009)	(0.0009)	(0.0009)	(0.0010)
Hearing	0.0278	0.0291	0.0295	0.0286	0.0304	0.0291
	(0.0037)	(0.0033)	(0.0031)	(0.0031)	(0.0032)	(0.0035)
Demand unobservable (ξ_j)	0.0133	0.0139	0.0141	0.0137	0.0146	0.0140
	(0.0018)	(0.0016)	(0.0015)	(0.0015)	(0.0015)	(0.0017)
Mean marginal cost (\$)	11,473	10,908	10,750	11,107	10,435	10,903
Std. dev. marginal cost (\$)	1,520	1,283	1,193	1,297	1,172	1,353
Observations	3,857	9,102	12,800	1,719	4,852	32,506

Notes: This table summarizes our estimates of plan-characteristic-level marginal costs – i.e. the parameters of Equation (14). We calculate these parameters by 'inverting' the first-order conditions for profit maximization with respect to prices and product characteristics. See Section 6.1 for details. Observations are county-year-plans. Reported values are means across the relevant firm. Standard deviations are in parentheses.

is an increase in mean utility.[29] [30]

Multiple equilibria are possible in our model. In Appendix Table G.2, we explore empirical relevance of this possibility by re-estimating the policy functions separately for each

[29] There are two exceptions: the deductible and vision. The policy function estimates imply an increase in the benchmark increases the deductible and decreases the provision of vision benefits. By the end of our sample, the vast majority of plans have a zero deductible. These estimates appear to be driven by the fact that after the implementation of the ACA, the number of plans with a positive deductible decreased while the benchmark rates also declined. It is also the case that the vast majority of plans offer vision benefits suggesting that there is limited variation to identify the vision policy function. While the coefficient estimates are counter-intuitive for these two benefits, in both cases the impact of a change in the benchmark is small in magnitude and, as is seen below, have little impact on the calculation of the optimal benchmark.

[30] Though we present a linear specification, we have explored non-linear functions of B_j, first-differenced specifications, and machine learning techniques and did not obtain meaningful improvements over this specification.

Census Region. While region-within-firm point estimates differ, the confidence intervals generally overlap. We conclude that the existence of isolated markets with widely disparate equilibrium behavior is unlikely.[31]

Table 6: Policy function estimation results

	(1) Part B reduction	(2) Deduct-ible	(3) Prim. copay	(4) Spec. copay	(5) Hospital copay	(6) Enhanced drug	(7) Dental	(8) Vision	(9) Hearing	(10) ξ
Benchmark ($000s) $\times$										
Aetna	0.0085	0.0195	-0.8575	-0.3474	-0.1524	0.2907	0.0289	-0.0091	0.0883	0.7442
	(0.0027)	(0.0067)	(0.1820)	(0.2764)	(0.0123)	(0.0116)	(0.0104)	(0.0049)	(0.0095)	(0.1190)
BCBS	0.0090	0.0212	-0.7922	-0.3882	-0.1696	0.2816	0.0221	-0.0203	0.0622	0.8741
	(0.0027)	(0.0067)	(0.1819)	(0.2769)	(0.0123)	(0.0116)	(0.0103)	(0.0050)	(0.0096)	(0.1195)
Humana	0.0088	0.0233	-1.0460	-0.5791	-0.1492	0.2741	0.0350	-0.0241	0.0238	0.8178
	(0.0027)	(0.0067)	(0.1821)	(0.2772)	(0.0123)	(0.0117)	(0.0103)	(0.0050)	(0.0096)	(0.1195)
Kaiser	0.0081	0.0137	-0.1747	-0.6441	-0.1164	0.3040	0.0500	-0.0149	0.0851	0.5623
	(0.0027)	(0.0066)	(0.1808)	(0.2748)	(0.0122)	(0.0115)	(0.0102)	(0.0049)	(0.0094)	(0.1183)
UHG	0.0071	0.0120	-0.9967	-0.3682	-0.1383	0.2754	-0.0084	-0.0104	0.0872	0.5690
	(0.0027)	(0.0067)	(0.1820)	(0.2775)	(0.0124)	(0.0117)	(0.0104)	(0.0050)	(0.0096)	(0.1195)
Other	0.0125	0.0165	-1.1637	-0.8191	-0.1793	0.2796	0.0192	-0.0128	0.0708	0.6625
	(0.0027)	(0.0067)	(0.1818)	(0.2765)	(0.0123)	(0.0116)	(0.0103)	(0.0050)	(0.0095)	(0.1193)
Obs.	37,998	37,998	37,998	37,998	37,998	37,998	37,998	37,998	37,998	37,998
R^2	0.1371	0.0748	0.2417	0.2967	0.2816	0.0725	0.1448	0.1270	0.3346	0.3035

Notes: This table reports firm-characteristic-level estimates of the effect of changes in the benchmark on product characteristics – i.e. estimates of the $\beta_{f,l}$ of Equation (21). Each column reports the coefficients of a regression with the specified product characteristic as the dependent variable. The independent variables for each regression include the benchmark interacted with firm indicators, lagged values of market average product characteristics, and market demographics. We omit other covariates in this table for space. We instrument for the benchmark with the Census Rural/Urban Continuum category interacted with year fixed effects. All dollars are 2017 dollars. No plan in our sample changed basic drug coverage. Observations are county-year-plans. Robust standard errors are in parentheses.

9 Optimal Geographic Variation in MA Subsidies

We now turn to solving the optimal subsidy schedule problem of Equation (1): setting the benchmark schedule to maximize aggregate consumer welfare keeping government expenditures constant. Consumer welfare is given by Equation (10) and government expenditures are given by Equation (18). Evaluating candidate solutions to Equation (1) requires finding equilibrium vectors (p, x, ξ) for counterfactual benchmarks in each market. Our payoff function satisfies the smoothness condition of Assumption WB. Conditional on the exclusion restriction for our instrument, our policy function estimators satisfy Assumption CE, and our demand and cost estimators recover unique unobservables and cost parameters for each plan, thus satisfying Assumption DGP with the information sets $\mathcal{I}_i$ consisting of the parameters of the demand system and cost functions. We therefore proceed by applying Proposition 1.

[31] We repeated this exercise splitting markets by the level of the benchmark and found similar results.

For a market m and counterfactual benchmark B'_m, we 1) use the estimated policy functions to update x and ξ; 2) use the estimated γ parameters to update c and therefore $\hat{\mathcal{I}}$; and 3) solve for equilibrium prices by searching for a fixed point in the firms' first-order conditions (a partitioned solver G).

We focus on the counterfactual results here and leave the discussion of the technical details to Appendix C. We test the performance of our approach with Monte Carlo exercises in Appendix D and explore examples of the counterfactual outcomes in Appendix E.

We begin in Section 9.1 by describing the consumer-welfare-maximizing county-level MA benchmark policy for 2017, which, for simplicity, we refer to as the "optimal policy." In Section 9.2, we quantify the contribution of changes in product characteristics in affecting consumer welfare, and in Section 9.3 we explore the selection of markets for increased or decreased benchmarks based on the derivatives of the consumer welfare and government expenditure functions. In Section 9.4, we consider the performance of a linear policy rule. We explore alternative social welfare functions in Section 9.5 and summarize other results and robustness checks in Appendix F.

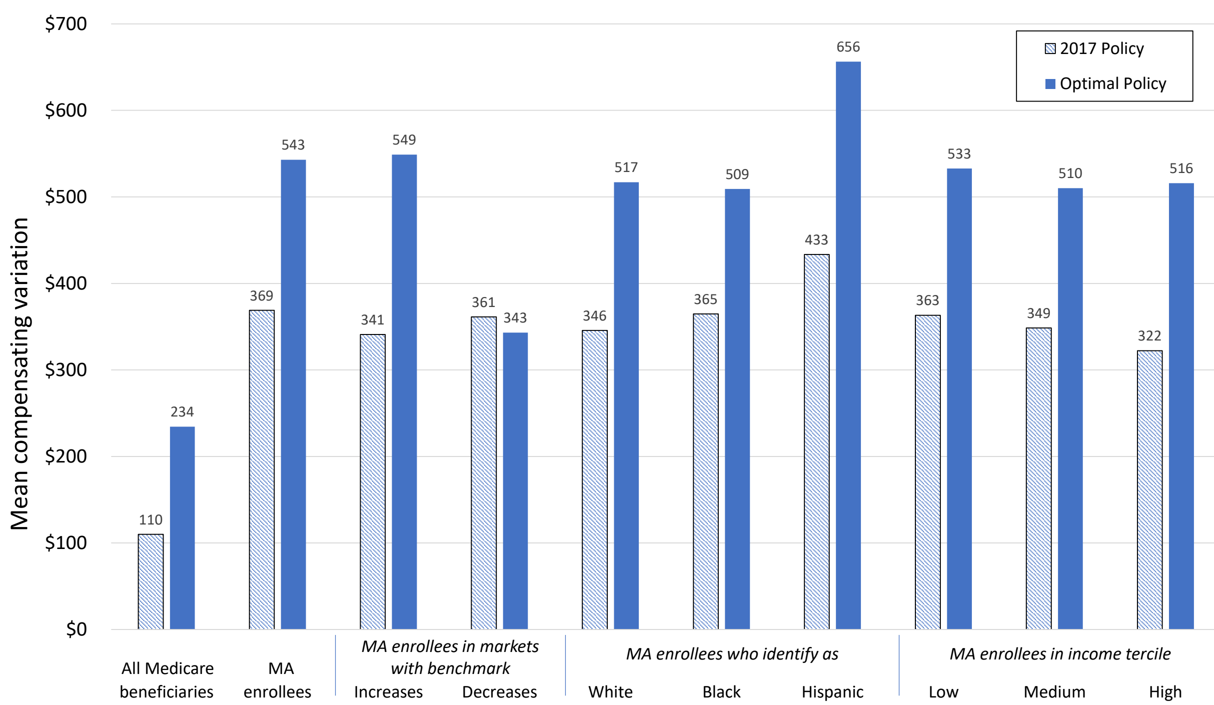
9.1 The consumer welfare maximizing benchmark schedule

Figure 2 presents our main results by comparing the annual compensating variation (CV) received by various groups of Medicare beneficiaries under the 2017 policy to the CV received under the policy that solves Equation (1). The first set of bars illustrates the mean CV for all Medicare beneficiaries and mean CV conditional on MA enrollment. The second set of bars splits markets by the direction of the benchmark change relative to the 2017 policy. The optimal policy increases the benchmark in 270 out of the 439 markets in our data, and in those markets the mean compensating variation for MA enrollees increases from \$340.99 to \$548.83, while in the 171 markets where the benchmark is lower, the mean compensating variation decreases from \$361.39 to \$343.18.

The third and fourth sets of bars examine the welfare changes by race and income. Previous work has identified inequalities in the Medicare and MA systems (Ayanian et al., 2014, Li et al., 2017), and, as a consequence, the welfare impact of redistributing the benchmarks on different demographic categories is of interest. While on average all groups gain welfare, compensating variation increases more in percentage terms for White and Hispanic enrollees than Black enrollees. High income beneficiaries benefit more in percentage terms than low or medium income beneficiaries. We note that this result does not imply that *uniform* changes in MA benchmarks would disproportionately affect any particular group.

This policy increases total MA share from 29.8% to 43.2% and increases aggregate annual consumer welfare (as measured by the sum of the compensating variation generated by MA) from \$4.62 billion to \$9.84 billion per year. The increase in compensating variation is driven both by moving individuals from TM to MA and by making MA more valuable to those already signed up for an MA plan. The increase in share in percentage terms (44.8%) roughly matches the percentage increase in the mean compensating variation per MA enrollee (47.2%) – thus these sources of gains are roughly equally important.

Figure 2: Compensating variation under the 2017 policy and the optimal policy

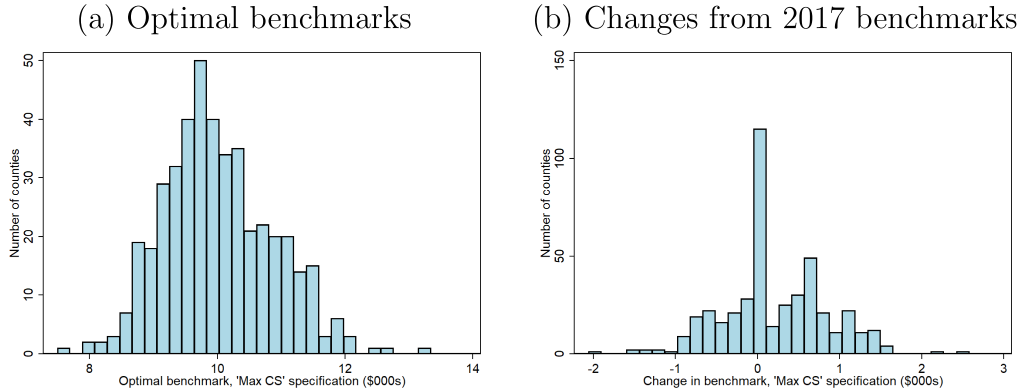


Notes: This chart illustrates the mean annual compensating variation for individuals in various groups under the 2017 policy and under the policy that solves Equation (1). We calculate compensating variation for all Medicare beneficiaries via Equation (10) and compensating variation conditional on MA enrollment via Equation (11). The optimal policy increases benchmarks in 268 markets and decreases benchmarks in 171 markets. All dollars are 2017 dollars. White, Black, and Hispanic groups are defined by CMS. We weight all calculations by the MCBS sample weights.

The benchmark changes are detailed in Figure 3. The left-hand histogram illustrates the distribution of benchmarks under the optimal policy and can be compared to Appendix

Figure G.1. The right-hand histogram shows the distribution of changes in the benchmark. The changes are generally modest – the interquartile range of the difference between the optimal and 2017 benchmarks is from -\$610 to \$627. The mean change is \$201 and the median is \$12. In percentage terms, roughly 87% of the changes in the benchmark are of less than 10% of the 2017 benchmarks; in 380 markets the absolute change is less than \$1,000.

Figure 3: The distribution of the optimal benchmarks across markets



Notes: Graph (a) illustrates the distribution of benchmarks under the policy that solves Equation (1). Graph (b) illustrates the distribution of differences between the optimal policy and the 2017 policy.

The magnitude of the changes in the mean surplus across the benchmark increase/decrease split suggests that the aggregate changes in these markets are also substantial. Table 7 reports the aggregate consumer welfare and government spending under the 2017 policy and the optimal policy by the direction of the benchmark change. The aggregate consumer welfare generated by MA increases in markets with benchmark increases from \$2.79 billion per year to \$8.87 billion per year. This change comes with a decrease in spending on TM of \$89.5 billion and an increase in MA spending of \$91.5 billion.[32] Aggregate consumer welfare decreases in the other markets from \$1.83 billion to \$0.97 billion, while spending transfers from MA to TM to balance the government's budget constraint.

[32]This transfer of resources from TM to MA may generate concerns about externalities with respect to the government's bargaining power (see e.g. Lakdawalla and Yin, 2015). However, as the government largely sets Medicare reimbursement rates nationally with local cost-of-living adjustments, these externalities are likely to be small over the range of TM and MA enrollment shares estimated in our counterfactuals. To check this, we regressed the log of the risk-adjusted TM cost on the log of the total MA share and county and year fixed effects. The coefficient on the MA share was -0.007 (t-value 2.19) indicating that this is likely a second-order concern.

Table 7: Aggregate market share, consumer welfare, and government spending under 2017 policy and optimal policy

Summary across all markets

	2017 policy	Optimal policy	% change
Total MA share (%)	29.8	43.2	44.8
Aggregate consumer welfare ($ billion)	4.62	9.84	113.2
Government spending on MA ($ billion)	126.3	195.8	55.1
Government spending on TM ($ billion)	314.2	244.7	-22.1
Total government expenditures ($ billion)	440.5	440.5	0.0

Markets in which benchmark increases (270 markets, 64.2% population share)

	2017 Policy	Optimal Policy	% Change
Total MA share (%)	28.6	57.3	100.4
Aggregate consumer welfare ($ billion)	2.79	8.87	218.3
Government spending on MA ($ billion)	78.1	169.6	117.0
Government spending on TM ($ billion)	215.9	126.4	-41.5
Total government expenditures ($ billion)	294.0	295.9	0.6

Markets in which benchmark decreases (169 markets, 35.8% population share)

	2017 Policy	Optimal Policy	% Change
Total MA share (%)	32.0	17.9	-44.0
Aggregate consumer welfare ($ billion)	1.83	.97	-47.0
Government spending on MA ($ billion)	48.2	26.3	-45.4
Government spending on TM ($ billion)	98.3	118.3	20.3
Total government expenditures ($ billion)	146.5	144.6	-1.3

Notes: This table reports aggregate statistics under the 2017 policy and the policy that solves Equation (1). We calculate all statistics at the individual level and aggregate using the MCBS sample weights. We calculate consumer welfare via Equation (10) and government spending using Equation (18). All dollars are 2017 dollars. Totals may differ due to rounding.

