MULTIDIMENSIONAL SCREENING FOR QUALITY WITH AN APPLICATION TO HEALTH INSURANCE

By

Hector Chade, Victoria Marone, Amanda Starc

and Jeroen Swinkels

June 2026

COWLES FOUNDATION DISCUSSION PAPER NO. 2541

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Multidimensional Screening for Quality with an Application to Health Insurance

Hector Chade, Victoria Marone, Amanda Starc, and Jeroen Swinkels*

June 30, 2026

Abstract

We analyze a multidimensional screening model in which a principal offers a menu of quality-price pairs to a consumer with multiple dimensions of private information and a quasilinear utility function. We derive necessary conditions for optimality, and use them to provide insight into optimal exclusion, positive trade, and screening. We then recast the problem in terms of incremental quality levels and prices, the so-called demand-profile approach (*DPA*). Under *DPA*, the problem decouples across increments and can be solved one at a time. We provide novel conditions under which *DPA* recovers the solution to the full problem exactly or approximately, and which make the necessary conditions sufficient for optimality: essentially, valuations must be sufficiently correlated across quality increments. Applied to empirical estimates of demand for health insurance, we show that *DPA* is approximately valid, and we apply it to understand equilibrium outcomes in a monopoly insurance market.

JEL: I11, C26, I13

*We are grateful to Zack Cooper, Stuart Craig, Michael Dickstein, David Dranove, Eilidh Geddes, Ben Handel, Kate Ho, Amanda Kowalski, Tim Layton, Ilse Lindenlaub, Neale Mahoney, Alessandro Pavan, Rob Porter, Mike Powell, Mark Shepard, Ashley Swanson, Ben Vatter, Mike Whinston and seminar participants at Northwestern, the Utah Winter Business Economics Conference, Stanford, UCLA, Paris School of Economics, Michigan State, Penn, the NBER IO program, University of Nevada-Reno, Arizona State, Karlsruhe, HSBC Business School, and Penn State for their helpful comments. Chade: hector.chade@asu.edu, Arizona State University; Marone: victoria.marone@yale.edu, Yale University and NBER; Starc: amanda.starc@kellogg.northwestern.edu, The Kellogg School of Management, Northwestern University, and NBER; Swinkels: j-swinkels@kellogg.northwestern.edu, The Kellogg School of Management, Northwestern University.

1 Introduction

Asymmetric information is central to our understanding of major sectors of the economy. Regulators oversee firms that have private information about costs and consumer tastes. Investors evaluate entrepreneurs with differing abilities and project qualities. And in a particularly critical sector, health insurers sell insurance contracts to consumers who know both their health status and their taste for insurance. In each of these settings, agents hold multiple important dimensions of private information, and the space of possible contracts across which they can be screened is potentially vast. Until recently, however, the majority of both theoretical and empirical papers on screening have either considered a one-dimensional hidden information problem (e.g., Rothschild and Stiglitz, 1976; Stiglitz, 1977; Mussa and Rosen, 1978; Maskin and Riley, 1984), or else considered a multidimensional problem with restrictive assumptions on payoff functions (e.g., Armstrong, 1996; Rochet and Choné, 1998; Manelli and Vincent, 2006) or highly restricted contract spaces (e.g., Einav et al., 2010; Veiga and Weyl, 2016).

One reason for this dearth of evidence is, of course, that screening problems become substantially less tractable with multidimensional types. Since Wilson (1993), theorists have attempted to characterize optima of multidimensional screening models. But despite being a natural and highly relevant generalization of the one-dimensional case, and despite the enormous progress made by the aforementioned seminal papers in the literature, there is still much work to be done. Existing results rely on assumptions imposed on consumer utility as a function of their type and allocation, such as linearity of utility in type, single-crossing property between quantity or quality and every dimension of type, multiplicative separability, and convexity conditions. Unfortunately, these assumptions fail in many applied settings. In health insurance, for example, the consumer's valuation of coverage is derived from her expected utility from a lottery over health outcomes, which is not likely to have the properties needed to apply the known results. In consequence, once one takes seriously both that agents can vary along several private dimensions and that a principal can flexibly design a menu of contracts, we have only a limited understanding of properties of optimal menus.

This paper studies what can be learned about the solution to general multidimensional screening problems in which a principal offers a menu of quality-price pairs. We make progress on two fronts. First, we derive necessary conditions for a solution in a setting with only minimal assumptions on primitive objects. These necessary conditions are natural generalizations of the textbook conditions in the one-dimensional case. Despite not being in general sufficient, they nonetheless have substantive economic content, shedding light on several properties of the solution that are of economic interest: optimal exclusion, the existence of trade, and the incentives to screen. Second, we build on the so-called *demand-profile approach* (*DPA*) introduced by Wilson (1993). This

approach recasts the problem in terms of incremental quality levels and prices, yielding a sequence of simple and tractable subproblems, one per increment. We provide novel conditions under which the candidate solution provided by *DPA* is actually a solution to the full problem, or else an approximation thereof. As a by-product, these conditions also make our necessary conditions sufficient for optimality in the full problem. We then apply this approach to a calibrated model of health insurance, where *DPA* closely approximates the full solution, while also substantially clarifying the economics behind it.

Our model is as follows. A principal can offer a set of products, indexed by a scalar quality level, to a consumer or a population thereof with multidimensional private information. A consumer's willingness to pay for quality is strictly increasing in one of the type dimensions, but its dependence on the others is unrestricted. The consumer's outside option is an exogenous "base" quality provided by a third party (which may be zero). The principal's payoff is a weighted average of consumer surplus, principal profits, and (the negative of) any costs borne by the third party. This formulation nests both a profit-maximizing monopolist, who focuses only on profits, and a utilitarian social planner, who values all three components equally. The timing is as follows. The principal offers a menu of price-quality pairs. Consumers observe the menu, learn their type, and choose a quality level.

We begin by deriving necessary conditions that any optimal menu must satisfy. The key assumption required is that consumer utility exhibits single crossing with respect to quality and *some* dimension of the consumer's type, with other interactions between the consumer's utility and private information arbitrary. These necessary conditions are the natural generalization of those in the one-dimensional case, and we show that they have significant economic content, both for interpreting optimal prices and for characterizing the optimal allocation. The conditions can be interpreted through the familiar lens of marginal benefit equals marginal cost. They also establish that a positive measure of types are generically excluded from trade with the principal, that there is always positive trade in the market, and that under a mild assumption and an economically motivated definition of "intermediate quality," all intermediate qualities are assigned to a positive measure of types. The exclusion result is known in the multidimensional mechanism design literature, but derived there under different (and in many ways more stringent) assumptions. Our setup affords a simple and intuitive proof that does not rely on multiplicative separability or convexity assumptions.

We then turn to the question of sufficiency, which has been a central challenge in the existing literature. We make progress on this front using the demand-profile approach (*DPA*) to nonlinear pricing problems. This approach recasts the screening problem in terms of incremental quality levels and prices, decoupling it into a sequence of simple subproblems, one per increment. The decoupling buys us two things. First, it yields substantial economic insight into the solution, which

is difficult to achieve if one focuses on the full problem at the level of generality we use. Second, it delivers comparative statics results with respect to the weights in the principal's objective function and with respect to the principal's incentives to screen. Of course, decoupling of the full problem into a sequence of incremental quality-price problems need not deliver an optimal solution to the full problem. A major result of our analysis is that we provide novel sufficient conditions on primitives (essentially, on the distribution of types) under which *DPA* to be valid. When these conditions hold, *DPA* solves the full problem exactly. A by-product of this result is that these same conditions make our necessary conditions in the full problem sufficient for optimality. But, these sufficient conditions are strong and hard to satisfy in empirical work. We thus also provide weaker conditions under which *DPA* is *approximate* valid. Both of these results (validity and approximate validity of *DPA*) rely on having substantial correlation among consumer values.

We apply our analysis to health insurance menu design. Here, the quality of a product (insurance contract) is its generosity of coverage. The insurer (principal) offers a menu of vertically differentiated contracts, and each consumer either purchases one of them or else remains in a base level of coverage provided by the government (third party). We study a simulated population of consumers calibrated to match demographics of the under-65 US population and parameter estimates from Marone and Sabety (2022). Using five potential contracts in our main analysis, we solve for the optimal menu across three focal principals (a monopolist insurer, a utilitarian planner, a planner facing an excess cost of public funds).[1] This numerical analysis serves three purposes. First, it lets us quantitatively evaluate the predictions derived from the necessary conditions—optimal exclusion, positive trade, and the use of every intermediate quality level—which are results on sign but not magnitude. Our numerical results imply that a monopolist excludes substantially more consumers than a social planner facing no excess cost of funds—37 percent of consumers are not served by the monopolist, whereas less than 1 percent are not served by the planner. Likewise, the monopolist screens more than the social planner, separating consumers across a wider range of coverage levels and offering much less coverage overall.

Second, the numerical analysis lets us assess the empirical plausibility of the sufficient conditions for approximate validity of *DPA*. These conditions are remarkably well satisfied: the distribution of consumer demand for incremental coverage is close to non-crossing, meaning consumers' marginal valuations across increments are substantially correlated in the population, and the implied optimal allocation features quantities that are decreasing in quality level, meaning the resulting final allocation (after integrating back across increments) is coherent. For each of our focal principals, payoffs under *DPA* agree with those of the true problem within one percent.

Finally, with *DPA* validated, we use it to interpret the underlying economics of the problem.

[1] We also establish that as the number of potential quality levels grows, the optimal price schedule and allocation converge to their continuum counterparts. In our setting, five contracts (quality levels) suffice to capture over 98 percent of the payoff achievable with 65 contracts.

Each increment of coverage becomes its own market, with its own demand, marginal revenue, and marginal cost curves. Optimal incremental pricing thus reflects familiar identities: price equals marginal cost for a planner, and price equals marginal cost plus a markup inversely proportional to the demand elasticity for a monopolist. We use this structure to show how the problem can be analyzed graphically, in the spirit of Einav et al. (2010). While their original analysis was restricted to only two potential contracts, our analysis provides the appropriate extension to an arbitrary number of vertically differentiated contracts.[2] Optimal outcomes at each level of coverage can be read directly from the demand, marginal revenue, and marginal cost curves for incremental quality. Viewing the problem through this lens provides intuition for how the principal's objective shapes the allocation at each level of coverage. The ratio of the monopolist's premium relative to average cost falls steadily as coverage rises, from 3.7 on the lowest coverage contract, where demand is least elastic, to 1.1 on the highest, where demand is most elastic. The monopolist's optimal menu reduces social welfare by \$728 per household per year (equal to 7 percent of household average total healthcare spending) relative to what can be achieved by the planner. As the cost of public funds rises, however, losses in the market become more costly for the planner, and it begins acting more like a monopolist. The comparative statics delivered by *DPA* are thus also confirmed quantitatively. As the weight on consumer surplus rises, the relevant margin at each increment shifts from marginal revenue toward demand, and the optimal quantity increases.

Our theoretical approach is related to the seminal works by Stiglitz (1977) (insurance), Mussa and Rosen (1978) (quality provision), and Maskin and Riley (1984) (quantity provision). These papers similarly analyze a principal-agent problem with private information, but consider only one-dimensional private information. There is a subsequent important literature on screening with multidimensional private information, including Wilson (1993), Armstrong (1996), Rochet and Choné (1998), and Manelli and Vincent (2006), which has been surveyed by Rochet and Stole (2003). Our class of problems belongs to Section 5 of Rochet and Stole (2003), which they call "the one-dimensional instrument" case. An early contribution to this class is the parametric example solved in Laffont et al. (1987), which is a special case of our general formulation. More recently, Deneckere and Severinov (2017) provide a solution for a class of problems with two-dimensional private information, and Veiga and Weyl (2016) characterize the solution to a problem very similar to ours, but in which the insurer is exogenously constrained to offering just one contract. In a recent independent contribution, Araújo and Perez (2025) analyzes a screening problem with a one-dimensional instrument and multidimensional types. They provide conditions on the agent's utility that imply that local incentive compatibility implies global incentive compatibility, and an algorithm to find the optimal allocation when types are two-dimensional. Finally, our proof of

[2]Note that with only two potential contracts, *DPA* is trivially valid, since each consumer's payoff function contains only two points (with the left point fixed at zero).

the validity of *DPA*, which basically shows that if *DPA* is valid under perfect correlation among the type dimensions then it is valid with high but imperfect one, bears some resemblance to the argument in Yang (2025), which in a multidimensional screening problem with two instruments provides conditions under which a single instrument is used under perfect correlation and then extends the result to stochastic correlation.

There is also a recent theoretical literature on competitive markets with multidimensional private information, such as Azevedo and Gottlieb (2017), who provide a new equilibrium concept in settings with adverse selection, and Farinha Luz et al. (2023), who focus on risk classification. Insurers in these papers are price-takers, while the insurer in our setting is a price-setter. In a recent independent contribution, Gottlieb and Moreira (2023) analyze optimal monopoly insurance with multidimensional types. They too derive an optimal exclusion result under permissive primitives, but only with binary losses. Moreover, they show that competitive firms provide less coverage than monopoly for those who have a higher willingness to pay for coverage. The two papers complement one another nicely with respect to properties of optimal menus.

Our approximation method relies on the pioneering work of Wilson (1993), which focused on nonlinear pricing in a setting without common values (that is, without selection). While this approach has been used in a number of theoretical contexts, there are few tests of its applicability in real-world settings.[3] Our finding that this approach provides an excellent approximation to the true multidimensional screening problem in health insurance markets presents a promising avenue for new theoretical and empirical exploration. This conversation between theory and practice—the theory provides empirical guidance and the data inform the theory—seems particularly relevant to menu design, where until now empirical analysis appears to have outpaced theoretical advancement.[4]

Our paper is also related to a large empirical literature on health insurance.[5] Our graphical analysis of the insurer's problem builds on the foundational framework in Einav et al. (2010), who

[3] One recent example is Gaynor et al. (2023), who use the demand-profile approach to finding the optimal nonlinear reimbursement contract to offer healthcare providers.

[4] This is especially true in health insurance markets. As noted by Einav and Finkelstein (2011) (and emphasized by Veiga and Weyl, 2016), "On the theoretical front, we currently lack clear characterizations of the equilibrium in a market in which firms compete over contract dimensions as well as price, and in which consumers may have multiple dimensions of private information." While endogenous determination of contract characteristics has been recognized as a centrally important force, applied researchers must often abstract from it in order to make empirical progress. Our approach accommodates this force, permitting tractable analysis in settings that have traditionally been considered prohibitively complex. In particular, our convergence result (Section 6) shows that as the number of potential quality levels grows, the optimal price schedule and allocation in the finite case converge to their continuum counterparts. Absent a fixed cost of offering additional contracts, competing over contract dimensions and competing over price are thus not distinct problems.

[5] Our model of consumer demand for health insurance builds on a workhorse introduced by Cardon and Hendel (2001), which has been used in several subsequent papers (for example, Einav et al., 2013; Azevedo and Gottlieb, 2017; Ho and Lee, 2023; Marone and Sabety, 2022). We enrich the model to allow a unified treatment of insurers with differing objective functions.

focus on competitive markets and two potential contracts, and Mahoney and Weyl (2017), who focus on imperfect competition with two potential contracts. A central contribution of our paper is to show under what conditions this approach can be extended to an arbitrary number of vertically differentiated contracts. Our focus on health insurance menu design for multidimensional consumers is also closely related to recent work by Marone and Sabety (2022) and Ho and Lee (2023), who each solve for the optimal menus of contracts that would be offered by a utilitarian planner in their respective empirical settings. We build on these findings by asking to what extent various features of those solutions will hold in general, and how optimal menus would change with the insurer objective function.

The paper is organized as follows. Section 2 describes the model. Section 3 derives necessary conditions that any optimal menu must satisfy. Section 4 recasts these conditions in more familiar economic terms and uses them to establish three implications that hold regardless of sufficiency: optimal exclusion, positive trade, and the use of every intermediate quality level. Section 5 develops the demand-profile approach and explores the economics it illuminates when valid. Section 6 extends the analysis to a continuum of quality levels and connects the finite and continuum cases. Section 7 presents the application of our analysis to the health insurance setting. While the theory informs the numerical exercise (and vice versa), Section 7 is largely self-contained for the applied reader. Proofs and additional technical material are available in the appendices.