9.2 The role of changes in product characteristics

A key element of our contribution is the ability to incorporate changes in product characteristics into our evaluation of different benchmark schedules. In Appendix A, we construct a simple model of a monopolist firm who sets premium and non-premium characteristics when the government provides a subsidy and who faces heterogeneous consumers. In this framework, treating non-price characteristics as given leads to biased estimates of the welfare of changing the subsidy.[33] Below, we analyze the role of non-price attributes in our welfare calculations and the optimal benchmark schedule. These results align with the results from the simple example – ignoring the endogeneity of non-price attributes leads to a meaningfully different optimal benchmark schedule.

Figure 4 shows the plan-level distribution of product characteristics under the 2017 benchmark schedule and under the schedule that solves Equation (1). To highlight the way in which benchmark changes drive changes in product characteristics, we divide markets into two categories based upon the directional change in the benchmark from the 2017 policy to the optimal policy. The first box-and-whisker in each category illustrates the distribution of the characteristic in 2017, and the second box-and-whisker illustrates the distribution of the characteristic under the counterfactual benchmark. In other words, the top (bottom) set of boxes-and-whiskers in each graph refers to the set of 4,614 (2,228) plans in markets which would receive an increase (decrease) in the benchmark over (under) the 2017 level under the optimal policy.[34]

Figure 4 illustrates three notable patterns. First, benchmark increases tend to move product characteristics in a more generous direction (e.g. premiums and copays decrease). Decreases have the opposite effect. This is not universal as, for example, the mean deductible increases when the benchmark increases. Second, some characteristics move more than others. For example, the median hospital copay across plans in markets which experience

[33] This is true even in the absence of a non-zero-price constraint.

[34] The 'enhanced prescription drug', 'dental', 'vision', and 'hearing' characteristics are coded as indicators in the data. In our counterfactual analyses, we interpret values between 0 and 1 as less generous than the average coverage of that type in the 2017 data, and values above 1 as more generous. We model these characteristics as continuous variables for consistency with other characteristics, and simplicity of exposition and computation. We have explored alternative specifications using logit- and probit-based policy functions for these characteristics and found similar results. Contact the authors for details.

benchmark increases changes from \$1,363 in the 2017 data to \$1,277 under the optimal policy (a 6.3% decrease), whereas the median specialist copay across the same set of plans changes from \$35.33 to \$35.00 (a 0.94% decrease).

Finally, benchmark changes have different effects on the variance in characteristics depending on both the characteristic and the direction of the change. For example, the standard deviation in annual premiums across markets which receive benchmark increases changes from \$738 to \$624 (a 15.4% decrease), but increases from \$720 to \$777 (a 7.9% increase) in markets which receive benchmark decreases. In contrast, the standard deviation in primary care copays decreases (increases) by 0.24% (0.55%) in markets which receive benchmark increases (decreases).

Next, we decompose the contribution to consumer welfare attributable to changes in product characteristics. We perform this decomposition as follows. For any set of benchmarks $\{B_m\}$, our approach calculates a corresponding set of premiums and product characteristics for each market, X_m. Dropping the market subscripts, let $X(t)$ be a line through product space with $X(0)$ reflecting outcomes in the data and $X(1)$ reflecting outcomes at the optimal policy. The total gains to consumers by moving from current policy to the optimal policy can be written as $CS(X(1)) - CS(X(0))$. To understand how the gains are realized across different product characteristics, we decompose the overall gain using the gradient theorem. For any line $X(t)$,

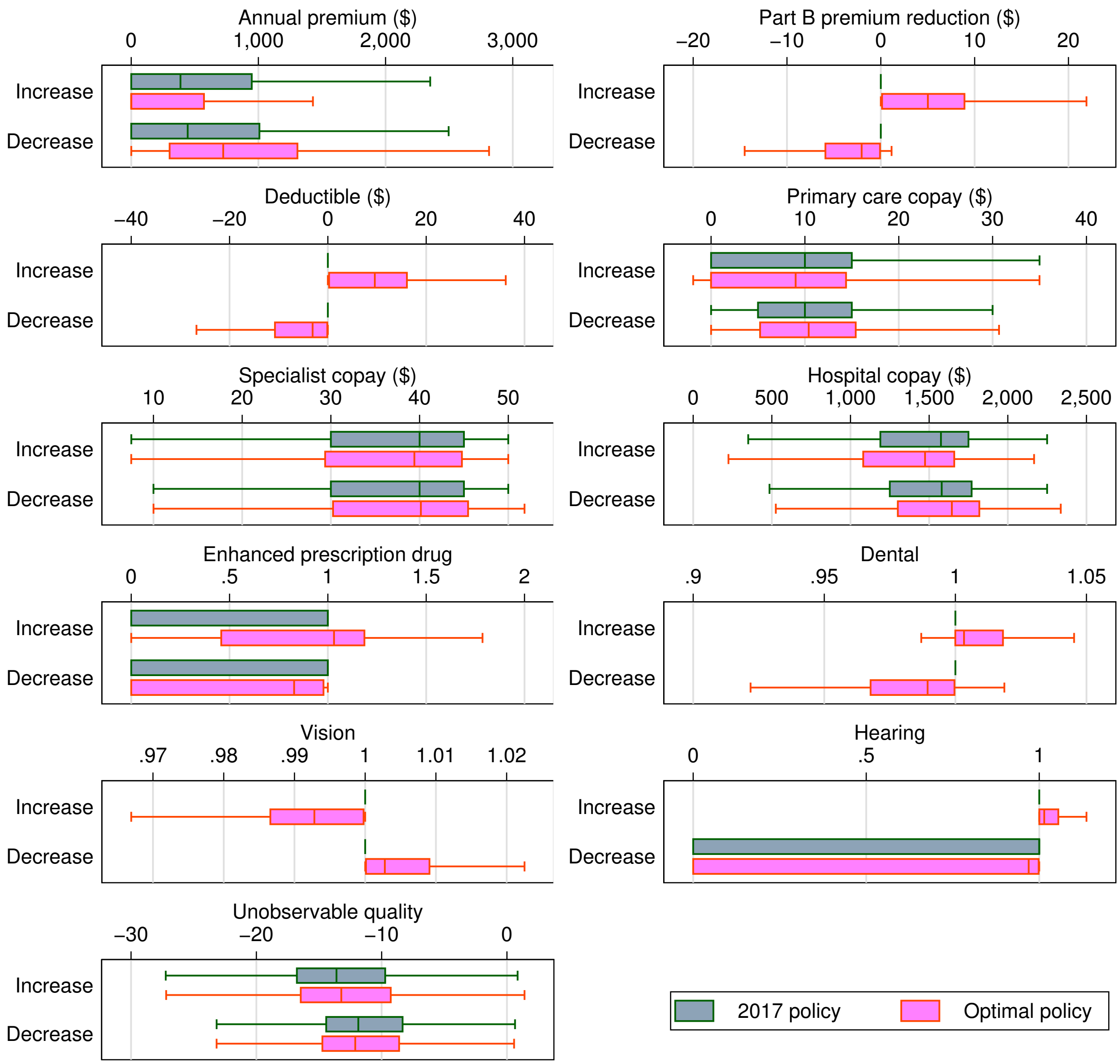
$$\begin{aligned} CS(X(1)) - CS(X(0)) &= \int_0^1 \nabla CS(X(t)) dX(t) dt \\ &= \nabla_X CS(X(0)) \cdot dX(1) + o(||dX(1)||), \end{aligned}$$

where $\nabla_X CS(X(0))$ is a $1 \times \#X$ vector of derivatives of CS with respect to premium and each product characteristic, and $dX(1)$ is a $\#X \times 1$ vector of changes in premiums and product characteristics. This can be rewritten as

$$\sum_i \sum_j \sum_k \nabla CS_{ijk}(X(0)) \cdot dX(1)_{jk} + o(||dX(1)||), \tag{22}$$

where i denotes consumers, j denotes products, and k denotes product characteristics (in-

Figure 4: The distribution of product characteristics under the 2017 policy and the optimal policy, by direction of benchmark change



Notes: These box-and-whiskers plots illustrate the unweighted plan-level distribution of product characteristics under the 2017 policy and the optimal policy. We divide markets according to the direction of the change in the benchmark when moving from the 2017 policy to the optimal policy. In other words, the set of 2017 box-and-whisker plots under the 'Increase' label illustrate the distribution of product characteristics in the data for the set of plans in markets which would receive an increase in the benchmark under the optimal policy. Outliers excluded for clarity.

cluding premiums). In short, the change in surplus for individual i is approximately equal to the sum of the effects of each product characteristic for each good. By re-arranging this expression and summing over goods for each product characteristic, we can compute the effects of changes in each product characteristic as the subsidy schedule changes from the actual policy to our calculated optimal policy—we approximate Equation (22) with

$$\frac{1}{\alpha}\left(\sum_{j}\sum_{k} s_{ij}(\delta(0),\mu(0)) + s_{ij}(\delta(1),\mu(1))\beta_{ik}\Delta X_{jk}\right).$$

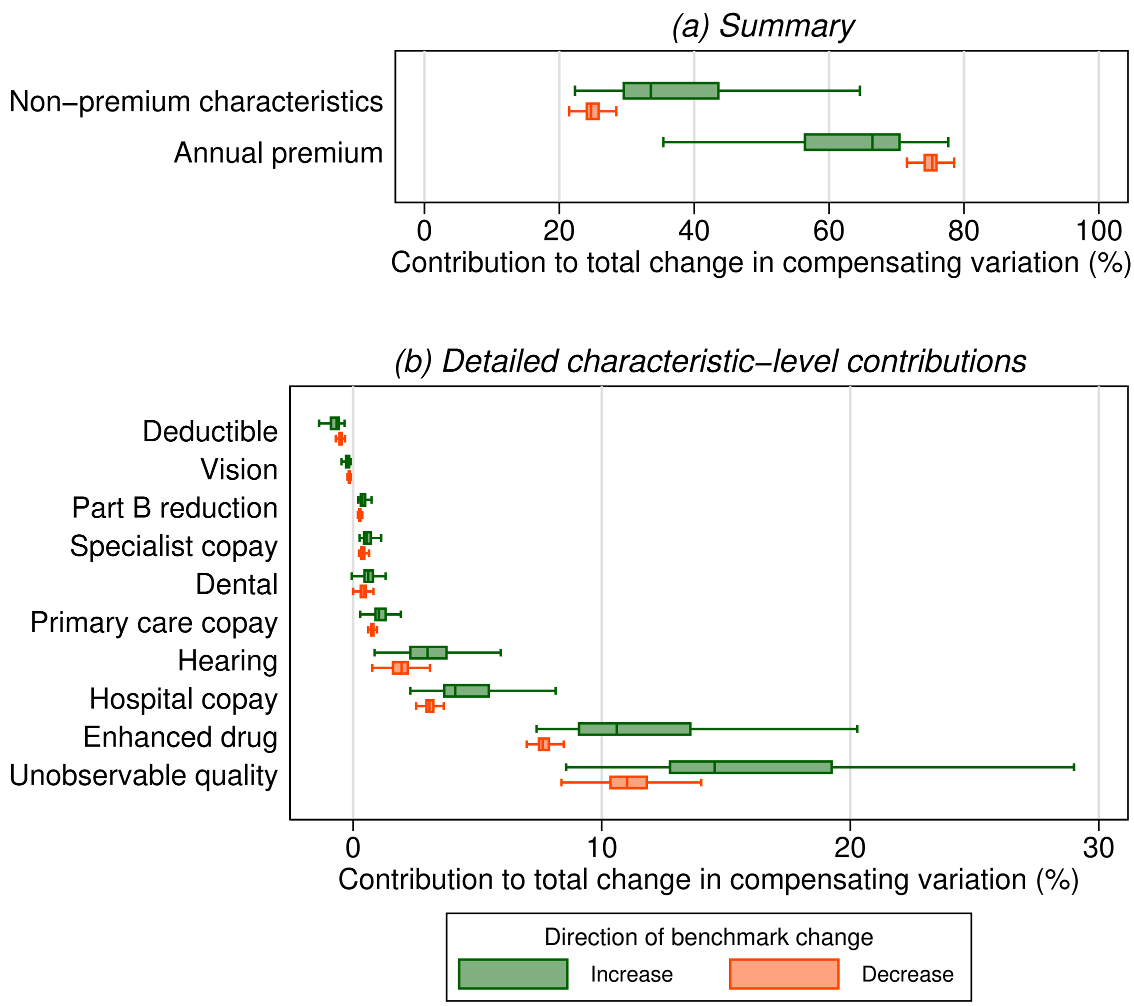
We present the results of this exercise in Figure 5 with box-and-whiskers plots of the beneficiary-level distribution of the contribution of product characteristics to the changes in consumer welfare. We group beneficiaries by the direction of the change in the benchmark and calculate percentage contributions to focus on the effect of differences in product characteristics rather than differences in consumer demographics which change the level of welfare.

In subfigure (a), we compare the consumer welfare impact of changes in all non-premium characteristics to changes in the annual premiums. These results highlight the importance of modeling product characteristics. In markets in which the benchmark increases (decreases), non-premium product characteristics contribute a net average of 41.2% (25.0%) of the change in compensating variation. This difference is driven in part by the zero-lower bound on premium. As the benchmark increases, the annual premium is driven to zero for more plans and so product characteristics take on a greater role. As the benchmark decreases, all plans can increase premiums.

In subfigure (b), we detail the consumer welfare contribution of each product characteristic. The largest contributions come from the unobservable product quality, ξ, enhanced prescription drug benefits, and hospital copays. On the other hand, changes in the deductible and vision coverage move in the opposite direction of the overall change in welfare but have a comparatively small impact of consumer well-being. This is driven by the policy function estimates reported in Table 6: for each firm, increases in the benchmark are estimated to (a) increase deductibles and (b) reduce the prevalence of vision coverage.

Despite these results, it is possible that our optimal policy is invariant to endogenizing

Figure 5: The contribution of product characteristics to changes in compensating variation



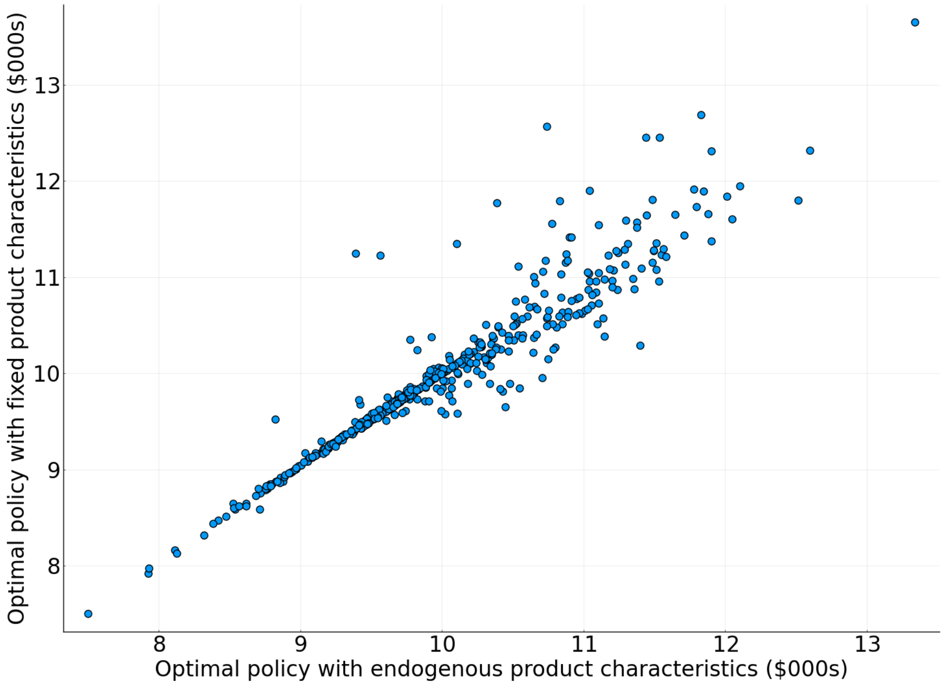
Notes: These box-and-whiskers plots illustrate the beneficiary-level distribution of the contribution of product characteristics to the changes in the compensating variation when moving from the 2017 policy to the optimal policy. To calculate these percentages, we decompose the total change in compensating variation into the changes stemming from changes in the premium and each of the product characteristics using the gradient theorem. A negative percentage indicates that changes in that characteristic moved compensating variation in the opposite direction from the total change. We weight the distributions using the MCBS sample weights. Outliers excluded for clarity.

product characteristics. We measure the contribution of product characteristics in an alternative way by recomputing the consumer surplus maximizing policy holding all non-premium characteristics fixed at 2017 levels. Changes in the benchmark are now pure premium subsidies. The mean absolute difference between the benchmarks under this "fixed characteristic" policy and our optimal policy is \$151.08, 75.1% of the mean change reported above. The pattern of differences between these policies is illustrated in Figure 6. In roughly half of the counties, the difference is greater than \$50. Relative to markets with smaller differences, these markets tend to have higher TM spending and more diverse product offerings (i.e. higher variance in non-price product characteristics within the market).

The differences between the policies is driven, in large part, by changes in pass-through behavior. When product characteristics are fixed, benchmark changes do not change underlying costs. However, when product characteristics change, underlying costs change, which changes the premium response. This, in turn, changes the effective pass-through from benchmarks to surplus and the 'bang for the buck' the optimizer 'sees' when determining the optimal policy. Intuitively, markets with more benchmark-sensitive plans (because of enrollee preferences for those characteristics or because of relative plan efficiencies) will receive larger benchmark increases under the optimal policy compared to a fixed characteristic setting. Furthermore, incorporating product characteristic responses likely leads to greater differences in predicted outcomes for those markets with greater product characteristic dispersion.[35] We note, however, that the optimal policy for any particular market depends on counterfactual behavior across all other markets. Indeed, markets with greater TM spending play a greater role in the overall budget constraint and thus the policies in such markets may be more subject to the estimated costs and benefits of changes in other markets.

[35]To see this, note that given the demand system of Equation (4), a uniform increase to all product characteristics will lead to different welfare gains in markets with dispersed product characteristics relative to markets with uniform product characteristics, though we note the sign and magnitude of the difference depends on the nature of the dispersion.

Figure 6: Comparing the optimal policies when product characteristics are endogenous and fixed



Notes: This graph compares solutions of Equation (1) when product characteristics vary as the benchmark varies (horizontal axis) and when product characteristics are held fixed (vertical axis). Each point is a market in 2017.

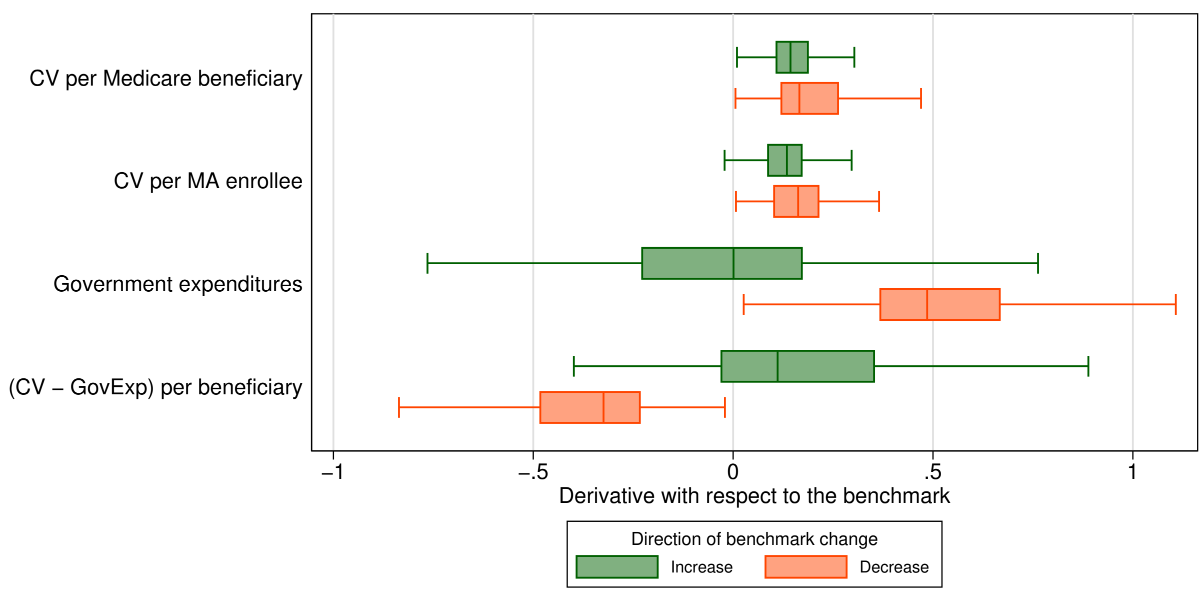
9.3 The selection of markets for benchmark increases and decreases

The nature of the solution to our optimal subsidy problem implies that markets should be selected for benchmark increases based upon the marginal impact of an change in the benchmark rate on consumer welfare and government expenditures. Figure 7 illustrates the distribution of the derivatives of CS, $\overline{CS}^{cond}$, $GovExp$, and 'total surplus' (defined as $CS - GovExp$) with respect to a \$1 increase in the benchmark rate from the 2017 level. For each function, we illustrate the distribution for markets that receive increases or decreases under the optimal policy separately.

The distributions of the derivatives of both CS and $\overline{CS}^{cond}$ overlap across the two sets of markets, suggesting that changes in compensating variation alone do not drive the results. Indeed, the mean change in CS is higher in the markets where the optimal policy decreases the benchmark. In contrast, both the distributions of the $GovExp$ and 'total surplus' derivatives are more separated across the two sets of markets. A total of 199 markets with benchmark increases also have positive total surplus derivatives at the 2017 policy level, and every market with a benchmark decrease has a negative total surplus derivative at the 2017 policy level. The derivative of $GovExp$ depends on premium elasticities, MA costs relative to TM costs, and the extent to which competition leads firms to pass-through increases in the benchmark to benefits. While it is difficult to disentangle these interrelated factors, to the extent that the optimizer seeks to maximize the "bang for the buck", these results suggest that the buck (how much more the government spends when the benchmark increases by a dollar) matters more than the bang (how much additional surplus consumers receive when the benchmark increases by a dollar).

Finally, we note that in many markets, the derivative of GS is negative at the 2017 benchmark. This can occur when the benchmarks are set lower than TM spending, and the cost-savings from beneficiaries that move from TM to MA outweighs the cost increases for inframarginal MA enrollees. We explore this in more detail in Section 9.5.

Figure 7: Derivatives of market-level welfare and spending functions with respect to the benchmark



Notes: These box-and-whiskers plots illustrate the market-level distribution of the derivatives of welfare and spending functions with respect to the benchmark. We calculate "CV per Medicare beneficiary" via Equation (10), "CV per MA enrollee" via Equation (11), and "Government expenditures" via Equation (18). To form these derivatives, we calculate these functions at the individual level, aggregate to the market level, and then take numeric derivatives by simulating a \$1 increase in the benchmark from the 2017 level. Markets are categorized according to the direction of the benchmark change when moving from the 2017 policy to the policy that solves Equation (1). We weight the resulting distributions by the MCBS sample weights. Outliers excluded for clarity.

Given the importance of these derivatives in determining the optimal benchmark schedule, it is natural to consider the extent to which market-level observables can explain the across market variance in these derivatives. We investigate this by modeling the derivatives of the consumer surplus and government expenditure functions as a linear function of our county-level observables from the Area Health Resources File. Appendix Table G.3 reports the means of these variables by benchmark quartile.

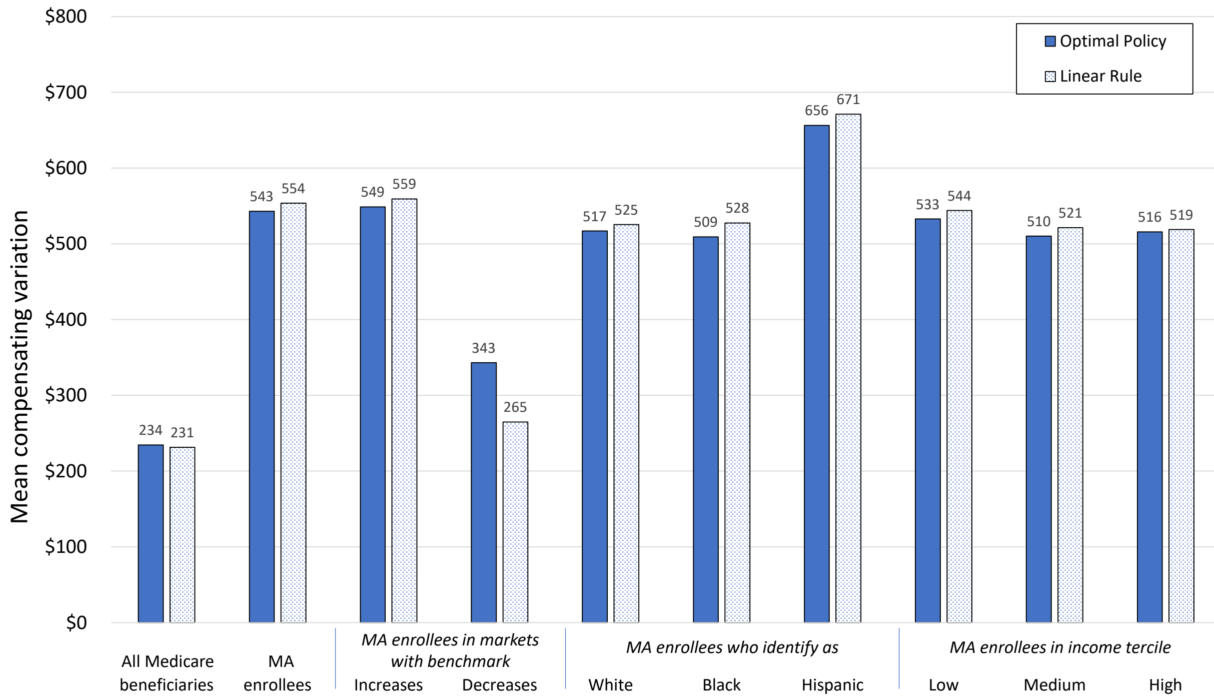
Columns (1) and (2) of Table 8 present standardized regression coefficients when the dependent variable is the derivative of consumer surplus and government expenditures, respectively. Across the two regressions, TM costs, measures of competition, per-capita income, and average risk score enter significantly.

9.4 A linear policy rule

We next examine the extent to which the optimal benchmark schedule can be approximated using linear regression with the the same set of county-level explanatory variables as in the derivative regressions discussed above. Analyzing such linear rules may be particularly policy-relevant because of the well known difficulties in implementing complex optimal tax and subsidy schemes (see e.g. Scott Morton, 1997, Decarolis, 2015, Jaffe and Shepard, 2017). Column (3) of Table 8 presents the coefficient estimates from regressing the optimal benchmark on the county-level variables. The explanatory variables include the number of firms offering plans but not the number of plans due to endogeneity concerns. The explanatory variables also include measures of race and ethnicity because of their importance in the demand system. The optimal benchmark is most strongly associated with measures of competition, TM costs, and per-income income. This regression fits the optimal benchmark schedule well – the R^2 is 0.88.

Figure 8 compares the optimal benchmark schedule to the policy generated by the fitted values of this regression. The linear rule closely matches the aggregate consumer welfare generated by the optimal policy. However, the linear rule increases government spending by 0.318%. Under this rule, 267 markets receive benchmark increases (as opposed to 270). The second set of bars shows that, on average, the linear rule results in changes which are too large relative to the optimal policy. Consumers in markets with benchmark decreases lose more surplus, and consumers in markets with benchmark increases gain more surplus.

Figure 8: Compensating variation under the optimal policy and the linear policy rule



Notes: This chart illustrates the mean compensating variation for individuals in various groups under the policy that solves Equation (1) and the policy fit by the regression in Column (3) of Table 8. We calculate compensating variation for all Medicare beneficiaries via Equation (10) and compensating variation conditional on MA enrollment via Equation (11). The markets with benchmark increases and decreases are not identical between the two policies. The optimal (linear) policy increases benchmarks in 268 (266) markets and decreases benchmarks in 171 (173) markets. The linear rule increases total government spending on Medicare from \$440.5 billion to \$441.9 billion. All dollars are 2017 dollars. White, Black, and Hispanic groups are defined by CMS. We weight all calculations by the MCBS sample weights.

9.5 Alternative social welfare functions

Using an objective function that simply maximizes aggregate consumer surplus results in a optimal policy schedule that creates large winners and losers relative to the current policy. It is therefore natural to explore other social welfare functions that consider the distributional impact of a change in benchmarks. We consider the optimal benchmark schedule that arises from using family of alternative objective functions given by

$$\max_{\{B_m\}} \zeta \sum_m \int_i CS_{im}(B_m)di - (1-\zeta)Var(CS) \quad \text{s.t.} \sum_m GovExp_m(B_m) = \bar{B}. \tag{23}$$

Table 8: Modeling the optimal policy and the derivatives of surplus and spending at the 2017 policy as a function of market-level observables

	(1)	(2)	(3)
	$\frac{\partial CS}{\partial B_m}$	$\frac{\partial GovExp}{\partial B_m}$	Log Optimal Benchmark
Log of Risk-adj. per-cap. TM costs	-0.2716	-0.7919	0.0877
	(0.0576)	(0.0362)	(0.0019)
Number of MA firms	0.1951	0.2256	-0.0030
	(0.0627)	(0.0394)	(0.0021)
Log of Medicare beneficiaries	0.1259	-0.0113	-0.0044
	(0.0835)	(0.0524)	(0.0027)
Share of 65+ population who is White	-0.5389	0.0794	0.0190
	(0.1849)	(0.1160)	(0.0061)
Share of 65+ population who is Black	-0.5188	0.1041	0.0136
	(0.1456)	(0.0914)	(0.0048)
Share of 65+ population who is Hispanic	-0.3752	0.0675	0.0106
	(0.1185)	(0.0744)	(0.0039)
Average risk score	0.1588	0.1849	-0.0043
	(0.0752)	(0.0472)	(0.0025)
Log of median household income	-0.0468	0.0206	-0.0084
	(0.0825)	(0.0518)	(0.0027)
Share of 65+ population in deep poverty	-0.0494	-0.0152	0.0008
	(0.0548)	(0.0344)	(0.0018)
Unemployment rate	-0.0006	-0.0969	-0.0006
	(0.0648)	(0.0407)	(0.0021)
Population density	0.0178	-0.1410	-0.0073
	(0.0537)	(0.0337)	(0.0018)
MDs per capita	-0.0217	-0.0265	0.0014
	(0.0634)	(0.0398)	(0.0021)
Medicare-qualified hospitals per capita	0.0004	0.0068	0.0016
	(0.0467)	(0.0293)	(0.0015)
Nursing facilities per capita	-0.0493	-0.0421	-0.0042
	(0.0595)	(0.0373)	(0.0020)
Hospice facilities per capita	-0.0232	-0.0226	-0.0001
	(0.0473)	(0.0297)	(0.0016)
Medicare hospital readmission rate	-0.0665	-0.0609	-0.0002
	(0.0633)	(0.0398)	(0.0021)
Observations	439	439	439
R-squared	0.1791	0.6765	0.8782

Notes: This table reports the relationships between various county-level demographic and marketplace variables and outputs from our model as estimated by OLS. The dependent variables for Columns (1) and (2) are the derivatives of the CS and $GovExp$ functions (respectively), as defined in Equations (10) and (18), taken at the 2017 policy with respect to the benchmark. The dependent variable in Column (3) is the subsidy schedule that solves Equation (1), in units of thousands of dollars per year. All independent variables and the dependent variables for Columns (1) and (2) have been normalized to have mean zero and unit variance. Robust standard errors are in parentheses.

Equation (23) is similar to Equation (1) with the addition of a penalty for variance in compensating variation parameterized by ζ. We also consider another social welfare function given by

$$\min_{\{B_m\}} \sum_m GovExp_m(B_m) \text{ s.t. } B_m \geq \bar{B}_m \forall m. \tag{24}$$

Equation (24) seeks to minimize government expenditures, potentially subject to a binding floor $\bar{B}_m$. We report aggregate share, welfare, and spending for solutions to these equations as well as the other benchmark schedules described above in Table 9, and illustrate the compensating variation for different groups of consumers compared to the policy that solves Equation (1) in Figure 9.

The effect of a penalty on the variance in compensating variation can be severe – a weight of merely 0.01 on the $Var(CS)$ term results in a reduction in benchmarks nearly everywhere in order to fund increases in a few markets. The aggregate consumer welfare generated by MA drops to \$2.37 billion, below the level of that generated by the 2017 benchmarks. Reducing the weight to 0.001 results in a policy that raises aggregate welfare to \$9.15 billion. The main difference in outcomes between this policy and the policy that solves Equation (1) is a reduction in surplus conditional on MA enrollment – the policy funds benchmark increases in 342 counties (as opposed to 270) and raises the total MA share to 47.4% (as opposed to 43.2%).