2 The Model

We wish to accommodate a wide variety of settings in which a monopolist *principal* designs a set of contracts amongst which a privately informed *agent* chooses. As will be seen in our setup, there is no harm in reinterpreting the model such that instead of a single agent, the principal faces a set of consumers, and we will refer to the agent as a consumer where that is the more natural interpretation.

To accommodate important applied settings, including health insurance, we allow for the possibility that there is a *government* that is affected by what occurs in the market, and we allow the principal to put arbitrary weights on the payoffs of different actors in the system. While the type space is multidimensional, the allocation to the agent is one-dimensional, and we henceforth refer to it as *quality*, which is indexed by $x \in [0, 1]$.[6]

The agent has type which we will for convenience write as (ψ, θ) where ψ lies in a closed and bounded interval $\Psi = [\underline{\psi}, \bar{\psi}]$ of the reals, while θ lies in a finite-dimensional cube Θ of $\mathbb{R}^N$. Let H be the distribution of (ψ, θ).

[6] In the terminology of Rochet and Stole (2003), this is a multidimensional screening problem with a one-dimensional instrument.

The agent's utility is linear in money, where an agent of type (ψ, θ) who receives quality x and pays t has payoff (consumer surplus) $v(x, \psi, \theta) - t$ where v is $\mathcal{C}^2$. Our model is of vertical differentiation, so that $v_x > 0$. The dependence of v on θ is arbitrary, but crucial to the tractability of our analysis (and the reason that we separate it out in the notation) is that increases in ψ strictly increase the marginal valuation of the agent for quality. That is, $v_{x\psi} > 0$. In the health insurance application in Section 7, we show that if we take ψ to be risk aversion, then $v_{x\psi}$ is indeed strictly positive, but in other settings ψ might for example be a parameter that indexes the intercept of the consumer's demand.

The principal is risk-neutral with respect to money and incurs financial cost $\gamma(x, \psi, \theta)$ when an agent of type (ψ, θ) chooses quality x, where γ is $\mathcal{C}^2$ and strictly increasing in x. The government has $\mathcal{C}^2$ financial cost $\eta(x, \psi, \theta)$ and is also risk-neutral. The principal puts weights $w \equiv (w_C, w_P, w_G) \geq 0$ on consumer surplus, their own profits, and government spending. For simplicity, we assume $w_P > 0$ and normalize it to one such that the principal has payoff

$$S(t, x, \psi, \theta) = w_C(v(x, \psi, \theta) - t) + t - \gamma(x, \psi, \theta) - w_G\eta(x, \psi, \theta). \tag{1}$$

A profit-maximizing principal has $w = (0, 1, 0)$. A utilitarian social planner has $w = (1, 1, 1)$ and a government running publicly financed health insurance, housing, or education from general revenue has $w = (1/\tau, 1, 1)$, where $\tau > 1$ reflects the cost of government funds. As another example, an employer running an employee health plan might have $w = (\tau, 1, 0)$, with $\tau < 1$.[7]

The principal is allowed to offer quality levels from an exogenously given finite set $X = \{x_0, x_1, x_2, \ldots, x_J\}$ with $x_0 < x_1 < x_2 < \cdots < x_J \leq 1$. Section 6 shows how to generalize to the case where there is a continuum of quality levels in $[0, 1]$. Quality $x_0 \in [0, 1]$ is available to the agent for free and represents the agent's outside option. Let $\rho_j \equiv \rho(x_j)$, $j = 0, 1, ..., J$. The principal chooses a price vector $\rho : X \to \mathbb{R}$ where $\rho_0 = 0$ and $\rho_{j+1} \geq \rho_j$. Let $\mathbb{P}$ be the set of admissible price vectors.[8] The principal also suggests a choice $\chi(\psi, \theta) \in X$ to each type. Incentive compatibility is that

$$\chi(\psi, \theta) \in \arg\max_{x \in X}(v(x, \psi, \theta) - \rho(x)) \tag{IC}$$

for each (ψ, θ). The principal's problem is then

$$\max_{\rho \in \mathbb{P}, \chi} \int S(\rho(\chi(\psi, \theta)), \chi(\psi, \theta), \psi, \theta) dH(\psi, \theta) \text{ s.t. } IC. \tag{$\mathcal{P}$}$$

[7] A setting with $w_P = 0$, or indeed $w_P < w_C$, would need to include a constraint on the losses of the principal, otherwise infinite transfers to the consumer would be optimal.

[8] Since $v_x > 0$, any non-monotone price vector ρ is equivalent to the one that for each j replaces $\rho(x_j)$ by $\min_{j' \geq j} \rho(x_{j'})$, and so monotonicity is innocuous.

Note that we restrict our attention to deterministic mechanisms. Except for notable exceptions (for example, Manelli and Vincent (2006)), this is a standard assumption in the literature on multidimensional mechanism design (see Rochet and Stole (2003) on this point). Stochastic mechanisms can indeed be useful to the principal when types are multidimensional (Manelli and Vincent, 2006), or when the type includes the agent's coefficient of absolute risk aversion (Kadan et al., 2017). We rule them out here both for reasons of tractability and because they are implausible in our health insurance application.[9]

2.1 Simplifying Assumptions

To facilitate the analysis of this problem, we will put some mild structure on H and v. Define $\mathcal{H}$ as the set of probability measures with support contained within $\Psi \times \Theta$ (potentially a strict subset) that satisfy the following three assumptions. With some abuse of notation, we will use H also for marginal and conditional distributions, and similarly h for their densities, but it will be clear from the context to which we are referring. Our first simplifying assumption is that H can be expressed in terms of continuous densities and conditional densities.

Assumption 1 (Continuous Densities) *H has a continuous density h. The conditional probability of one set of the coordinates of $\Psi \times \Theta$ on the realization of another can be expressed as a density $h(\cdot|\cdot)$ which is continuous on $\Psi \times \Theta$.*

Second, we impose a condition that, facing any given price vector, only very rarely does the agent have three optimal choices at once.

Assumption 2 (Diffuse Conditional Willingness to Pay) *Fix three quality levels $x_\ell < x_m < x_h$ contained in X, and τ_1 and τ_2 strictly positive. Let $\hat{\Theta} \subseteq \Theta$ be the set of θ where for some single value of ψ, (ψ, θ) is simultaneously willing to pay τ_1 to increase quality from x_ℓ to x_m and τ_2 to increase quality from x_m to x_h. Then, $H(\hat{\Theta}) = 0$.*

Appendix B.1 gives two permissive sets of conditions that ensure Assumption 2. Indeed, we doubt there are economically meaningful examples where it fails.[10]

It also clarifies what follows if we assume that when ψ is fixed at one of $\underline{\psi}$ or $\bar{\psi}$, the willingness to pay for any given quality increment does not take on any given value with positive probability.

[9] We find it implausible that the insurer in Section 7 would be allowed to run lotteries to determine premium and coverage for different consumers (many regulations prevent charging identical consumers different premiums).

[10] For a graphical intuition, form the level sets of points (ψ, θ) such that the agent is willing to pay τ_1 to increase quality from x_ℓ to x_m, and the level sets of points such that the agent is willing to pay τ_2 to increase quality from x_m to x_h. Since $v_{x\psi} > 0$, on each level set, there is at most one ψ associated with any given θ. Assumption 2 says that the associated ψ does not coincide on a positive measure subset of Θ.

Assumption 3 (No Atoms on Boundary) *Fix $x'' > x'$ and $\tau > 0$. Then,*

$$H(\{\theta | v(x'', \bar{\psi}, \theta) - v(x', \bar{\psi}, \theta) = \tau\}) = 0 \text{ and } H(\{\theta | v(x'', \underline{\psi}, \theta) - v(x', \underline{\psi}, \theta) = \tau\}) = 0.$$

Appendix B.1 argues that this is equally mild in our multidimensional setting.[11] Of course this condition is intrinsically about a multidimensional setting, since it fails automatically when the problem is one-dimensional with ψ as the only active parameter, so that θ takes on a single-value.

Finally, at various points in what follows, we will transition from prices to quantities, and that is cleaner if the demand for any increment in quality is strictly decreasing over the relevant range.