Solutions to Equation (24) seek to minimize government expenditures. Where the 2017 MA payments are larger than the cost of TM, this is done by reducing the benchmark. However, there are markets where the 2017 policy results in MA payments that are on average lower than TM costs. Increasing the benchmark results in both intensive and extensive margin changes to MA payments. The government must pay more for consumers who were already enrolled in an MA plan, and it must transfer payments from the TM system to the MA system for consumers who switch. In 175 markets, the extensive margin impact is larger than the intensive margin impact, and therefore an increase in the benchmark rate decreases total government expenditures.[36] As a result, when the government can freely choose benchmarks, it can reduce total spending on TM and MA by \$4.2 billion, though

[36] In Appendix E we explore in detail how government expenditures varies with the benchmark for three markets, including a market which features this behavior.

at the cost of \$5.40 billion in aggregate consumer welfare relative to the optimal policy. However, if the government is prohibited from reducing benchmarks below 2017 benchmarks, total consumer welfare increases from the 2017 policy to \$6.15 billion. At the same time, government spending is reduced by \$1.8 billion. We note that we consider the cost of public funds in this exercise for consistency with our baseline counterfactual. As long as the cost of public funds is constant across the geographies where those funds are spent, including those costs will not change this policy.

Table 9: Aggregate share, consumer welfare, and government spending under alternative benchmark policies and social welfare functions

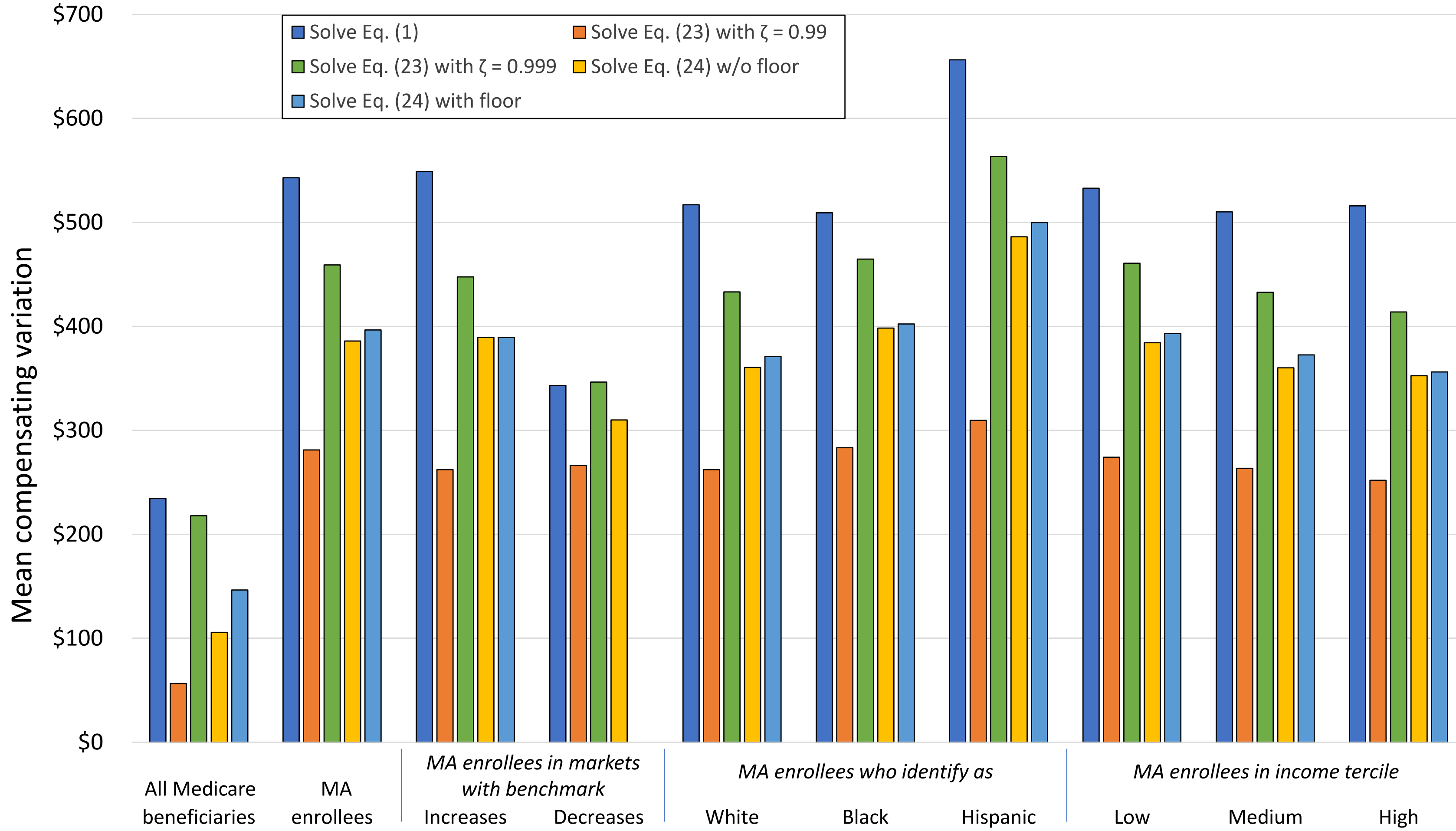
	(1) 2017 policy	(2) Optimal policy	(3) Linear rule	(4) Eq. 23 $\zeta = 0.99$	(5) Eq. 23 $\zeta = 0.999$	(6) Eq. 24 w/o floor	(7) Eq. 24 w/ floor
MA share (%)	29.8	43.2	41.8	20.1	47.4	27.4	33.2
$\sum CS$ (\$ billion)	4.62	9.84	9.72	2.37	9.15	4.44	6.15
$\sum GovExp$ (\$ billion)	440.5	440.5	441.9	440.5	440.5	436.3	438.7
Mkts. w/ increases	N/A	270	267	188	342	178	178

Notes: This table reports aggregate outcomes under a variety of benchmark schedules. The schedule in Column (1) is the 2017 benchmarks. The schedule in Column (2) solves Equation (1). The schedule in Column (3) is fit by the regression in Column (3) of Table 8. The schedules in Columns (4) and (5) solve Equation (23) (consumer welfare maximization with a penalty on the variance of welfare) with different choices of ζ. The schedules in Columns (6) and (7) solve Equation (24) (government expenditure minimization) with a non-binding and binding floor set to the 2017 benchmarks, respectively. We calculate all statistics at the individual level and aggregate using the MCBS sample weights. We calculate consumer welfare via Equation (10) and government spending using Equation (18). All dollars are 2017 dollars.

10 Conclusion

Seeking to reduce the perceived inefficiency of government-provided goods and services, policy makers have implemented public-private partnerships in which the government provides subsidies to private firms that are tied to consumers' choices. The idea is that market forces will bring down the total cost and increase the benefits of the goods over time. In many cases, the goods are meaningfully differentiated. Additionally, these goods may be offered in geographies with consumers who have substantially different preferences. While the subsidy rates are generally set according to measures of government costs, the optimal subsidies conditional on a fixed budget depend on equilibrium interactions between endogenously dif-

Figure 9: Compensating variation under the optimal policy and the linear policy rule



Notes: This chart illustrates the mean compensating variation for individuals in various groups under the policy that solves Equation (1), policies that solve Equation (23) (consumer welfare maximization with a penalty on the variance of welfare) with different choices of ζ, and policies that solve Equation (24) (government expenditure minimization) with a non-binding floor and with a binding floor set equal to the 2017 policy. We calculate compensating variation for all Medicare beneficiaries via Equation (10) and compensating variation conditional on MA enrollment via Equation (11). The markets with benchmark increases and decreases are not identical between the various policies. All dollars are 2017 dollars. White, Black, and Hispanic groups are defined by CMS. We weight all calculations by the MCBS sample weights.

ferentiated firms and heterogeneous consumers.

We provide a framework for calculating the optimal market-level subsidy schedule that takes into account both supply and demand responses to alternative subsidy rates. We model demand with a discrete-choice system and avoid the curse of dimensionality in computing counterfactual product characteristics with an approach that combines policy function estimation with a first-order condition solver.

We apply our framework to Medicare Advantage in the United States, through which almost half of U.S. seniors obtain Medicare benefits, and estimate our model using micro-level panel data. We derive demand instruments from detailed data on costs, which are likely to be available in other managed competition environments, and from the panel nature of our product data. We derive policy function instruments from county-level differences in urban/rural status following Duggan et al. (2016). In contrast to previous work on MA, we fully endogenize both premiums and product characteristics and consider firm-level characteristic policy functions and product-level characteristic costs.

We find that the optimal (budget-neutral) subsidy schedule differs substantially from the one currently employed by the government. The 2017 policy generates an average of \$109.95 in consumer welfare per Medicare beneficiary per year as measured by compensating variation. By maximizing the mean consumer welfare while holding government spending constant, we find a policy that results in an average of \$234.40 in welfare per beneficiary per year. These gains come both from finding markets where it is easy to move people from TM to MA (in the sense of needing fewer government dollars) and from improving benefits for people already utilizing MA. We show that freely-available market-level observables can be used to approximate the optimal policy with a linear rule that captures over 95% of the consumer welfare gains at the cost of an increase in government expenditures of 0.318% relative to the optimal policy. Apart from finding a particular consumer-surplus-maximizing policy, which may not be implementable for political or other reasons, our framework can be used to evaluate the outcomes of any proposed subsidy schedule.

Accounting for endogenous product characteristics is important. Changes in non-price product characteristics account for over 35% of the changes in total surplus, and a policy computed under the assumption that non-price characteristics remain constant is significantly different than the policy that allows non-price characteristics to adjust. Modeling product characteristics is particularly important in markets with more diverse product offerings and/or more heterogeneous responses to changes in subsidy policies.

Our model does not explicitly take risk selection into account. Both our work and the work of others indicates that risk selection is likely second-order in the MA setting because of the risk adjustment mechanism. However, risk selection may play a greater role in other contexts. In those cases, if the appropriate data is available, a modification of our approach to solving for the optimal across market subsidy can be applied. Incorporating selection effects directly into evaluations of potential subsidy schedules requires detailed data on the risk profiles of consumers and their associated cost differentials. We also note that risk selection could interact with the social welfare function: a social planner may wish to up-weight the expected surplus of those with poor health status.

Our framework can be adopted to any market in which subsidized firms offer differentiated products and where the necessary regularity conditions we characterize apply. For example, many charter schools offer specialized curricula which may appeal to different sets of parents. With data on family characteristics and choices, the benefits created by these schools and the outcomes of alternative voucher-style policies could be calculated.

References

Adrion, E. R. (2020). Politics and policymaking in medicare part c. *Medical care*, 58(3):285–292.

Aizawa, N. and Kim, Y. S. (2018). Advertising and risk selection in health insurance markets. *American Economic Review*, 108(3):828–67.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2):3–30.

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

Ashman, J. J., Rui, P., and Okeyode, T. (2019). Characteristics of office-based physician visits, 2016. *NCHS Data Brief, no 331*.

Ayanian, J. Z., Landon, B. E., Newhouse, J. P., and Zaslavsky, A. M. (2014). Racial and ethnic disparities among enrollees in medicare advantage plans. *New England Journal of Medicine*, 371(24):2288–2297.

Baicker, K., Mullainathan, S., and Schwartzstein, J. (2015). Behavioral hazard in health insurance. *The Quarterly Journal of Economics*, 130(4):1623–1667.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Baum-Snow, N. and Marion, J. (2009). The effects of low income housing tax credit developments on neighborhoods. *Journal of Public Economics*, 93(5):654–666.

Benkard, C. L., Bodoh-Creed, A., and Lazarev, J. (2018). Simulating the dynamic effects of horizontal mergers: U.s. airlines. *Working Paper*.

Berenson, R. A. (2008). From politics to policy: A new payment approach in medicare advantage. *Health Affairs*, 27(Suppl1):w156–w164.

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

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

Brown, J., Duggan, M., Kuziemko, I., and Woolston, W. (2014). How does risk selection respond to risk adjustment? new evidence from the medicare advantage program. *American Economic Review*, 104(10):3335–64.

Bundorf, M. K., Levin, J., and Mahoney, N. (2012). Pricing and welfare in health plan choice. *American Economic Review*, 102(7):3214–48.

Cabral, M., Geruso, M., and Mahoney, N. (2018). Does privatized medicare benefit patients or producers? evidence from the benefits improvement and protection act. *American Economic Review*.

Curto, V., Einav, L., Levin, J., and Bhattacharya, J. (2021). Can health insurance competition work? evidence from medicare advantage. *Journal of Political Economy*, 129(2):570–606.

Dalton, C. M., Gowrisankaran, G., and Town, R. (2018). Salience, myopia, and complex dynamic incentives: Evidence from Medicare Part D. Technical report, National Bureau of Economic Research.

Decarolis, F. (2015). Medicare part d: Are insurers gaming the low income subsidy design? *American Economic Review*, 105(4):1547–80.

Decarolis, F. and Guglielmo, A. (2017). Insurers' response to selection risk: Evidence from Medicare enrollment reforms. *Journal of Health Economics*, 56:383–396.

Decarolis, F., Polyakova, M., and Ryan, S. P. (2020). Subsidy design in privately provided social insurance: Lessons from medicare part d. *Journal of Political Economy*, 128(5):1712–1752.

Diamond, P. A. and McFadden, D. L. (1974). Some uses of the expenditure function in public finance. *Journal of Public Economics*, 3(1):3–21.

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

Einav, L., Finkelstein, A., and Tebaldi, P. (2018). Market design in regulated health insurance markets: Risk adjustment vs. subsidies. *MIT Working Paper*.

Enthoven, A. C. (1978). Consumer-Choice Health plan (second of two parts). A national-health-insurance proposal based on regulated competition in the private sector. *N. Engl. J. Med.*, 298(13):709–720.

Ericson, K. M. and Starc, A. (2015). Measuring consumer valuation of limited provider networks. *American Economic Review*, 105(5):115–19.

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

Geruso, M. and Layton, T. J. (2017). Selection in health insurance markets and its policy remedies. *Journal of Economic Perspectives*, 31(4):23–50.

Goolsbee, A. and Petrin, A. (2004). The consumer gains from direct broadcast satellites and the competition with cable tv. *Econ*, 72(2):351–381.

Gruber, J. (2017). Delivering public health insurance through private plan choice in the united states. *Journal of Economic Perspectives*, 31(4):3–22.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–82.

Hicks, J. R. (1945). The generalised theory of consumer's surplus. *The Review of Economic Studies*, 13(2):68–74.

Hoxby, C. M. (2000). Does competition among public schools benefit students and taxpayers? *American Economic Review*, 90(5):1209–1238.

Jaffe, S. P. and Shepard, M. (2017). Price-linked subsidies and health insurance markups. Working Paper 23104, National Bureau of Economic Research.

Kaiser Family Foundation (2021). State health facts: Medicare service use: Hospital inpatient services. Technical report, Kaiser Family Foundation.

Laffont, J.-J. and Tirole, J. (1993). *A theory of incentives in procurement and regulation*. MIT press.

Lakdawalla, D. and Yin, W. (2015). Insurers' negotiating leverage and the external effects of medicare part d. *Review of Economics and Statistics*, 97(2):314–331.

Li, Y., Cen, X., Cai, X., Thirukumaran, C. P., Zhou, J., and Glance, L. G. (2017). Medicare advantage associated with more racial disparity than traditional medicare for hospital readmissions. *Health Affairs*, 36(7):1328–1335.

Loewenstein, G., Friedman, J. Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., Beshears, J., Choi, J. J., Kolstad, J., Laibson, D., et al. (2013). Consumers' misunderstanding of health insurance. *Journal of Health Economics*, 32(5):850–862.

Lustig, J. (2010). Measuring Welfare Losses from Adverse Selection and Imperfect Competition in Privatized Medicare. Working paper, Yale University.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, pages 105–142.

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

McKean, R. N. and Minasian, J. R. (1966). On achieving pareto optimality-regardless of cost! *Economic Inquiry*, 5(1):14.

MedPAC (2017). Status report on the medicare advantage program. In *Report to the Congress: Medicare Payment Policy*, chapter 13, pages 345–379. MedPAC.

MedPAC (2019). *Report to the Congress, Medicare Payment Policy*. Medicare Payment Advisory Commission.

Meyers, D. J., Mor, V., Rahman, M., and Trivedi, A. N. (2021). Growth in medicare advantage greatest among black and hispanic enrollees. *Health Affairs*, 40(6):945–950.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, pages 395–421.

Nevo, A. and Whinston, M. D. (2010). Taking the dogma out of econometrics: Structural modeling and credible inference. *Journal of Economic Perspectives*, 24(2):69–82.

Newhouse, J. P., Landrum, M. B., Price, M., McWilliams, J. M., Hsu, J., and McGuire, T. G. (2019a). The comparative advantage of medicare advantage. *American journal of health economics*, 5(2):281–301.

Newhouse, J. P., McWilliams, J. M., Price, M., Huang, J., Fireman, B., and Hsu, J. (2013). Do medicare advantage plans select enrollees in higher margin clinical categories? *Journal of health economics*, 32(6):1278–1288.

Newhouse, J. P., Price, M., Hsu, J., McWilliams, J. M., and McGuire, T. G. (2015). How much favorable selection is left in medicare advantage? *American journal of health economics*, 1(1):1–26.

Newhouse, J. P., Price, M., Huang, J., McWilliams, J. M., and Hsu, J. (2012). Steps to reduce favorable risk selection in medicare advantage largely succeeded, boding well for health insurance exchanges. *Health Affairs*, 31(12):2618–2628. PMID: 23213145.

Newhouse, J. P., Price, M., McWilliams, J. M., Hsu, J., and McGuire, T. (2014). How much favorable selection is left in medicare advantage? Working Paper 20021, National Bureau of Economic Research.

Newhouse, J. P., Price, M., McWilliams, J. M., Hsu, J., Souza, J., and Landon, B. E. (2019b). Adjusted mortality rates are lower for medicare advantage than traditional medicare, but the rates converge over time. *Health affairs*, 38(4):554–560.

Nosal, K. (2012). Estimating switching costs for medicare advantage plans. Working paper, University of Arizona.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110:705–729.

Poterba, J. M. (1996). Government intervention in the markets for education and health care: how and why? In *Individual and social responsibility: Child care, education, medical care, and long-term care in America*, pages 277–308. University of Chicago Press.

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

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033.

Scott Morton, F. M. (1997). The interaction between a most-favored-customer clause and price dispersion: an empirical examination of the medicaid rebate rules of 1990. *Journal of Economics & Management Strategy*, 6(1):151–174.

Smith, P. C., Stepan, A., Valdmanis, V., and Verheyen, P. (1997). Principal-agent problems in health care systems: an international perspective. *Health policy*, 41(1):37–60.

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

Spence, A. M. (1975). Monopoly, quality, and regulation. *The Bell Journal of Economics*, pages 417–429.

Spence, M. (1976). Product differentiation and welfare. *The American Economic Review*, 66(2):407–414.

Sweeting, A. (2007). Dynamic product repositioning in differentiated product markets: The case of format switching in the commercial radio industry.

Tebaldi, P. (2017). Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the aca. Working paper, University of Chicago.

Town, R. and Liu, S. (2003). The welfare impact of Medicare HMOs. *Rand J Econ*, 34(4):719–736.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406.

Appendices

A The role of non-price characteristics

The analysis in this paper is predicated on the idea that firms may change price and non-price characteristics of their products in response to per-capita subsidies offered by the government. In this Appendix we explore this phenomenon in the form of a simplified model relative to the one presented starting Section 5.

Consider a monopolist offering a single product at price p to a measure of consumers. The product has a single non-price 'quality' characteristic x, also chosen by the firm. Demand for the product follows a discrete choice framework. Consumer i has price sensitivity α_i and preference for quality β_i. The utility of purchasing the good is given by $u_i = \alpha_i p + \beta_i x + \epsilon_i$, where ϵ_i is taken to be Type-I Extreme Value. Given a distribution of consumer preferences, demand is given by $s(p, x) = \int \frac{\exp(\alpha_i p + \beta_i x)}{1+\exp(\alpha_i p + \beta_i x)} di$.

The marginal cost of providing the product is a function of quality: $c(x) = \exp(c_b + c_x x)$ where c_b and c_x are parameters. The firm earns a per-capita subsidy from the government B. The firm's problem is therefore

$$\max_{p,x}(p - c(x) + B)s(p, x).$$

Suppose that the market consists of two types of consumers with equal frequency. Type 1 consumers have $\alpha_1 = -1.5, \beta_1 = 0.25$. Relative to type 1 consumers, type 2 consumers are more price sensitive and have a stronger taste for quality: $\alpha_2 = -4.0, \beta_2 = 2.0$. Let $c_b = -0.5$ and $c_x = 1.0$. Figure A.1 illustrates outcomes under these assumptions. In each subfigure, the solid lines indicate outcomes when the firm can freely choose both the price and the quality of the good. The dashed lines indicate outcomes when quality is held fixed at the level the firm chooses when it is unconstrained and it receives zero subsidy.

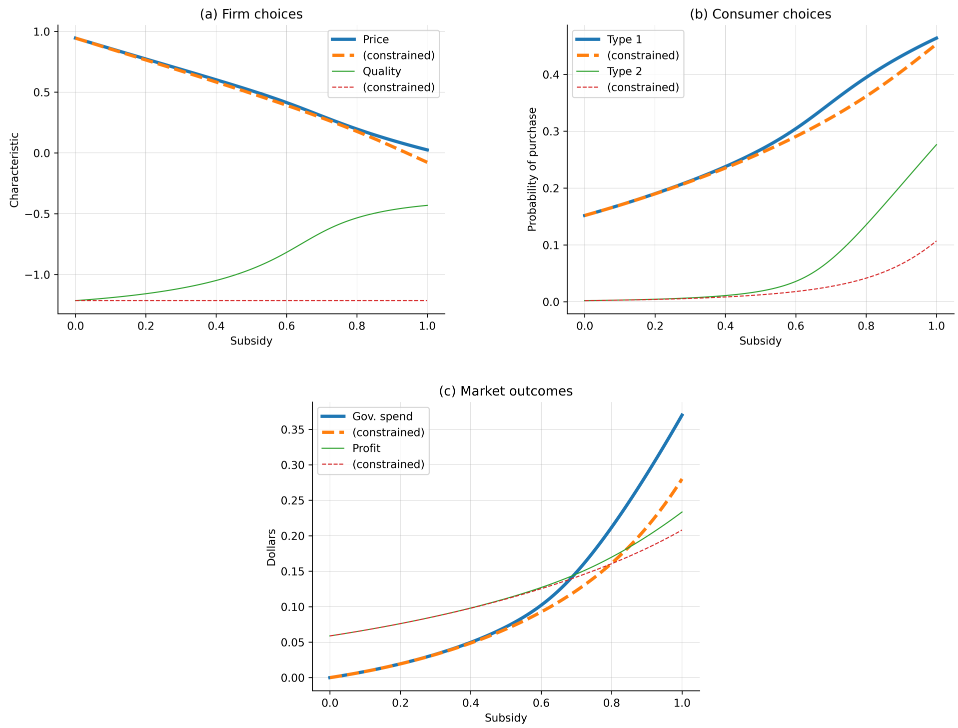
Subfigure (a) illustrates the firm's optimal choices. As the subsidy increases, the price decreases and the quality increases. We note that this need not always hold. For some combinations of consumer preferences and firm costs the firm's quality policy function is decreasing in the subsidy. The price decreases more under the constrained scenario than the unconstrained scenario, though we note the difference in prices remains small relative to the price level even as the subsidy increases.

Subfigure (b) illustrates the choices of the two types of consumers. Type 1 consumers are more likely to purchase the good across the range of subsidies, and the difference between their choices in the constrained and unconstrained scenarios remains small relative to the likelihood of purchase. A significantly larger difference is seen in the behavior of Type 2 consumers. As the subsidy increases, the increase in the probability of purchase under the unconstrained scenario is more than double the increase in the probability of purchase under the constrained scenario.

Finally, subfigure (c) illustrates government spending and firm profits. As the subsidy increases, the firm's profits increase. This increase is larger in the unconstrained scenario. At a subsidy of $B = 1.0$, the firm earns 12% higher profits when it is able to change both the price and quality of the product. The gap in government spending is even larger. At $B = 1.0$, the government spends 32% more when the firm can change its quality.

To summarize, in this model, taking non-price product characteristics as given when they

Figure A.1: Choices and outcomes in a simplified model



Notes: This figure plots choices and outcomes for the model described in Appendix A. In the model, a monopolist offers a single product with a single non-price 'quality' characteristic to a measure of consumers consisting of two types. See text for details and parameters. In each graph we plot outcomes in two scenarios. In the first scenario (solid lines), we allow the firm to freely choose both price and quality. In the second scenario (dashed lines), we hold the quality fixed at the same level as the optimal choice when the firm is not subsidized.

are in fact endogenous would significantly underestimate the impact of subsidies on consumer welfare and government spending. As these are key inputs into the optimal subsidy schedule problem, it is likely that such an analysis would generate an optimal subsidy schedule that is considerably different than the schedule generated by an analysis that incorporated endogenous non-price characteristics.

These differences are apparent even in the absence of the zero-lower-bound on prices or any heterogeneity in costs across consumers. In other words, risk selection incentives are not necessary to generate responses in non-price characteristics. Here, the response is driven solely by the heterogeneity in demand. As the subsidy increases, the firm chooses to attract more type 2 consumers by increasing the quality of their product.

B Accounting for the CMS Bidding System

In Section 6, we model firms as choosing prices and product characteristics while receiving a subsidy as a function of those product characteristics. In reality, firms choose a bid b_{jt} and a premium p_{jt} in addition to product characteristics. In this Appendix, we show how the CMS bidding rules can be used to transform the bidding problem into the price-setting problem.

The government pays the firm the amount of their bid plus a rebate payment if their bid is less than the benchmark. After taking into account risk adjustment, using Equation (3) we can write the rebate payment as a function of the bid b_{jt} and the plan-level benchmark B_{jt} with

$$reb(b_{jt}; B_{jt}, \lambda_{ft}) = \begin{cases} \lambda_{ft}(B_{jt} - b_{jt}) & \text{if } b_{jt} < B_{jt} \\ 0 & \text{if } b_{jt} \geq B_{jt} \end{cases}, \tag{25}$$

where λ_{ft} is the rebate percentage.

CMS requires that rebate payments be used to fund benefits beyond those offered by TM. Both supplemental benefits and cost-sharing reductions may be paid for with rebate funds, but cost-sharing reductions *must* be paid for with rebate funds. We assume this constraint is binding in the following sense: the rebate the firm receives is exactly equal to the incremental cost of providing the plan's cost-sharing benefits over the cost of providing a 'base' TM-equivalent plan. Let $x_{c,jt}$ be the subvector of product characteristics capturing cost-sharing reductions. We write the incremental cost function as $incr_{jt}(x_{c,jt})$ which is greater than zero as firms must provide at least TM-equivalent coverage, though it may vary by firm. We assume

$$reb_{jt} = incr_{jt}(x_{c,jt}). \tag{26}$$

Combined, Equations (25) and (26) imply that the bid itself is determined by the choice of cost-sharing benefits $x_{c,jt}$ and the incremental cost function $incr_{jt}$. In other words, our assumption can be reinterpreted to mean that insurers bid in such a way that the rebate they receive exactly pays for the incremental cost of providing cost-sharing benefits. There is therefore a continuous function mapping the choice of $x_{c,jt}$ onto the bid b_{jt}:

$$b(x_{c,jt}; B_{jt}, \lambda_{ft}, \cdot) = B_{jt} - \frac{1}{\lambda_{ft}} incr_{jt}(x_{c,jt}). \tag{27}$$

As a consequence, we can model the firm as simply choosing a premium, p_j, and product characteristics. The government pays the minimum of the bid plus the rebate and the benchmark: $sub(b_{jt}; B_{jt}, \lambda_{ft}) = \min\{b_{jt} + \lambda_{ft}(B_{jt} - b_{jt}), B_{jt}\}$. Thus, since $incr_{jt}(\cdot) \geq 0$,

$$sub(x_{c,jt}; B_{jt}, \lambda_{ft}, incr_{jt}) = B_{jt} - \frac{1-\lambda_{ft}}{\lambda_{ft}} incr_{jt}(x_{c,jt}). \tag{28}$$

Finally, we define the incremental cost function. Let $c_{b,ft} = \gamma_f + \gamma_s + \gamma_r + \gamma_m + \gamma_t$. Let $\gamma_{xc,jt}$ be the subvector of cost parameters associated with $x_{c,jt}$ and let $\omega_{c,jt}$ be the unobservable component of costs associated with cost-sharing benefits. The log-cost of cost-sharing benefits is then $c_{c,jt} = \gamma'_{xc,jt} x_{c,jt} + \omega_{c,jt}$.[37] We define the incremental cost function as

$$incr(x_{c,jt}; x_{b,ft}, \gamma_{xc,jt}, \omega_{c,jt}) = \exp[c_{b,ft} + c_{c,jt}(x_{c,jt})] - \exp(c_{b,ft}). \tag{29}$$

Since $incr(\cdot)$ is smooth, $sub(\cdot)$ is smooth, which is necessary to satisfy Assumption SP.

B.1 Estimating plan-characteristic-level costs

We estimate plan-characteristic-level costs using Equation (17) and the above definition of $sub(\cdot)$. In particular, note that for the set of characteristics which are not cost-sharing characteristics, denoted by $x_{-c,jt}$, $\frac{\partial sub}{\partial x_{-c,jt}} = 0$. For the cost-sharing characteristics,

$$\frac{\partial sub}{\partial x_{c,jt}} = -\frac{1-\lambda_{ft}}{\lambda_{ft}} \exp[c_{b,ft} + c_{c,jt}(x_{c,jt})] \gamma_{xc,jt} c_{c,jt}. \tag{30}$$

This suggests a two-step approach for estimating cost parameters. First, use Equation (17) with $\frac{\partial sub}{\partial x_{-c,jt}} = 0$ to recover parameters $\gamma_{x-c,jt}$. Then calculate $c_{b,ft} + c_{c,jt}(x_{c,jt}) = c_{jt} - \gamma'_{x-c,jt} x_{-c,jt}$ and use Equation (17) with Equation (30) to recover parameters $\gamma_{xc,jt}$.

C Computational details

In this Appendix we provide computational details for our counterfactual approach. Let m be the market under consideration with market-level benchmark B_m and plan-level benchmarks B_j taken from the data. Given a counterfactual market-level benchmark B', we calculate

[37] We have explored specifications with linear cost functions and found similar results.

plan-level counterfactual benchmarks $B'_j = B \times \phi_j$. We then calculate counterfactual product characteristics using the demand, supply, and policy function estimates of Section 8. For plan j and characteristic l, we calculate $\widehat{x_{lj}(B'_j)} = x_{lj}(B_j) + \hat{\beta}_{f,l}(B'_j - B_j)$ where $x_{lj}(B_j)$ is the value observed in the data. This ensures that when we input the benchmarks in the data, our routine recreates the product characteristics in the data.

As we observe rebates for each plan, after estimating the marginal cost function, we can estimate ω_{jt} and use Equations (26) and (29) to estimate $\hat{\omega}_{c,jt}$ for each plan. Then, after updating x, we calculate $incr(\cdot)$ using this estimated $\hat{\omega}_{c,jt}$. In other words, given $\widehat{X(B')}$, we counterfactual estimate plan costs $\widehat{c_j(B')} = c_j(\widehat{X(B')})$. With costs in hand, we use Equation (27) to calculate bids and Equation (28) to calculate government payments to firms.

Given non-price product characteristics, costs, and subsidies, we solve firms' price-setting problems by searching for a fixed point in the firms' first-order conditions. This fixed point search is the G algorithm of Section 2. In general, multiple equilibria in the price-setting game are possible. Assumption PRS, and in particular the Assumption ES of Bajari et al. (2007) interpretation, provides a way forward. We use observed actions as a basis for calculating counterfactual equilibria by discarding equilibria inconsistent with our data in the sense that a small change in the benchmark generates a larger change in equilibrium outcomes than is observed in the data. In practice, this does not bind in our setting as we do not find evidence of multiple equilibria in the data nor in the counterfactual simulations.

With equilibrium vectors in hand, we calculate CS and $GovExp$ using Equations (10) and (18), respectively. These functions are non-linear, and so we implement the government spending constraint with a penalty function and address the possibility of multiple local maxima with a multi-start procedure. We restrict the set of counterfactual benchmarks we consider to the range of benchmarks we observe in the data, though in practice this constraint does not bind.