Assumption 4 (Demand Curve Slopes Down) *Fix $x'' > x'$, and let $\mathcal{S}$ be the support of H. Then, $H(\{(\psi, \theta) | v(x'', \underline{\psi}, \theta) - v(x', \underline{\psi}, \theta) > \tau\})$ is strictly decreasing in τ for $\tau \in (\min_{(\psi,\theta)\in\mathcal{S}}(v(x'', \underline{\psi}, \theta) - v(x', \underline{\psi}, \theta)), \max_{(\psi,\theta)\in\mathcal{S}}(v(x'', \underline{\psi}, \theta) - v(x', \underline{\psi}, \theta)))$.*

In Sections 3–4, $\mathcal{H}$ denotes the set of distributions satisfying Assumptions 1–3. When the problem is recast in quantities or inverse demand functions, we also impose Assumption 4.

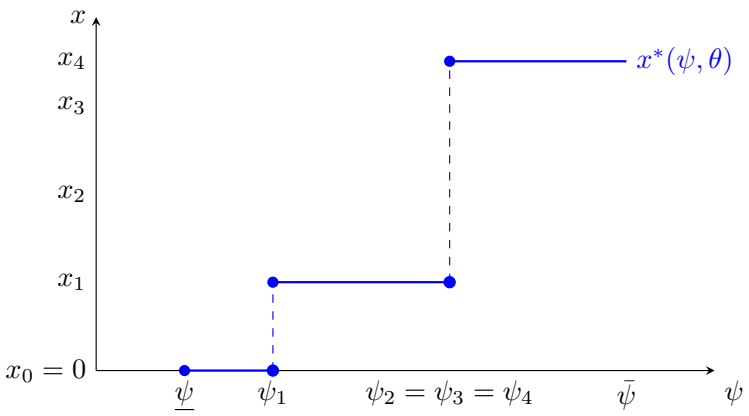
3 Optimal Menu Design

We now derive necessary conditions that any optimal menu must satisfy. They generalize the familiar screening conditions in Mussa and Rosen (1978) and Maskin and Riley (1984) for the one-dimensional case. It is common to derive optimality conditions by varying the allocation and tracing through the effects on prices. But, as Wilson (1993) pp. 212-214 points out, this approach is incoherent with multidimensional types. To illustrate the point, consider pricing seats on an airplane, where the willingness to pay extra for a first-class seat depends on the consumer's height, taste for better food, and value of early boarding. For any given incremental price of a first-class seat, there is a level-set ψ (height) as a function of θ (taste for better food and early boarding) such that a consumer with θ strictly prefers first class if and only if their height is above the level set. Any allocation that does not coincide with one of these level sets cannot be supported by any price vector. Because of this, we work directly with price vectors, and track the optimal choices of the agent as these price vectors are perturbed.

Fix some θ and ρ and consider the agent's optimal choices as a function of ψ. An example is given in Figure 1. Because $v_{x\psi} > 0$, if a given ψ prefers some quality level to a lower one, then this preference is strict for any higher ψ. Thus, the agent's optimal choice is unique except at some jump points. But, a jump may skip some quality levels. Because the conditional distribution of

[11] An alternative sufficient for our purposes is that $h(\bar{\psi}|\theta) = h(\underline{\psi}|\theta) = 0$ for each θ.

Figure 1. Optimal Choice



Notes: The consumer's optimal choice as a function of ψ for given ρ and θ.

ψ is atomless, it does not matter what quality the types at jumps choose, and so the principal's payoff is tied down completely by the values of the jump points. Let $\psi_j(\rho, \theta)$ be the point at which the optimal choice changes from something strictly below x_j to x_j or above.[12] The agent has unique optimal quality j on the open interval between $\psi_j(\rho, \theta)$ and $\psi_{j+1}(\rho, \theta)$. Let $\underline{x}_j(\rho, \theta)$ and $\bar{x}_j(\rho, \theta)$ be the lowest and highest optimal choices at $\psi_j(\rho, \theta)$. In the figure, at ψ_1 the optimal choice switches from x_0 to x_1 and thus $\underline{x}_1 = x_0$ and $\bar{x}_1 = x_1$. But, at the next jump point, the optimal choice jumps from x_1 to x_4, and so ψ_2, ψ_3, and ψ_4 agree, with $\underline{x}_j = x_1$ and $\bar{x}_j = x_4$ for $j \in \{2, 3, 4\}$.

Define

$$\mathcal{V}(x, \psi, \theta) \equiv v(x, \psi, \theta) - \gamma(x, \psi, \theta) - w_G \eta(x, \psi, \theta).$$

We shall refer to this object as the *weighted social surplus* in that when $w_G = 1$ it is the social surplus created by allocating type (ψ, θ) to quality x, but otherwise, the cost to the government is weighted. In the one-dimensional case, this object is familiar when $w_G = 0$.

We can then rewrite the principal's payoff as

$$S(t, x, \psi, \theta) = \mathcal{V}(x, \psi, \theta) - (1 - w_C)(v(x, \psi, \theta) - t). \tag{2}$$

The first term is the weighted social surplus. The term $v(x, \psi, \theta) - t$ is the consumer surplus of

[12] If no $\psi > \underline{\psi}$ has an optimal choice strictly below x_j set $\psi_j = \underline{\psi}$ and if no ψ has an optimal choice at or above x_j set $\psi_j = \bar{\psi}$.

the agent. The weight $1 - w_C$ is the cost to the principal of transferring money to the agent. Because of this, if we fix a price function ρ, and assume that some type (ψ, θ) is indifferent given ρ between x_ℓ and $x_h > x_\ell$, then the impact of a switch from x_ℓ to x_h by (ψ, θ) on the principal (that is, $S(\rho(x_h), x_h, \psi, \theta) - S(\rho(x_\ell), x_\ell, \psi, \theta)$) is $\mathcal{V}(x_h, \psi, \theta) - \mathcal{V}(x_\ell, \psi, \theta)$. To see this, note that the term involving the change in the consumer surplus drops out since the agent is indifferent, and so in particular, $v(x_h, \psi, \theta) - v(x_\ell, \psi, \theta) = \rho(x_h) - \rho(x_\ell)$.

Let

$$r_j(\rho, \theta) \equiv \frac{\mathcal{V}(\bar{x}_j(\rho, \theta), \psi_j(\rho, \theta), \theta) - \mathcal{V}(\underline{x}_j(\rho, \theta), \psi_j(\rho, \theta), \theta)}{v_\psi(\bar{x}_j(\rho, \theta), \psi_j(\rho, \theta), \theta) - v_\psi(\underline{x}_j(\rho, \theta), \psi_j(\rho, \theta), \theta)}$$

if $\psi_j(\rho, \theta)$ is interior, and let $r_j(\rho, \theta) \equiv 0$ otherwise. Note that $\mathcal{V}$ does not depend directly on the transfer, and hence neither does r_j.

Theorem 1 (Optimality Condition) *Let $H \in \mathcal{H}$, and let (ρ, χ) be optimal. Then, for $1 \leq j \leq J$, if $\rho_j > \rho_{j-1}$ then*

$$(NC) \qquad \int_\Theta [(1 - w_C)(1 - H(\psi_j(\rho, \theta)|\theta)) - r_j(\rho, \theta) h(\psi_j(\rho, \theta)|\theta)] h(\theta) d\theta = 0,$$

and it is weakly less than zero if $\rho_j = \rho_{j-1}$. The allocation χ is characterized H-a.e. by $\{\psi_j\}_{j=1}^J$.

The proof of this is simple and contains economic intuition, and so we present it here. Fix j and change the price of all qualities x_j or higher by a constant ε, where when $\rho_{j-1} = \rho_j$ we must have $\varepsilon > 0$ to respect monotonicity, but otherwise we can vary ε in either direction. Let the agent of any given type reallocate to his (almost everywhere) unique optimal choice facing the perturbed price vector. Consider the derivative of the principal's payoff with respect to ε at $\varepsilon = 0$. To do so, for each θ we calculate how the agent's jump points vary with ε and use this to calculate the derivative of the principal's payoff. We will argue that except on a probability-zero subset of Θ, the square-bracketed term in *NC* is indeed this derivative. We then integrate with respect to θ.[13]

Consider first those θ such that $\psi_j(\rho, \theta) = \underline{\psi}$. By Assumption 3, except on a set of θ with probability zero, any $x < x_j$ is suboptimal facing ρ for $\underline{\psi}$.[14] Thus, for small ε, these qualities remain suboptimal.[15] Thus, *a fortiori* the optimal response for any other ψ is also x_j or above.

[13] Exchanging the order of the integral and derivative is harmless since X is finite and v, γ, η and the conditional densities are continuous on compact sets, so the difference quotients are uniformly bounded near $\varepsilon = 0$ (except for a measure zero set of values of θ). Therefore, the Lebesgue dominated convergence theorem can be applied.

[14] For all relevant θ, $(\underline{\psi}, \theta)$ weakly prefers their optimal choice to anything strictly below x_j, and by Assumption 3, the set of θ such that that $(\underline{\psi}, \theta)$ is indifferent between any two qualities facing ρ has zero measure, and so except on this set, the preference is strict.

[15] Let $\hat{x} \geq x_j$ be an optimal choice of $\underline{\psi}$. Then, for each $x < x_j$, $v(\hat{x}, \psi, \theta) - \rho(\hat{x}) > v(x, \psi, \theta) - \rho(x)$ and so for small ε, $v(\hat{x}, \psi, \theta) - \rho(\hat{x}) - \varepsilon > v(x, \psi, \theta) - \rho(x)$ and thus x is not optimal. Choose ε small enough that this holds for each $x < x_j$.

But then, since the relative prices of qualities at or above x_j have not changed, all ψ make the same optimal choices as before, and the only effect of the perturbation is to transfer ε from the agent of type θ to the principal. The derivative of the principal's payoff facing θ at $\varepsilon = 0$ is thus $1 - w_C$, which is what the bracketed term in *NC* reduces to, using that $r \equiv 0$ and $1 - H(\psi_j|\theta) = 1$.[16] Similarly, if $\psi_j = \bar{\psi}$, then except on a set of θ with probability zero, the allocation is unchanged, and no money is transferred, and the bracketed term is appropriately zero.

So, consider ψ_j interior. By Assumption 2, except on a set of θ of zero probability, $\underline{x}_j$ and $\bar{x}_j$ are the only optimal choices for ψ_j. For example, in the figure, x_3 is suboptimal where the optimal choice switches from x_1 to x_4. But then, on some interval $[\psi_j - \delta, \psi_j + \delta]$ around ψ_j, only $\underline{x}_j$ and $\bar{x}_j$ are candidates for optimality facing ρ and thus remain so for ε small. The division point between $\underline{x}_j$ and $\bar{x}_j$ is thus $\tilde{\psi}(\varepsilon|\theta)$ defined by

$$v(\bar{x}_j(\rho,\theta), \tilde{\psi}(\varepsilon|\theta), \theta) - v(\underline{x}_j(\rho,\theta), \tilde{\psi}(\varepsilon|\theta), \theta) = \rho(\bar{x}_j(\rho,\theta)) + \varepsilon - \rho(\underline{x}_j(\rho,\theta)). \tag{3}$$

so that $\tilde{\psi}(\varepsilon|\theta)$ values the increase in quality from $\underline{x}_j$ to $\bar{x}_j$ by ε more than does ψ_j. By the implicit function theorem, $\tilde{\psi}_\varepsilon(0)$ is then one over the denominator of r_j, which is strictly positive since $v_{x\psi} > 0$. Evaluated at $\varepsilon = 0$, the rate of types switching their choice from $\bar{x}_j$ to $\underline{x}_j$ is thus $h(\psi_j|\theta)$ divided by the denominator of r. And, as argued above, the difference between $S(\rho(\cdot), \cdot, \psi_j, \theta)$ evaluated at $\bar{x}_j$ and $\underline{x}_j$ is just the difference in $\mathcal{V}$'s in the numerator of r_j. Thus, $r_j h(\psi_j|\theta)$ captures the effect on the principal's payoff from types who change their choice. Finally, $1 - H(\psi_j|\theta)$ is the mass of agents who pay the higher price without changing their allocation. The bracketed term is thus again the derivative of the principal's payoffs facing θ, and we are done.[17]

It is of great interest to know when these necessary conditions are also sufficient beyond the one-dimensional case where there are well-known primitives. One approach is to look for conditions under which the principal's objective function is concave in the price vector ρ. This is near-hopeless. To see one important reason why, let there be only two qualities x_0 and x_1 with $v(x_1, \psi, \theta) = \theta\psi$ and $v(x_0, \psi, \theta) = 0$, and with ψ uniform on $[0, 1]$ independent of θ. Then, the demand for x_1 is linear with choke price θ and maximum quantity one and so the profits facing θ of a pure monopolist with zero marginal costs are concave on $[0, \theta]$.[18] But, as soon as we take an expectation of profits across different θ, we have to consider prices beyond any given type's choke price, and since the profits facing θ are zero at prices above θ, there is an upward kink in profits facing θ at θ. This non-concavity for a given θ can easily destroy concavity when one takes expectations across different θ's. In this highly simplified example, there are conditions on the

[16] Per Footnote 11, note that if $h(\underline{\psi}|\theta) = 0$, then the bracketed term also reduces to $1 - w_C$.

[17] One can use *NC* in a standard way to produce a candidate solution. For each active set of increments that are strictly positive, set to zero the remaining increments and solve *NC* for those that are strictly positive. The solution to this system yields a candidate ρ, and thresholds ψ_j's, which then produce a candidate allocation χ.

[18] Demand for x_1 is given by $\Pr(\{\theta\psi > p\}) = \Pr(\{\psi > \frac{p}{\theta}\}) = 1 - \frac{p}{\theta}$.

distribution of θ that smooth these kinks and restore concavity. But, we see no hope of doing this in general and with multiple quality levels.

Deneckere and Severinov (2017) consider a two-dimensional setting and put enough structure on the problem to allow a reduction to a tractable optimal-control problem. This reduction is useful but seems unlikely to substantially generalize beyond two dimensions. Figalli et al. (2011) and McCann and Zhang (2019) use tools from optimal transport and provide conditions for concavity, but (see Pass and Halim (2025)) in our setting these basically reduce the problem to a single dimension, in that there exists a one-dimensional index such that if one type has a higher index than another then that type has higher marginal willingness to pay for all quality levels. We will substantially extend this idea when we provide conditions for the validity and near-validity of the demand-profile approach, and hence for the sufficiency of our necessary conditions.

4 Interpretation and Implications of the Necessary Conditions

In this section, we relate the necessary conditions *NC* to the one-dimensional case. Then, we recast them in terms of more economically familiar objects. This allows us to derive some important implications of the necessary conditions that hold regardless of sufficiency. Additional implications of optimality will be derived later when we add further structure to the problem.

Relating *NC* to the One-Dimensional Case. To start to connect our optimality conditions with existing results, consider the case in which the set of qualities is large and in which adjacent qualities are close by. Assume further (and this is an assumption on an endogenous object) that for each θ, as ψ increases the agent switches from one quality to the next, so that each quality is used by some range of ψ. Then, noting that r (by the Cauchy mean value theorem) can be expressed as $\mathcal{V}_x/v_{x\psi}$ for some intermediate point between x_j and x_{j-1}, one arrives at the standard Mussa and Rosen (1978) condition except for the expectation over θ. This provides leverage to generalize insights from the one-dimensional case to our multidimensional setting.

But, complications arise because it is not clear what conditions will ensure that no quality level is skipped (by any type θ) in the multidimensional setting. Further, one must deal with the fact that as the set of qualities becomes dense there may still be a given quality that is chosen by many types, and issues similar to ironing arise. In Section 6, we derive the analogue to the standard optimality conditions with and without ironing in the continuum case. In Section 5.3, we explore conditions under which no type θ skips any qualities as ψ varies. In such case, our necessary conditions *NC* are also sufficient for optimality.