Finally, we obtain substantial computational efficiency by noting that our problem is separable: choices in one market do not affect other markets. We proceed by solving equilibria for each market on a grid of benchmarks in a first stage and then evaluating candidate benchmark schedules using grid interpolation. Both of these steps benefit from parallel computing.

D A Monte Carlo Analysis of our Counterfactual Approach

In Section 2, we describe a theoretical approach for appoximately counterfacual equilibrium outcomes using policy function estimation, potentially augmented with a first-order condition solver. In our MA application, we predict changes in the plan characteristics with estimated policy functions and then solve for the equilibrium in prices—we detail computational specifics in Section C. An obvious question is: How well does this approach work in practice? In this Appendix, we explore the performance of our approach. We write a simplified model and simulate market-level data (where markets differ in part by the level of the subsidy offered to firms) by solving for equilibria explicitly.[38] We use the results of these simulations as the bases for estimating policy functions and approximating outcomes when the subsidies change. We then compare the exact solution to our estimated solution both in terms of the firms' actions (the objects being predicted) and in terms of consumer welfare (the object of interest). We also explore the gains offered by our 'augmented' approach by comparing the approximation error when prices and product characteristics are *all* approximated by policy functions and the error when *only* product characteristics are approximated by policy functions and prices are solved through first-order conditions taking those characteristics as given.

To focus our attention on the approximation error introduced by our technique and to ensure tractability we simplify the model of Sections 5 and 6. Markets are denoted by m. Each market has a unit measure of consumers, denoted by i, and F firms denoted by f. Each firm is present in each market, and offers a constant number of products J, denoted by j. Each product consists of a price p_{jm} (which varies by market) and an $X \times 1$ vector of non-price characteristics δ_{jm}. The choice-specific utility obtained by consumer i when purchasing product j is

$$u_{ijm} = \alpha_i p_{jm}^2 + \beta_i' \delta_{jm} + \epsilon_{ijm} \tag{31}$$

where α_i is the price sensitivity of consumer i, β_i is i's $X \times 1$ vector of preferences for non-price product characteristics, and ϵ_{ijm} is consumer i's idiosyncratic unobservable preference for product j, assumed to be i.i.d. Type-I Extreme Value.

[38] Julia code that implements the exercise outlined in this Appendix and generates the accompanying exhibits can be downloaded at `https://keatonmiller.org/s/mc_release.zip`.

Marginal costs are assumed to be constant at the product level and given by

$$c_{jm} = \exp(-0.3 + \gamma' \delta_{jm}^2 + \nu_f + \omega_m) \tag{32}$$

where δ_{jm}^2 is element-wise squaring of product characteristics, γ is a $X \times 1$ vector of per-characteristic costs, where each component is drawn from i.i.d. $N(0.1, 0.05)$, ν_f is a firm-specific cost shock that is constant across markets, drawn from i.i.d. $N(0, 0.01)$, and ω_m is a market-specific cost shock that is constant across firms, drawn from i.i.d. $N(0, 0.1)$.

Firms receive a subsidy payment b_m from the government and make decisions at the market level. As decisions made in each market are independent, the firm's problem is

$$\max_{p_{fm}, \delta_{fm}} \pi_{fm} = \sum_{j=1}^{J} \int_i (p_{jm} - c_{jm}(\delta_{jm}) + b_{mt})\, s_{ijm}(p_m, \delta_m) di \tag{33}$$

where p_{fm} and δ_{fm} represent the vectors of prices and product characteristics for the firm in that market and s_{ijm} is the probability that consumer i purchases product j as a function of all of the prices and product characteristics in the market. As ϵ_{ijm} is Type-I Extreme Value, s_{ijm} takes a logit form.

Equilibrium in m is a vector (p, δ) for all firms such that each firm's choices solve Equation (33) when taking the competitors' choices and the subsidy level as given. While this model abstracts from common empirical issues, it is a framework which allows us to explore the performance of our approximation approach not merely as a function of the size of the dataset (here represented by the number of markets simulated), but also as a function of the number of firms in each market, the number of products offered by each firm, and the number of non-price product characteristics. Each of these factors can potentially affect the structure of the equilibrium, and in particular may affect the shape of the response of firms to changes in the subsidy.[39]

The existence of multiple products and product characteristics raises the potential for both multiple equilibria and for "trivial" equilibria, in which firms offer identical products. We address this issue through the distribution of consumer preferences: if X is the number of product characteristics, we define X consumer types with equal proportions among the consumer population. Each consumer type n has a strong preference for it's 'own' characteristic

[39] We thank an anonymous referee for this point.

and a weak preference for 'other' characteristics per

$$\beta_{nd} \sim \begin{cases} N(2.5, 0.1) & \text{if } n = x \\ N(0.1, 0.01) & \text{if } n \neq x \end{cases} \tag{34}$$

where β_{nx} is the xth element of the preference vector for consumers of type n. We set $\alpha_n = 2.5$ for all consumers. These choices ensure that when $J \leq X$, there exists an equilibrium in which firms' strategies involve products that are differentiated and specialized. Each product features a high value of a single product characteristic and low values of other product characteristics.

We solve for equilibria by iterating over best response functions. We generate data by giving each market a subsidy ranging from 0 to 1 with equal spacing. We then test the performance of our approach by considering a counterfactual in which the order of subsidies is reversed, so that the market which received the highest subsidy in the first period now receives the lowest and vice versa.

We implement our approach using the generated data. Specifically, we use the equilibrium outcomes to estimate demand and invert the first-order conditions for price-setting to recover an estimate of marginal costs $\hat{c}_{jm}$. We regress this estimate on δ_{jm} and firm and market fixed effects to recover the cost parameters.[40] We then estimate policy functions for each product characteristic. Let δ_{fjx} be the xth product characteristic of product j for firm f. We fit

$$\delta_{fjx} = \theta_{jx} b_m + FE_f + \epsilon_{fjx} \tag{35}$$

where θ_{jx} is the parameter of interest, FE_f is fixed effects, and ϵ_{fjx} is an error term.

We approximate non-price characteristics in the counterfactual by using the estimated $\hat{\theta}_{jx}$ with

$$\hat{\delta}^{ctf}_{fjx} = (b^{ctf}_m - b_m)\hat{\theta}_{jx} + \delta_{fjx} \tag{36}$$

where the ctf superscript refers to counterfactual policies and actions. We calculate marginal costs at these estimated characteristics using our estimates of the cost parameters. We explore the performance of our procedure over two cases. First, we apply this policy function

[40] Note that since there are no product-level unobservables in this specification, this procedure recovers exact parameters.

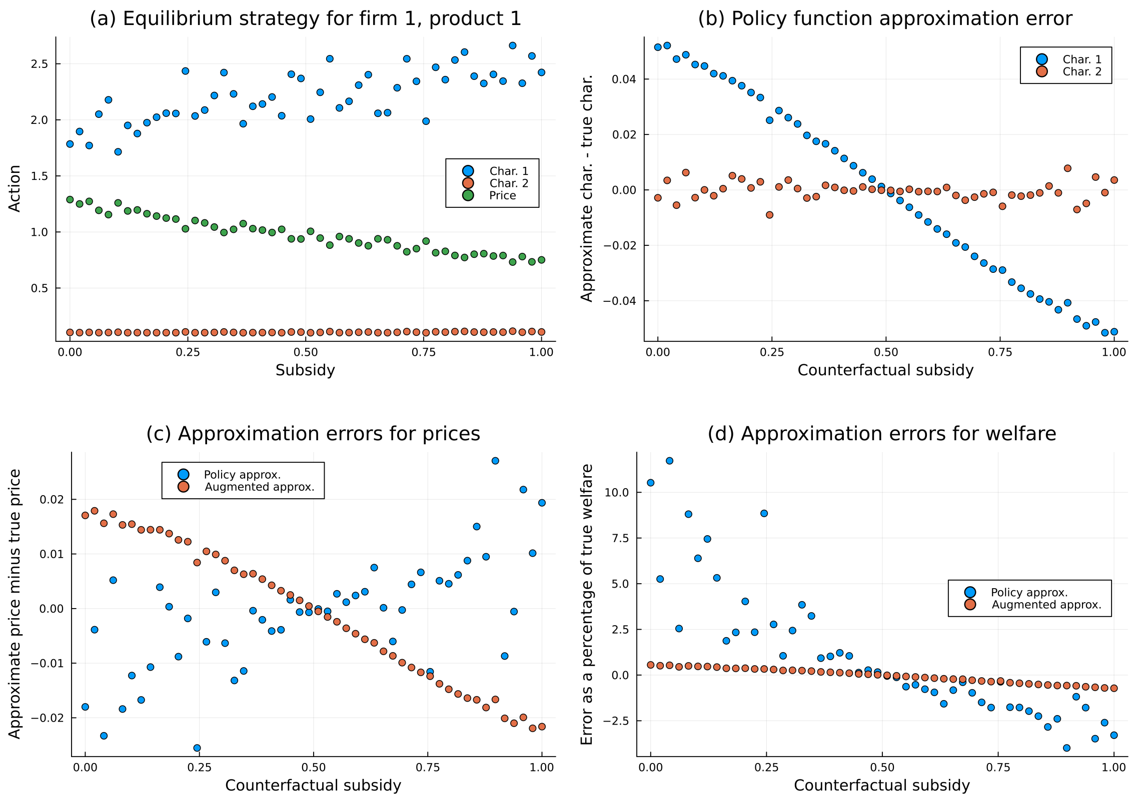
approach to both prices and product characteristics. Second, we apply policy function estimation to only the product characteristics and then solve for an equilibrium in prices with marginal costs and product characteristics taken as given.

Figure D.1 illustrates the output of this procedure for a 50-market run. Each market has four firms, each of which offers two products with two non-price characteristics. There are two types of consumers. Graph (a) illustrates the product design for firm 1's first product in each market; this product is 'designed' to attract consumers of type 1. A higher subsidy leads to lower prices and an increased level of characteristic 1, though there is some variation due to market-specific cost shocks. The other three graphs illustrate approximation errors when computing counterfactual equilibria. Graph (b) illustrates the error coming from policy function approximation: the reported values are the difference between the approximated characteristics in each market and the true counterfactual characteristics. The noise observed in graph (a) generates some attenuation in the estimate of θ_{11}, which means that the policy function slightly overestimates (underestimates) the value of the characteristic when the counterfactual subsidy is low (high). Graph (c) illustrates the difference between approximated prices and true prices under two approaches. First, we consider using policy functions for prices. A similar phenomenon happens here. The policy function underestimates (overestimates) prices when the counterfactual subsidy is low (high). However, the augmented approach slightly overestimates (underestimates) prices when the counterfactual subsidy is low (high).

The result of these patterns in approximation errors is apparent in graph (d), which illustrates the error in the consumer welfare in the market as a percentage of the true value. Note that when using the policy function approach alone, the patterns of errors in graphs (b) and (c) are 'additive:' when the counterfactual subsidy is low (high), the product is both too generous (not generous enough) and too cheap (too expensive) relative to the true counterfactual solution, implying that welfare is substantially higher (lower) than the true outcome. In contrast, the augmented approach for prices generates much smaller errors in consumer welfare, as the error in prices works in the opposite direction as the error in product characteristics.

To evaluate the performance of the approximation approach in a systematic way, we define two statistics. First, we define the mean approximation error in action space by calculating the Euclidean distance between the approximate and exact action vectors and dividing it

Figure D.1: Example Monte Carlo simulation output



Notes: These graphs illustrate outcomes from a sample Monte Carlo simulation run. For this run, 50 markets were simulated, each with four firms (identical across markets). Each firm offered two products, each of which had two non-price product characteristics. There were two consumer types. The top-left graph illustrates the policy function for firm 1 product 1 across markets, including the price decisions and both product characteristics. The top-right graph compares the exact policy function to the approximated policy function under counterfactual subsidies. The bottom-left graph compares prices when product characteristics are exact and approximated. The bottom-right graph compares the consumer welfare in each market under the exact and approximated solutions. See text for details and discussion. Code to replicate these figures can be found at `https://keatonmiller.org/s/mc_release.zip`.

by the magnitude of the true action vector. If $D(x, y)$ is the distance between x and y, and a_m is the vector of prices and product characteristics for all products in market m, we define $Err^a = \frac{1}{M}\sum_m D(a_m^{Exact}, a_m^{Approx})/||a_m^{Exact}||$. Second, we calculate consumer welfare per Equation (10) under both the exact solution and the approximated solution and define the mean absolute logarithm error as $Err^{CW} = \frac{1}{M}\sum_m |\log(CW_m^{Exact}) - \log(CW_m^{Approx})|$. Both of these measures are strictly positive and are designed to capture the amount of error relative to the size of the approximated object. Finally, we calculate these for two approximations, one in which policy functions are used for prices and product characteristics (denoted with the subscript *pol*) and one in which we solve for prices taking approximate product characteristics as given (subscript aug).

The results of our simulations are presented in Table D.1. We explore the performance of our approach by varying the complexity of the market both in terms of the number of firms present, as well as the complexity of the product offerings of those firms. We also vary the number of markets simulated to evaluate the performance with datasets of differing sizes. Across scenarios, the performance corresponds with sensible priors. Increasing the complexity of markets tends to increase the mean simulation error when the number of markets is low. Increasing the number of markets tends to decrease the error. Even in the worst case, however, the error in the action vector remains under 2%, and the error in the consumer welfare is less than 1% when using the augmented approach, though it is higher when using policy functions to approximate prices. This is consistent with Figure D.1 – simulation error in the policy function for characteristics is magnified when prices are approximated by policy functions and offset when solving for prices taking approximate characteristics as given.

Table D.1: Monte Carlo simulation results

(a) Approximation error in action space

Firms per mkt.	Prods. per firm	Chars. per prod.	50 Markets Err^a_{pol}	50 Markets Err^a_{aug}	100 Markets Err^a_{pol}	100 Markets Err^a_{aug}	200 Markets Err^a_{pol}	200 Markets Err^a_{aug}
1	1	1	0.0067	0.0051	0.0109	0.0099	0.0127	0.0113
1	2	2	0.0073	0.0056	0.0118	0.0107	0.0136	0.0121
1	4	4	0.0083	0.0071	0.0134	0.0126	0.0153	0.0143
2	1	1	0.0062	0.0062	0.0126	0.0132	0.0144	0.0151
2	2	2	0.0063	0.0064	0.0129	0.0135	0.0147	0.0154
2	4	4	0.0070	0.0072	0.0141	0.0147	0.0158	0.0166
4	1	1	0.0069	0.0065	0.0135	0.0139	0.0152	0.0157
4	2	2	0.0070	0.0067	0.0138	0.0142	0.0155	0.0160
4	4	4	0.0079	0.0079	0.0152	0.0159	0.0169	0.0177

(b) Absolute logarithm error in consumer welfare

Firms per mkt.	Prods. per firm	Chars. per prod.	50 Markets Err^{CW}_{pol}	50 Markets Err^{CW}_{aug}	100 Markets Err^{CW}_{pol}	100 Markets Err^{CW}_{aug}	200 Markets Err^{CW}_{pol}	200 Markets Err^{CW}_{aug}
1	1	1	0.0385	0.0018	0.0486	0.0035	0.0580	0.0040
1	2	2	0.0416	0.0019	0.0525	0.0037	0.0627	0.0043
1	4	4	0.0426	0.0021	0.0540	0.0041	0.0644	0.0046
2	1	1	0.0164	0.0025	0.0336	0.0055	0.0368	0.0063
2	2	2	0.0167	0.0025	0.0339	0.0055	0.0370	0.0063
2	4	4	0.0166	0.0026	0.0343	0.0058	0.0373	0.0065
4	1	1	0.0143	0.0018	0.0275	0.0040	0.0295	0.0046
4	2	2	0.0144	0.0019	0.0276	0.0041	0.0295	0.0046
4	4	4	0.0146	0.0020	0.0282	0.0044	0.0299	0.0049

Notes: In all simulations, the number of consumer types is equal to the number of non-price product characteristics. Err^{δ} and Err^{CW} are defined in the text; reported metrics are means across all markets in the simulation. For ease of comparison, preference and firm-level cost draws are identical across rows, while market-level cost draws are constant across columns. Code to replicate this table can be downloaded at `https://keatonmiller.org/s/mc_release.zip`.