An Elementary Economic Interpretation of *NC*. In the spirit of Bulow and Roberts (1989), let us see *NC* in more familiar terms. Fix any given ρ and j. Consider a setting in

which the principal's sole tool is to adjust the j^{th} price increment, $\rho(x_j) - \rho(x_{j-1})$, to p_j, but ρ is otherwise unaffected (this is a reparametrization of the perturbation underlying *NC*). Regardless of p_j, and using Assumption 2, type θ will choose between $\underline{x}_j(\rho,\theta)$ and $\bar{x}_j(\rho,\theta)$ facing the perturbed price function for p_j in a neighborhood of $\rho(x_j) - \rho(x_{j-1})$ and so for simplicity, let us consider an (artificial) setting in which $\underline{x}_j(\rho,\theta)$ and $\bar{x}_j(\rho,\theta)$ are the only choices available to θ. Let $\Delta(\theta,\rho)$ be the price to move from $\underline{x}_j(\rho,\theta)$ to $\bar{x}_j(\rho,\theta)$. That is,

$$\Delta(\theta, p_j, \rho) = \rho(\bar{x}_j(\rho,\theta)) - \rho(\underline{x}_j(\rho,\theta)) - (\rho(x_j) - \rho(x_{j-1})) + p_j.$$

Suppressing temporarily the dependence on ρ, let $Q^j(p_j|\theta)$ be the fraction of types ψ who purchase the high quality for each θ, and in a minor abuse of notation, let $Q^j(p_j) \equiv \int Q^j(p_j|\theta)h(\theta)d\theta$ be total demand.[19] Then, the principal's incremental revenue if q_j consumers choose their high quality can be expressed as

$$R^j(Q^j(p_j)) = \int \Delta(\theta, p_j)Q^j(p_j|\theta)h(\theta)d\theta$$

Similarly, (and continuing to suppress ρ), when p_j is set so that q_j consumers choose their high quality option, let $C^j(q_j)$ be the principal's incremental cost, and similarly, let $C^j_G(q_j)$ be the government's incremental cost and $CS^j(q_j)$ the incremental consumer surplus. Let MC^j, MC^j_G, and MCS^j be the respective derivatives with respect to q_j. In calculating these expressions, we track that the low and high qualities can depend on θ, that costs and values will in general depend on the ψ which is pivotal, and that the relative weights on these types in calculating margins will depend on how quickly the pivotal type ψ changes with the price and on the density of that type.

The principal now chooses q_j to maximize

$$w_C CS^j(q_j) + R^j(q_j) - C^j(q_j) - w_G C^j_G(q_j).$$

and so has first-order condition

$$w_C MCS^j(q_j|\rho) + MR^j(q_j|\rho) = MC^j(q_j|\rho) + w_G MC^j_G(q_j|\rho), \tag{4}$$

where we now make the dependence on ρ explicit. This is just the standard first-order condition for a monopolist adjusted for the possibility that the principal might put weight on either the utility of the consumer or the costs of the government.

Proposition 1 (Restatement of Necessary Conditions) *Let $H \in \mathcal{H}$. For each j, ρ with associated p satisfies NC if and only if $q_j = Q^j(p_j|\rho)$ satisfies (4).*

[19] Formally, $Q^j(p_j|\theta) = H(\{\psi | v(\bar{x}_j, \psi, \theta) - v(\underline{x}_j, \psi, \theta) > \Delta(\theta, p_j, \rho)\}|\theta)$.

To see the idea, note that as we vary ε in the perturbation underlying NC, we are doing the exact same thing as when we vary p_j here.[20]

Optimal Exclusion. One implication of our necessary condition is that if the weight on profit is higher than that on the consumer, then the principal optimally excludes a strictly positive mass of types. That is, a strictly positive mass of types receive x_0 at the optimal menu.

Proposition 2 (Exclusion) *Let $H \in \mathcal{H}$, and let $w_C < 1$. Then any optimal menu results in strictly positive exclusion.*

The idea of the proof (see Appendix A) is that under Assumption 3, if no one is using x_0 (so that $\psi_1 = \underline{\psi}$) then a zero measure set of types are indifferent between their choice and x_0. But then, the principal has a first-order profit from raising the price of all contracts x_1 and above and a second-order loss, as reflected in the fact that the $j = 1$ instance of NC is violated.

There are many instances of optimal exclusion in multidimensional screening problems in the literature, starting with Armstrong (1996), who rely on quasilinearity and convexity assumptions on the payoff functions and on the distribution of types, the first of which was significantly relaxed by Barelli et al. (2014). Recent contributions to optimal exclusion in multidimensional screening problems appear in Deneckere and Severinov (2017) and, in an insurance context, in Gottlieb and Moreira (2023). A feature of our result is that, unlike the literature mentioned, Assumption 3 plus a simple and intuitive perturbation argument deliver the well-known optimal exclusion result in a setting with multidimensional types.

Positive Trade. Let us next consider when we can be sure that at least some trade occurs and begin to explore the structure of that trade. A trivial example with no trade is where $w_C = w_G = 0$ and where for all types of the agent, the principal's incremental cost of moving the agent from x_0 to any higher quality is higher than the consumer's incremental value of that quality. More significantly, adverse selection can bite so hard that the consumers who want to trade at any given price are unprofitable. Sufficient for positive trade is that for some j, the demand curve for x_j versus x_0 is somewhere strictly above the average cost curve, since then the principal can make strictly positive profits (compared to no trade) by selling just x_j at an appropriate price.

[20] Indeed, if we take define (ρ_{-j}, p_j) as the price function were (as above) the j^{th} increment is replaced by p_j, then, as shown in the proof of Theorem 1, $Q(p_j|\rho,\theta) = 1 - H(\psi_j(\rho_{-j}, p_j, \theta)|\theta)$ and so $-Q_{p_j}(p_j|\rho,\theta)$ is $h(\psi_j(\rho_{-j}, p_j, \theta)|\theta)$ divided by the denominator of r. The numerator of r is $N \equiv v(\bar{x}_j) - v(\underline{x}_j) - (\gamma(\bar{x}_j) - \gamma(\underline{x}_j)) - w_G(\eta(\bar{x}_j) - \eta(\underline{x}_j))$, where for example $v(\bar{x}_j)$ is shorthand for $v(\bar{x}_j(\rho,\theta), \psi_j(\rho_{-j}, p_j, \theta), \theta)$. Thus, the bracketed expression in (NC) is $(1 - w_C)Q(p_j|\rho,\theta) - Q_{p_j}(p_j|\rho,\theta)N$. Integrating over θ, dividing by $Q_{p_j}(p_j|\rho)$, and rearranging one arrives at (4), since for example,

$$\int (\gamma(\bar{x}_j) - \gamma(\underline{x}_j)) \frac{(-Q_{p_j}(p_j|\rho,\theta))}{-Q_{p_j}(p_j|\rho)} h(\theta) d\theta = MC^j(Q_j(p_j)).$$

To see an economically sensible condition under which we can say more about the extent of trade, consider the set $\{x_j \in \{x_1, \ldots, x_J\} | \mathcal{V}(x_j, \bar{\psi}, \theta) - \mathcal{V}(x_{j'}, \bar{\psi}, \theta) > 0 \text{ for each } j' < j\}$. Assume this set is nonempty and let x^h be its largest element. An example where x^h is equal to the highest possible quality level x_J is an insurance market without moral hazard, since then the risk aversion of the consumer implies that full insurance has higher marginal value than marginal cost of serving any type.[21]

Proposition 3 (Positive Trade) *Let $H \in \mathcal{H}$. In any optimal menu, a strictly positive mass of types are allocated x^h or above.*

The idea is to start from a menu where all types choose strictly below x^h. Begin by lowering $\rho(x^h)$ until the first point where x^h starts to attract some types. This price is strictly higher than that of any quality that is used. Now, lower the price of x^h further by a small ε, and consider the types who are switching as ε changes. These types will have ψ very close to $\bar{\psi}$.[22] Thus, by continuity and compactness, for small ε and by definition of x^h, $\mathcal{V}(x^h, \psi, \theta) - \mathcal{V}(x_{j'}, \psi, \theta) > 0$ for the quality $x_{j'} < x^h$ from which they switch. But, as argued above, this is the principal's payoff from this switch, and since this is true for all $\varepsilon > 0$ sufficiently small, this contradicts optimality.[23]

Every Intermediate Quality Level is Allocated. The exclusion result shows that, with the minimal structure of our background assumptions, x_0 is assigned to a positive measure set of types. Proposition 3 gives conditions under which high quality levels are also used. In this section, we give conditions under which an interval of other qualities is used as well.

Say that quality level x_j is *intermediate* if for all θ, $\mathcal{V}(\cdot, \underline{\psi}, \theta)$ has a maximizer in $[0, 1]$ at or below x_j and $\mathcal{V}(\cdot, \bar{\psi}, \theta)$ has a maximizer at or above x_j. That is, for each θ, an efficient level of quality is (weakly) below x_j when ψ is low, and (weakly) above x_j when ψ is high. Note that the set of intermediate values is consecutive. To see an example where all qualities are intermediate, consider an insurance market (such as the one we formalize below) with mild moral hazard. Let $\underline{\psi}$ correspond to risk-neutrality so that the presence of moral hazard implies that $\mathcal{V}(\cdot, \underline{\psi}, \theta)$ is maximized at x_0, and let high values of ψ correspond to risk-aversion that is severe enough to swamp moral-hazard considerations so that $\mathcal{V}(\cdot, \bar{\psi}, \theta)$ is maximized at x_J.

[21] Once one sees the proof, one will be tempted to try to weaken the condition in the definition of x^h by restricting attention to those values of θ which maximize $v(x_j, \bar{\psi}, \theta) - v(x_{j'}, \bar{\psi}, \theta)$ for some j'. Selection effects kill this idea, since the relevant θ's may be choosing some other quality.

[22] Type (ψ, θ) has willingness to pay for x_j equal to $v(x_j, \psi, \theta) - \max_{j' < j}(v(x_{j'}, \psi, \theta) - \rho(x_{j'}))$. This is strictly increasing in ψ by the envelope theorem.

[23] To connect more formally to our necessary conditions note that what we have just done is to evaluate for each $\varepsilon > 0$ sufficiently small the sign of the sum of *NC* when one lowers the price of all qualities x^h and above by ε and when one raises the price of all qualities strictly above x^h by ε. Note also that this sum is zero at $\varepsilon = 0$ since by Assumption 3 only a zero measure set of types are on the margin of switching. So in some sense, we are exploiting a second-order necessary condition.

The following assumption is the key to our result:

Assumption 5 *$\mathcal{V}_x$ and $\mathcal{V}_x/v_{x\psi}$ strictly decrease in x for each (ψ, θ).*

The first condition is simply that the weighted social surplus is strictly concave in quality for any given consumer. The second term will turn out to be closely related (via Cauchy's mean value theorem) to the behavior of r. In Online Appendix B.2, we show that $\mathcal{V}_x/v_{x\psi}$ decreases with a few technical assumptions when v_x is log-supermodular in x and ψ. To interpret, assume that v is concave in quality, and note that for a given consumer of type (ψ, θ), $v_x(\cdot, \psi, \theta)$ is the consumer's demand curve for quality. Log-supermodularity says that as ψ increases, the demand curve moves up *in percentage terms* more at high x, which is where demand is lowest. So for example, if ψ shifts demand up by a constant, then the condition is trivially satisfied.

Proposition 4 (All Intermediate Levels Used) *Let $H \in \mathcal{H}$ and let Assumption 5 hold. Then, in every optimal menu, every intermediate quality is allocated with strictly positive probability.*

One way of phrasing this result is that some exclusion takes place at each intermediate step.[24] Note that as one adds contracts, the range of contracts that are intermediate is not changing. Hence, this result implies that all intermediate qualities are used even as one moves towards a continuum of qualities. To see the idea of the result, let x_{j^*} be intermediate but not used. Lower the price of x_{j^*} until it attracts just a few switchers. If these switchers have $\psi_{j^*} = \bar{\psi}$ then because there is an efficient allocation for $(\bar{\psi}, \theta)$ above x_{j^*}, concavity of $\mathcal{V}$ in x implies that moving these types up to x_{j^*} benefits the principal. Similarly, if the switchers have $\psi_{j^*} = \underline{\psi}$ then moving these types down to x_{j^*} is beneficial. Finally, if the switchers have ψ_{j^*} interior then as the price of x_{j^*} is lowered, types just above ψ_{j^*} change their choice downward to x_{j^*} and types just below ψ_{j^*} change their choice upward to x_{j^*}. Arguing as in the derivation of r, the total effect of this has the sign of

$$\frac{\mathcal{V}(x_{j^*}) - \mathcal{V}(\underline{x}_{j^*})}{v_\psi(x_{j^*}) - v_\psi(\underline{x}_{j^*})} - \frac{\mathcal{V}(\bar{x}_{j^*}) - \mathcal{V}(x_{j^*})}{v_\psi(\bar{x}_{j^*}) - v_\psi(x_{j^*})}$$

where both sides are evaluated at the given θ and at ψ very close to ψ_{j^*}. This is strictly positive using Cauchy's mean value theorem and Assumption 5.

A Useful Restatement of $\mathcal{P}$. In the spirit of the second part of this section, replace ρ by the vector of incremental prices $p = (p_1, \ldots, p_J) \in \mathbb{R}^J_+$. Let $\mathcal{V}^j(\psi, \theta) \equiv \mathcal{V}(x_j, \psi, \theta) - \mathcal{V}(x_{j-1}, \psi, \theta)$. Define also the consumer's payoff of moving from x_{j-1} to x_j as $v_j(\psi, \theta) \equiv v(x_j, \psi, \theta) - v(x_{j-1}, \psi, \theta)$.

[24] Steps between x_0 and the lowest intermediate quality may be skipped, as may steps above the highest intermediate quality and x_J.

We can then re-express the principal's problem as

$$\max_{p\in\mathbb{R}^J_+,\chi} \sum_{j=1}^{J} \int_{\{(\psi,\theta)|x_j\leq\chi(\psi,\theta)\}} (\mathcal{V}^j(\psi,\theta) - (1-w_C)(v_j(\psi,\theta)-p_j))dH(\psi,\theta) \text{ s.t. } IC \tag{5}$$

That is, for each j, one tallies the incremental payoff the principal receives from moving the agent from $j-1$ to j over those agents who are allocated x_j or above. The first term in square brackets is the social surplus created by moving the consumer up an increment, while the second term tallies the principal's cost of the new consumer surplus.[25]

5 The Demand-Profile Approach

In this section, we explore what we can learn about our problem by applying the so-called demand-profile approach (*DPA*). *DPA* was pioneered by Wilson (1993) and forms the basis for the preponderance of empirical work on screening. We will see that when *DPA* is valid, our necessary conditions *NC* simplify substantially and provide additional economic insight into our problem. Moreover, our exploration of when *DPA* is valid yields novel and practically relevant conditions under which the necessary conditions are sufficient.

We proceed in four steps. Section 5.1 studies the economics of *DPA* when it is valid. Section 5.2 explains how DPA can fail to capture the full problem. Section 5.3 gives conditions for the exact and approximate validity of *DPA*. Section 5.4 connects the validity of *DPA* to sufficiency of the necessary conditions of the full problem.

Under *DPA*, one simply takes a given price schedule ρ, recasts it as an incremental price schedule p, and asks which consumers would be willing to pay p_j for the incremental quality increase from x_{j-1} to x_j on each of the J increments. That is, one takes the complicated object $\{(\psi,\theta)|x_j \leq \chi(\psi,\theta)\}$ in (5) and replaces it with $\{(\psi,\theta)|v_j(\psi,\theta) > p_j\}$. With this replacement, the sum in (5) decomposes to J separate items, each of which depends only on the corresponding single price. Thus for each j, we maximize

$$\hat{\Pi}^j(p_j) \equiv \int_{\{(\psi,\theta)|v_j(\psi,\theta)>p_j\}} \left(\mathcal{V}^j(\psi,\theta) - (1-w_C)(v_j(\psi,\theta)-p_j)\right) dH(\psi,\theta) \tag{6}$$

by choice of p_j. Let $Q^j(p_j) \equiv H(\{(\psi,\theta)|v_j(\psi,\theta) > p_j\})$ be the fraction of consumers willing

[25] To see this, note that for any j^* and (ψ,θ),

$$\mathcal{V}(x_{j^*},\psi,\theta)-(1-w_C)(v(x_{j^*},\psi,\theta)-\rho(x_{j^*})) = \mathcal{V}(x_0,\psi,\theta)-(1-w_C)v(x_0,\psi,\theta)+\sum_{j=1}^{j^*}\left(\mathcal{V}^j(\psi,\theta)-(1-w_C)(v_j(\psi,\theta)-p_j)\right)$$

where the first term is out of the control of the principal and hence can be discarded.

to pay p_j for increment j. Then, it is equivalent to choose q_j to maximize $\tilde{\Pi}^j(q_j)$ defined by $\tilde{\Pi}^j(Q_j(p_j)) = \hat{\Pi}^j(p_j)$.[26] We are thus left with the problem

$$(\tilde{\mathcal{P}}) \qquad \max_{q=(q_1,\ldots,q_J)} \sum_{j=1}^{J} \tilde{\Pi}^j(q_j).$$

Let $\tilde{q} = (\tilde{q}_1, \ldots, \tilde{q}_J)$ and $\tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_J)$ be the quantity and price vectors that solve $\tilde{\mathcal{P}}$. Say that *DPA* is *valid* if any solution to $\tilde{\mathcal{P}}$ is also a solution to $\mathcal{P}$.