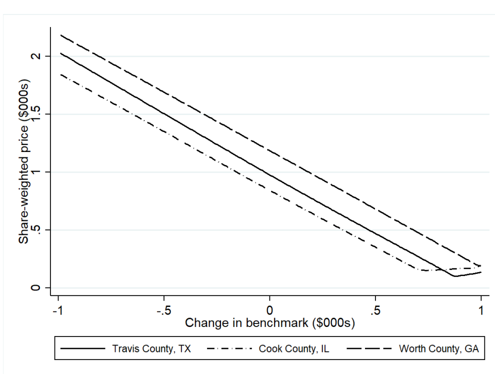
E The Non-Local Behavior of Surplus and Expenditures

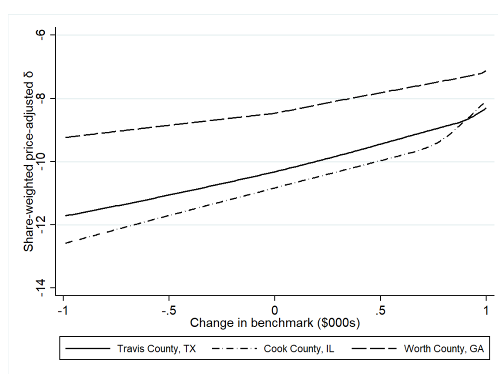
The results in Section 9.3 show that the derivatives of the CS and $GovExp$ functions at the 2017 policy point to the direction of the optimal policy. However, the derivatives do not provide sufficient information to calculate the optimal policy. Though the policy function approximations we use are linear, the pricing behavior and share functions are not and so we should expect the CS and $GovExp$ functions to be non-linear as well. In this Appendix, we explore the non-local behavior of CS and $GovExp$ through illustrated examples.

Figure E.1 illustrates components of the CS function for Travis County, TX (containing Austin), Cook County, IL (containing Chicago), and Worth County, GA (a rural county near Albany), in 2017. We chose these counties due to their different sizes and the typical nature of their counterfactual equilibria. The left-hand graph depicts the share-weighted premium and shows non-monotonicity as the benchmark increases. As the benchmark increases, prices near the zero lower bound. The right-hand graph illustrates the share-weighted $\delta'_j \equiv \delta_j - \alpha_0 p_j$ (i.e. the net utility impact of product characteristics) as a function of the benchmark. All three counties show increases as the benchmark increases. The slope, however, is higher in Cook County and Travis County than in Worth County due in part to differences in the firms present and their associated policy functions, and in part to the amount of competition.

Figure E.2 illustrates components of $GovExp_m$ for the same counties. The left-hand graph depicts the share-weighted plan cost. Cook County sees the highest increases. The middle graph shows the share-weighted average bid. Cook County's bids increase nearly

Figure E.1: Prices and product characteristics under counterfactual benchmarks, selected counties

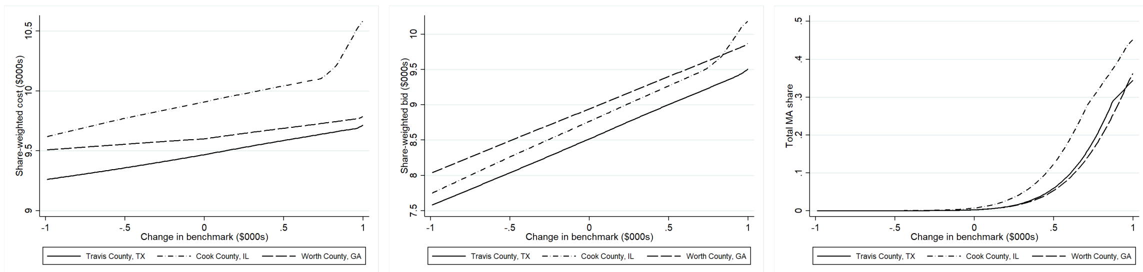




Notes: The horizontal axis is the change in the benchmark relative to 2017. The lines illustrate share-weighted averages across plans.

1-for-1 with an increase in the benchmark, whereas the average bids in Travis and Worth Counties increase with a slightly shallower slope. The right-hand graph illustrates the total share of MA (relative to TM). Cook County's share increases the fastest.

Figure E.2: Average plan costs and bids and MA share under counterfactual benchmarks, selected counties



Notes: The horizontal axis is the change in the benchmark relative to the 2017 level.

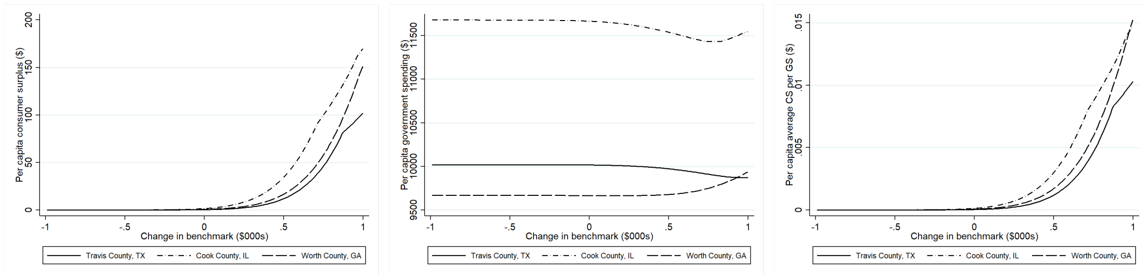
Figure E.3 combines these components into the CS_m and $GovExp_m$ functions. The first graph shows per-capita consumer surplus. Under the current policy, the three counties receive similar surplus. As the benchmark in each county is increased, the average surplus in Cook County grows faster than the others. The second graph illustrates per-capita government expenditures. This graph illustrates the potential gains noted in the previous section. Cook and Travis Counties have a flat or even decreasing level of government expenditures for modest increases in the benchmark rate. These graphs suggest that significant gains are possible in some markets simply by incentivizing switches from TM to MA.

The final graph of Figure E.3 combines the two functions to show the average MA surplus delivered to consumers per dollar spent by the government on the Medicare program. The slope of this line is related to the marginal impact of spending an extra dollar in a particular county through the MA benchmark mechanism, which is the key margin explored by the constrained maximization algorithm of our optimal policy search. Over small increases in the benchmark, Cook County experiences the largest gains in surplus per expenditures.

F Other counterfactual results and robustness

In this appendix, we briefly describe a number of other results. We begin with investigating the robustness of our counterfactual equilibria, as while our approach finds equilibria in premiums taking product characteristics as given, the first-order conditions with respect to other characteristics may not be satisfied with equality. To investigate our counterfactual

Figure E.3: Per-capita consumer surplus and government expenditures under counterfactual benchmarks, selected counties



Notes: The horizontal axis is the change in the benchmark relative to the 2017 level.

equilibrium, we calculate the change in variable profits firms would obtain by deviating from the predicted values for each product characteristic at the optimal policy by 1% of the sample mean, and subtract it from the change in profits for deviations from the observed 2017 characteristics. The difference represents error introduced by our counterfactual approach.

The results are reported in Appendix Table F.1. Across all markets, deviations in the unobservable characteristic pose the largest potential profit: a 1% increase in ξ leads to a mean decease in profits of -0.0213%. For all other characteristics the mean profitability change is less than 0.005% in absolute value, indicating that our approximation error is second order. The Monte Carlo results in Appendix D suggest that approximation error increases as the magnitude of the benchmark change increases. The fourth and fifth sets of rows of Appendix Table F.1 split markets into those with changes in the benchmark of more and less than $1,000. As expected, the 60 markets with changes greater than $1,000 have more profitable deviations – for example the mean change in variable profit from a 1% change in hospital copays is 0.0161%.

To provide context for these changes, the bottom row of Appendix Table F.1 reports the mean change in profits at the 2017 benchmarks (i.e. the mean gradient of the profit function). Profits are most sensitive to changes in ξ. At the 2017 benchmarks, a 1% increase in ξ results in a 0.27% decrease in profits on average. These changes may be driven in part by adjustment costs which we do not model. For example, if ξ is primarily determined by hospital and provider networks, it may be the case that contract negotiations between insurers and hospitals takes place at a different frequency than the MA benchmark update interval.[41]

[41] We thank an anonymous referee for this point.

The optimal subsidy schedule may have political economy implications if the changes result in a large-scale redistribution of government expenditure dollars and consumer welfare across states. Indeed, past changes to the MA payment formula have likely resulted from political considerations during the legislative process (Berenson, 2008). To explore these issues, we summarize the total consumer surplus and government expenditures by state in Appendix Table F.2. Of the 41 states (plus Washington D.C.) included in the 2017 MCBS, 29 receive increases to aggregate consumer welfare. The results suggest the optimal policy does not split cleanly along political divisions.

Appendix Table F.3 reports firm-level market shares and total variable profits under the 2017 policy, the optimal policy, and other policies explored in Figure 9. The optimal policy increases aggregate insurer variable profits from \$3.10 billion to \$5.15 billion. The absolute shares of each of the five largest insurers increase, though Humana and UnitedHealth Group lose share relative to others. The linear rule results in \$5.04 billion in aggregate variable profits. The 'minimize spending without decreasing benchmarks' specification results in \$4.05 billion in aggregate variable profits. Again, the absolute shares of the largest insurers increases.

Table F.1: The profitability of a 1% deviation from predicted product characteristics at the optimal policy

	An 1% increase in the characteristic at the optimal policy...								
	Enhcd. drug	Dental	Vision	Hearing	ξ	Part B reduc.	Deduct.	Prim. copay	Hosp. copay
All markets									
Mean	-0.0020	-0.0014	-0.0009	-0.0025	-0.0213	-0.0001	0.0004	0.0017	0.0036
25th Percentile	-0.0017	-0.0012	-0.0008	-0.0021	-0.0148	-0.0002	-0.0005	-0.0021	-0.0045
Median	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0000
75th Percentile	0.0004	0.0002	0.0001	0.0005	0.0083	0.0002	0.0005	0.0020	0.0043
270 markets with increases									
Mean	-0.0046	-0.0032	-0.0021	-0.0056	-0.0417	-0.0004	0.0012	0.0051	0.0107
25th Percentile	-0.0062	-0.0043	-0.0028	-0.0077	-0.0678	-0.0007	-0.0001	-0.0005	-0.0011
Median	-0.0003	-0.0002	-0.0002	-0.0003	-0.0017	-0.0000	0.0000	0.0001	0.0003
75th Percentile	-0.0000	-0.0000	-0.0000	-0.0000	0.0018	0.0000	0.0020	0.0086	0.0183
169 markets with decreases									
Mean	0.0032	0.0023	0.0015	0.0040	0.0210	0.0004	-0.0012	-0.0052	-0.0109
25th Percentile	0.0001	0.0001	0.0000	0.0001	0.0002	0.0000	-0.0011	-0.0047	-0.0100
Median	0.0006	0.0004	0.0003	0.0007	0.0021	0.0000	-0.0001	-0.0006	-0.0012
75th Percentile	0.0035	0.0025	0.0016	0.0042	0.0178	0.0004	-0.0000	-0.0000	-0.0001
64 markets with changes greater than $1,000									
Mean	-0.0077	-0.0054	-0.0035	-0.0095	-0.0729	-0.0006	0.0018	0.0076	0.0161
25th Percentile	-0.0166	-0.0113	-0.0072	-0.0210	-0.2425	-0.0024	-0.0024	-0.0104	-0.0225
Median	-0.0014	-0.0010	-0.0007	-0.0014	-0.0008	-0.0002	0.0007	0.0030	0.0063
75th Percentile	0.0005	0.0002	0.0000	0.0009	0.0886	0.0008	0.0072	0.0307	0.0645
375 markets with changes less than $1,000									
Mean	-0.0011	-0.0007	-0.0005	-0.0013	-0.0125	-0.0001	0.0002	0.0007	0.0015
25th Percentile	-0.0007	-0.0005	-0.0004	-0.0008	-0.0066	-0.0001	-0.0004	-0.0019	-0.0040
Median	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0001
75th Percentile	0.0003	0.0002	0.0001	0.0005	0.0058	0.0001	0.0002	0.0008	0.0017
Mean base deviation	-0.0027	-0.0013	-0.0005	-0.0041	-0.2732	-0.0073	-0.0216	-0.0932	-0.1976

...leads to a X% change in variable profit relative to the same deviation at the 2017 policy.

Notes: This table investigates the approximation error introduced by our policy function approach. For each plan, we calculate the percentage change in the variable profit earned by the firm that offers the plan in response to an increase of 1% in each of the product characteristics starting from the prices and characteristics at the optimal subsidy schedule. We update marginal costs and hold prices and all other product characteristics fixed for all plans in the market. To isolate the effect of our approximation approach, we calculate the percentage difference both at the optimal policy and at the 2017 policy, and subtract the two. For comparison, the last line of the table reports the mean deviation at the 2017 policy.

Table F.2: State-level surplus and expenditures changes from 2017 policy to optimal benchmark schedule

	# counties	Sum(weight)	Consumer surplus ($ M)			Government expenditures ($ M)		
State	in sample	(000,000s)	2017	Optimal	%Δ	2017	Optimal	%Δ
Alabama	14	11.61	78.2	44.8	-42.7	11,379	11,278	-0.89
Arizona	5	11.79	142.5	238.9	67.7	11,027	11,066	0.35
Arkansas	3	5.27	56.1	9.1	-83.7	4,806	4,693	-2.35
California	17	34.05	704.1	1,674.3	137.8	39,379	39,858	1.22
Colorado	8	3.87	68.9	157.3	128.4	3,729	3,784	1.46
Connecticut	4	6.18	63.6	355.9	459.4	6,876	7,028	2.20
District of Columbia	1	1.42	1.6	11.6	607.0	1,753	1,745	-0.42
Florida	26	36.95	587.5	675.2	14.9	41,228	41,299	0.17
Georgia	18	11.46	70.5	145.3	105.9	11,146	11,166	0.17
Illinois	15	13.49	69.5	223.2	221.0	14,735	14,813	0.53
Indiana	3	0.66	2.4	3.0	27.0	639	636	-0.47
Iowa	4	2.68	3.7	1.4	-61.9	2,317	2,312	-0.19
Kansas	2	3.25	13.4	32.1	138.9	3,078	3,093	0.48
Kentucky	12	7.52	45.0	94.0	108.6	7,634	7,648	0.19
Louisiana	7	4.01	89.8	63.2	-29.7	4,393	4,348	-1.03
Maryland	8	4.82	3.0	123.6	3,986.9	5,577	5,555	-0.39
Massachusetts	9	8.49	47.0	665.5	1,315.4	10,158	10,285	1.25
Michigan	29	22.85	179.3	366.3	104.3	24,463	24,429	-0.14
Minnesota	14	8.22	35.7	57.0	59.5	8,140	8,116	-0.29
Missouri	15	8.47	83.2	62.6	-24.7	8,468	8,400	-0.80
Nebraska	6	2.53	9.3	56.7	513.2	2,642	2,623	-0.71
Nevada	2	6.21	116.1	116.7	0.5	7,187	7,188	0.01
New Hampshire	16	16.65	76.6	814.2	962.3	20,143	20,302	0.79
New Jersey	6	9.50	126.7	124.1	-2.1	7,176	7,172	-0.06
New Mexico	26	26.93	389.0	1,078.4	177.2	30,768	30,242	-1.71
New York	18	18.67	230.8	219.3	-5.0	19,281	18,952	-1.70
North Carolina	24	17.45	172.2	521.3	202.8	17,888	18,218	1.85
Ohio	6	3.81	6.6	2.8	-57.5	3,649	3,640	-0.25
Oklahoma	1	0.02	0.1	0.1	-47.5	14	14	-0.69
Pennsylvania	25	19.55	309.3	784.0	153.5	20,556	20,866	1.51
Rhode Island	7	3.06	9.8	6.2	-36.7	2,908	2,897	-0.38
South Carolina	1	0.02	0.0	0.1	250.7	19	19	0.33
South Dakota	12	10.36	88.8	81.5	-8.2	9,950	9,902	-0.48
Texas	37	29.18	248.4	777.1	212.9	31,570	31,758	0.60
Utah	1	0.01	0.2	0.2	-0.4	13	13	-0.01
Vermont	1	0.02	0.0	0.2	380.7	16	16	0.38
Virginia	11	7.34	35.7	41.8	17.0	7,168	7,170	0.04
Washington	8	21.19	181.4	120.5	-33.6	18,643	18,445	-1.07
West Virginia	5	3.69	23.1	54.3	135.7	3,502	3,504	0.06
Wisconsin	11	15.88	246.9	33.2	-86.5	14,552	14,090	-3.18
Wyoming	1	0.83	1.3	7.0	424.4	685	688	0.40
Total	439	420	4,618	9,844	113.2	439,254	439,272	0.00

Notes: The MCBS uses a sample of counties and weights observations to be nationally representative; the first column reports the number of counties included in the MCBS in each state, and the second column reports the total MCBS sample weight in the state. Consumer surplus is calculated via Equation (10). Government expenditures include expenditures on TM and MA and are calculated via Equation (18). Surplus and expenditure statistics are calculated using MCBS sample weights.

Table F.3: Market shares and total variable profits for selected firms under 2017 policy and alternative policies

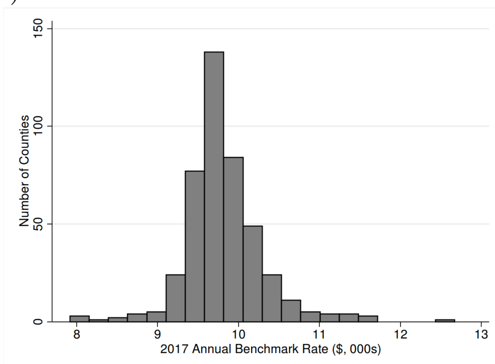
	(1) 2017 policy	(2) Optimal policy	(3) Linear rule	(4) 0.99 CS - 0.01 Var	(5) 0.999 CS - 0.001 Var	(6) Min. *GovExp*	(7) Min exp. w/ floor
Aetna							
Shr. of Medicare benes. (%)	1.16	3.10	3.21	1.01	3.00	1.70	1.95
Shr. of MA enrollees (%)	3.90	7.19	7.67	5.01	6.33	6.23	5.27
Total var. profits ($ bill.)	.117	.396	.415	.127	.384	.216	.234
Blue Cross Blue Shield							
Shr. of Medicare benes. (%)	4.22	7.85	7.34	3.04	8.67	4.08	5.59
Shr. of MA enrollees (%)	14.2	18.2	17.6	15.1	18.3	14.9	15.1
Total var. profits ($ bill.)	.340	.858	.789	.264	.890	.414	.500
Humana							
Shr. of Medicare benes. (%)	5.97	7.79	7.56	4.37	9.71	4.52	6.84
Shr. of MA enrollees (%)	20.0	18.0	18.1	21.8	20.5	16.5	18.5
Total var. profits ($ bill.)	.456	.642	.626	.339	.785	.402	.523
Kaiser Permanente							
Shr. of Medicare benes. (%)	1.83	2.43	2.50	1.05	2.31	2.13	2.16
Shr. of MA enrollees (%)	6.15	5.65	5.98	5.21	4.87	7.78	5.85
Total var. profits ($ bill.)	.342	.505	.580	.177	.463	.426	.433
UnitedHealth Group							
Shr. of Medicare benes. (%)	6.12	6.97	6.47	3.79	7.66	5.47	7.45
Shr. of MA enrollees (%)	20.5	16.2	15.5	18.9	16.1	20.0	20.2
Total var. profits ($ bill.)	.761	.904	.828	.489	.946	.742	.931
All Others							
Shr. of Medicare benes. (%)	10.5	15.0	14.7	6.82	16.1	9.47	13.0
Shr. of MA enrollees (%)	35.2	34.8	35.2	34.0	33.9	34.6	35.1
Total var. profits ($ bill.)	1.08	1.85	1.80	.702	1.84	1.12	1.43
Total							
Shr. of Medicare benes. (%)	29.8	43.2	41.8	20.1	47.4	27.4	36.9
Total var. profits ($ bill.)	3.10	5.15	5.04	2.10	5.31	3.32	4.05

Notes: Variable profits are computed via Equation (12) using equilibrium prices and estimated marginal costs adjusted for changes in product characteristics under alternative policies. The entries for Blue Cross Blue Shield include all members of the Blue Cross Blue Shield Association. All statistics are weighted by the MCBS sample weights.

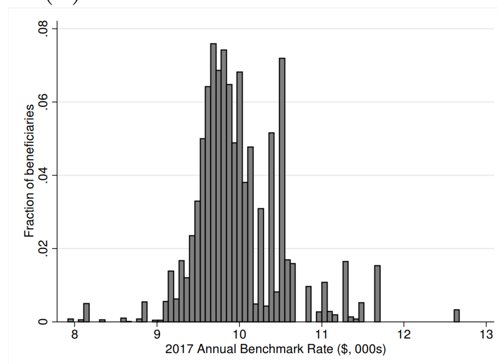
G Additional Tables and Figures

Figure G.1: Medicare Advantage Benchmark Distribution, 2017

(a) Benchmarks across counties



(b) Benchmarks across beneficiaries



Note: Includes only those counties included in the 2017 Medicare Current Beneficiary Survey.

Table G.1: Marginal cost parameter estimates using pricing first order conditions alone

	$\ln(c_j)$					
Variable	Aetna	BCBS	Humana	Kaiser	UHG	Other
Cost-sharing characteristics						
Part B reduction (per $1000)	-0.0001	-0.0000	0.0000	0.0000	0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Deductible (per $1000)	-0.0176	-0.0208	-0.0247	0.0000	-0.0095	-0.0263
	(0.0028)	(0.0020)	(0.0015)	(0.0000)	(0.0034)	(0.0011)
Primary care copay	-0.0010	-0.0008	-0.0010	0.0019	-0.0000	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0000)
Specialist copay	-0.0001	-0.0000	-0.0004	-0.0014	0.0002	-0.0011
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0000)
Hospital copay (per $1000)	-0.0246	-0.0211	-0.0035	-0.0266	-0.0260	-0.0177
	(0.0018)	(0.0015)	(0.0012)	(0.0029)	(0.0016)	(0.0006)
Supplemental coverage characteristics						
Prescription drug	0.0420	0.0230	0.0651	N/A	-0.0011	0.0279
	(0.0032)	(0.0019)	(0.0016)	(0.0000)	(0.0023)	(0.0011)
Enhanced prescription drug	0.0128	0.0172	0.0090	0.0000	0.0292	0.0171
	(0.0024)	(0.0015)	(0.0013)	(0.0000)	(0.0022)	(0.0010)
Dental	0.0120	0.0190	0.0098	0.0459	0.0205	0.0196
	(0.0018)	(0.0013)	(0.0012)	(0.0058)	(0.0013)	(0.0009)
Vision	0.0034	0.0309	0.0169	0.0225	-0.0262	0.0012
	(0.0175)	(0.0021)	(0.0012)	(0.0162)	(0.0097)	(0.0016)
Hearing	0.0445	0.0361	0.0239	0.0000	0.0046	0.0285
	(0.0086)	(0.0017)	(0.0011)	(0.0000)	(0.0053)	(0.0010)
Demand unobservable (ξ_j)	0.0153	0.0135	0.0116	0.0109	0.0089	0.0120
	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0001)
Fixed effects	Star rating, firm, county, year					
Observations	64,538					
R^2	0.8740					

Notes: Observations are county-year-plans. Estimates are formed via OLS. All Kaiser plans in our sample had prescription drug coverage. Robust standard errors are in parentheses.

Table G.2: Policy functions for product characteristics by census region

	(1) Part B reduction	(2) Deduct-ible	(3) Prim. care copay	(4) Hospital copay	(5) Enhanced drug	(6) Dental	(7) Vision	(8) Hearing	(9) ξ
Aetna									
Northeast	0.0141	0.0390	-0.9575	-0.1569	0.1514	0.0641	-0.0136	0.0693	0.6849
	(0.0029)	(0.0069)	(0.1851)	(0.0126)	(0.0111)	(0.0102)	(0.0050)	(0.0095)	(0.0987)
Midwest	0.0139	0.0373	-1.1700	-0.1760	0.1596	0.0439	-0.0134	0.0690	0.5873
	(0.0029)	(0.0069)	(0.1875)	(0.0127)	(0.0110)	(0.0104)	(0.0049)	(0.0094)	(0.0994)
South	0.0121	0.0421	-1.0256	-0.1577	0.1585	0.0602	-0.0148	0.0677	0.5443
	(0.0029)	(0.0069)	(0.1862)	(0.0126)	(0.0110)	(0.0102)	(0.0050)	(0.0095)	(0.0989)
West	0.0131	0.0373	-1.2843	-0.1773	0.1558	0.0639	-0.0129	0.0704	0.3595
	(0.0029)	(0.0070)	(0.1925)	(0.0128)	(0.0111)	(0.0104)	(0.0050)	(0.0095)	(0.1010)
BCBS									
Northeast	0.0145	0.0417	-0.9800	-0.1862	0.1446	0.0617	-0.0167	0.0585	0.6407
	(0.0029)	(0.0069)	(0.1845)	(0.0126)	(0.0110)	(0.0102)	(0.0050)	(0.0096)	(0.0985)
Midwest	0.0138	0.0334	-0.9288	-0.1799	0.1453	0.0754	-0.0330	0.0560	0.5706
	(0.0029)	(0.0069)	(0.1849)	(0.0125)	(0.0110)	(0.0101)	(0.0050)	(0.0095)	(0.0982)
South	0.0117	0.0502	-0.9722	-0.1736	0.1502	0.0251	-0.0298	0.0188	0.5542
	(0.0029)	(0.0071)	(0.1865)	(0.0126)	(0.0111)	(0.0103)	(0.0050)	(0.0096)	(0.0990)
West	0.0129	0.0411	-1.2165	-0.1756	0.1415	0.0527	-0.0210	0.0334	0.5842
	(0.0028)	(0.0069)	(0.1868)	(0.0126)	(0.0111)	(0.0102)	(0.0051)	(0.0097)	(0.0988)
Humana									
Northeast	0.0146	0.0384	-1.0963	-0.1323	0.1345	0.0784	-0.0254	0.0119	0.5822
	(0.0029)	(0.0070)	(0.1870)	(0.0128)	(0.0113)	(0.0104)	(0.0053)	(0.0098)	(0.0994)
Midwest	0.0146	0.0493	-1.3379	-0.1497	0.1332	0.0751	-0.0290	0.0109	0.5844
	(0.0029)	(0.0069)	(0.1847)	(0.0126)	(0.0111)	(0.0102)	(0.0051)	(0.0095)	(0.0981)
South	0.0136	0.0430	-1.2088	-0.1655	0.1404	0.0616	-0.0310	-0.0011	0.5524
	(0.0030)	(0.0069)	(0.1852)	(0.0126)	(0.0110)	(0.0102)	(0.0051)	(0.0096)	(0.0985)
West	0.0133	0.0391	-1.2410	-0.1571	0.1477	0.0626	-0.0291	0.0055	0.6271
	(0.0028)	(0.0069)	(0.1857)	(0.0126)	(0.0111)	(0.0102)	(0.0051)	(0.0096)	(0.0986)
Kaiser									
Northeast				*Kaiser had no presence in the Northeast region*					
Midwest	0.0138	0.0364	-1.0883	-0.1550	0.1620	0.0643	-0.0144	0.0641	-0.3968
	(0.0028)	(0.0067)	(0.1790)	(0.0121)	(0.0107)	(0.0098)	(0.0049)	(0.0093)	(0.0953)
South	0.0145	0.0314	-0.8989	-0.1589	0.1722	0.0681	-0.0368	0.0426	0.4715
	(0.0029)	(0.0069)	(0.1992)	(0.0125)	(0.0109)	(0.0104)	(0.0062)	(0.0101)	(0.1006)
West	0.0112	0.0321	-0.3409	-0.1245	0.1706	0.0812	-0.0187	0.0160	0.5463
	(0.0028)	(0.0068)	(0.1822)	(0.0124)	(0.0108)	(0.0100)	(0.0049)	(0.0095)	(0.0968)
UHG									
Northeast	0.0140	0.0345	-1.0138	-0.1491	0.1361	0.0237	-0.0149	0.0717	0.5633
	(0.0029)	(0.0071)	(0.1865)	(0.0127)	(0.0112)	(0.0103)	(0.0051)	(0.0096)	(0.0995)
Midwest	0.0139	0.0294	-1.0690	-0.1390	0.1399	0.0229	-0.0148	0.0715	0.6249
	(0.0029)	(0.0068)	(0.1852)	(0.0125)	(0.0111)	(0.0103)	(0.0050)	(0.0095)	(0.0976)
South	0.0095	0.0313	-1.2485	-0.1406	0.1357	0.0280	-0.0135	0.0676	0.5959
	(0.0029)	(0.0069)	(0.1854)	(0.0127)	(0.0111)	(0.0103)	(0.0050)	(0.0096)	(0.0985)
West	0.0109	0.0318	-1.3601	-0.1595	0.1475	0.0161	-0.0171	0.0611	0.5539
	(0.0028)	(0.0068)	(0.1838)	(0.0126)	(0.0111)	(0.0101)	(0.0050)	(0.0095)	(0.0982)
Other firms									
Northeast	0.0141	0.0401	-1.3511	-0.1969	0.1419	0.0524	-0.0174	0.0547	0.6398
	(0.0029)	(0.0068)	(0.1841)	(0.0125)	(0.0110)	(0.0101)	(0.0050)	(0.0095)	(0.0986)
Midwest	0.0144	0.0405	-1.4959	-0.1984	0.1497	0.0454	-0.0252	0.0522	0.5342
	(0.0029)	(0.0070)	(0.1843)	(0.0126)	(0.0110)	(0.0101)	(0.0051)	(0.0096)	(0.0979)
South	0.0226	0.0340	-1.3084	-0.1809	0.1448	0.0527	-0.0164	0.0486	0.5545
	(0.0029)	(0.0068)	(0.1850)	(0.0126)	(0.0110)	(0.0102)	(0.0050)	(0.0096)	(0.0986)
West	0.0137	0.0310	-1.3560	-0.1809	0.1413	0.0506	-0.0151	0.0521	0.6238
	(0.0029)	(0.0069)	(0.1846)	(0.0126)	(0.0110)	(0.0102)	(0.0050)	(0.0095)	(0.0983)
Observations	37,974	37,974	37,974	37,974	37,974	37,974	37,974	37,974	37,974
R-squared	0.1696	0.0920	0.2518	0.3016	0.0692	0.1799	0.1429	0.3379	0.1039

Notes: This table reports the results of multiple policy function regressions. The dependent variable in each column is the product characteristic. The independent variables include the market-level benchmark (measured in thousands of 2017 dollars) interacted with firm-by-Census-Region indicators, the relevant cost shock, the last period demand unobservable, and market average demographics. Estimates are formed via OLS. No plans in our sample changed basic drug coverage. Robust standard errors are in parentheses.

Table G.3: Mean county characteristics by benchmark quartile, 2017 policy and optimal policy

2017 policy	**0-25th**	**26-50th**	**51-75th**	**76-100th**
Risk-adj. TM costs per capita ($)	9,472	9,830	10,225	11,047
Average risk score	.991	.978	1.00	1.01
Beneficiaries	43,833	43,964	72,933	98,264
Median household income ($)	53,857	52,610	58,314	61,417
Percent in deep poverty, 65+	2.53	2.75	2.50	2.76
Unemployment rate	5.65	5.77	5.33	5.19
Population density (per mi^2)	755	1,280	1,098	1,409
Urban/Rural continuum code	2.90	3.52	2.25	2.25
Resources per 10,000 people				
MDs	21.3	20.0	22.7	24.5
Medicare hospitals	.028	.036	.045	.035
Skilled nursing facilities	.573	.679	.571	.590
Hospice facilities	.132	.190	.113	.101
Medicare hospital readmission rate	17.5	17.3	17.6	18.1
Preventable hospital admission rate	52.0	52.7	51.6	55.0
2017 benchmark ($)	9,318	9,660	9,868	10,421
Number of MA plans	14.8	13.1	17.0	17.5
Number of MA firms	7.2	7.0	9.0	9.3
Observations	110	110	110	109
Optimal policy	**0-25th**	**26-50th**	**51-75th**	**76-100th**
Risk-adj. TM costs per capita ($)	9,027	9,755	10,309	11,486
Average risk score	.992	.995	.989	1.001
Beneficiaries	31,934	50,225	72,406	104,483
Median household income ($)	49,631	54,294	55,773	66,548
Percent in deep poverty, 65+	2.74	2.48	2.62	2.70
Unemployment rate	5.99	5.47	5.44	5.04
Population density (per mi^2)	323	815	761	2,653
Urban/Rural continuum code	3.45	2.57	2.55	2.35
Resources per 10,000 people				
MDs	16.6	22.1	22.6	27.3
Medicare hospitals	.029	.032	.046	.037
Skilled nursing facilities	.614	.555	.612	.633
Hospice facilities	.154	.143	.119	.120
Medicare hospital readmission rate	17.2	17.5	17.6	18.1
Preventable hospital admission rate	51.7	59.9	52.8	52.7
Optimal benchmark ($)	8,998	9,681	10,220	11,178
Number of MA plans	12.6	16.2	15.9	17.6
Number of MA firms	6.5	8.9	8.6	8.5
Observations	110	110	110	109

Notes: This table reports county characteristics from CMS, Census, and Area Health Resource File data across benchmark quartiles. The top panel defines benchmark quartiles according to the 2017 policy. The bottom panel defines quartiles according to the optimal policy.