Below, we first explore the economics that *DPA* unlocks when it is valid, or in other words, the economics of the problem $\tilde{\mathcal{P}}$. We then provide sufficient conditions under which *DPA* is either exactly or approximately valid. The analysis yields novel conditions under which our necessary conditions *NC* are also sufficient for a solution to $\mathcal{P}$.

5.1 The Economics of DPA

Recall that in $\tilde{\Pi}^j$, aggregate demand for quality increment j is given by $Q^j(p_j) = H(\{(\psi,\theta)|v_j(\psi,\theta) > p_j)$. Let the inverse of Q^j be P^j.[27] Further, let $\gamma_j(\psi,\theta) \equiv \gamma(x_j,\psi,\theta) - \gamma(x_{j-1},\psi,\theta)$ and $\eta_j(\psi,\theta) \equiv \eta(x_j,\psi,\theta) - \eta(x_{j-1},\psi,\theta)$. The principal's cost function C^j now has the simple form

$$C^j(q_j) = \int_{\{(\psi,\theta)|v_j(\psi,\theta)>P^j(q_j)\}} \gamma_j(\psi,\theta) dH(\psi,\theta),$$

and analogously for C^j_G. Consumer surplus is given by

$$CS^j(q_j) = \int_{\{(\psi,\theta)|v_j(\psi,\theta)>P^j(q_j)\}} \left(v_j(\psi,\theta) - P^j(q_j)\right) dH(\psi,\theta).$$

Thus,

$$\tilde{\Pi}^j(q_j) = w_C CS^j(q_j) + R^j(q_j) - C^j(q_j) - w_G C^j_G(q_j),$$

which has first-order condition

$$(7) \qquad w_C MCS^j(q_j) + MR^j(q_j) = MC^j(q_j) + w_G MC^j_G(q_j).$$

The left hand side of (7) is the principal's marginal benefit of serving an additional consumer at increment j. The right-hand side is principal's marginal cost. This first-order condition can be

[26] Prices below $\min v_j$ are suboptimal. Prices above $\max v_j$ earn zero. And, by Assumptions 1-4, between $\min v_j$ and $\max v_j$ demand is strictly decreasing, continuous, with values on all of $[0,1]$, and so is one-to-one with price.
[27] Since Q^j is strictly decreasing, P^j is defined by $P^j(Q_j(p_j)) = p_j$.

rewritten as

$$P^j(q_j)\left(1+(1-w_C)\frac{1}{\epsilon^j(q_j)}\right)=MC^j(q_j)+w_G MC_G^j(q_j), \tag{8}$$

where $\epsilon^j \equiv P^j/(P^j_{q_j}q_j) < 0$ is the price-elasticity of demand.[28] We thus have a familiar markup formula, adjusted for the fact that as the price falls (q_j increases), the principal places value w_C on transfers to the consumer and for the fact that the principal may put some weight on the marginal cost to the government.

The form of (8) yields several insights of economic interest. First are some simple comparative statics on the solution with respect to the weight put on consumer's utility and government.

Proposition 5 (Comparative Statics with Respect to Weights) *Let $H \in \mathcal{H}$. Except at corner solutions, the optimal quantity served at each increment increases in w_C, and if C_G^j is increasing in q_j for each j, then the optimal quantity served at each increment decreases in w_G.*

This result follows immediately, since as w_C increases, the *lhs* of (8) increases, from which it follows that the optimal quantity at each increment increases (and the price decreases) in the weight put on the consumer increases.[29] In Figure 2, the marginal benefit curve is a blend between the demand curve and the marginal revenue curve. It rotates upwards when w_C rises. When $w_C = 0$, the marginal benefit curve coincides with the marginal revenue curve. This case occurs for a self-interested monopolist and as a limit case for a social planner with a high cost of public funds. When $w_C = 1$ (as might occur if the principal is a social planner with no excess cost of funds) then the marginal benefit curve coincides with the demand curve. Similarly, the *rhs* rises as w_G increases, and so optimal quantity goes down and price rises.

Second, we can now derive stronger results with respect to exclusion and screening. Indeed, we can show that if $w_C < 1$, then $\tilde{q}_j < 1$ at *every* increment. The key is that the demand curve for each increment of quality becomes arbitrarily steep near $q = 1$, because to have a near-minimum value for v_j, the consumer has ψ close to $\underline{\psi}$ *and* θ close to minimizing $v_j(\underline{\psi}, \cdot)$, where by Assumption 3 the set of θ's that minimize $v_j(\underline{\psi}, \cdot)$ has zero mass. But then, raising the price a little from $P^j(1)$ has a first-order gain from consumers paying more but a second-order loss from consumers who no longer consume.[30]

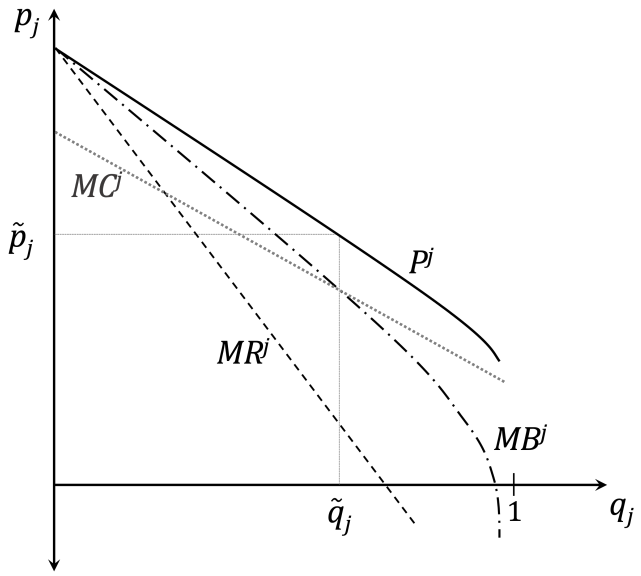
Proposition 6 (Incentives to Screen) *Let $H \in \mathcal{H}$. If $w_C < 1$, then $\tilde{q}_j < 1$ for all j.*

[28] In particular, $MCS^j(q_j) = -P^j_{q_j}(q_j)q_j$, where the Leibniz term is zero since $v_j(\psi,\theta) - P^j(q_j) = 0$ for new entrants.

[29] By standard monotone comparative statics results, this is true even if marginal benefit does not cross marginal cost once from above.

[30] In our insurance example, this effect is reflected by the fact that the marginal revenue curve diverges to $-\infty$ as q_j goes to 1.

Figure 2. Optimal Choice of q_j



Notes: The figure shows the inverse demand curve P^j, the marginal revenue curve MR^j, the principal's marginal cost curve $MC^j + w_G MC_G^j$ and marginal benefit curve MB^j in the market for incremental quality amount j. The optimal quantity $\tilde{q}_j$ obtains where the marginal benefit curve intersects the marginal cost curve.

The proof is in Appendix A.2. This result implies that in particular, if $w_C < 1$, then some consumers will be allocated x_0. It is tempting to conclude from the proposition that all qualities x_j for which $\tilde{q}_{j+1} < 1$ are allocated to some consumers, corresponding to the results in Section 4 on trade at all intermediate quality levels. But, this need not be true, since it is possible in principle that all consumers who are told to take increment j are also told to take increment $j+1$, and no type is told to stop at x_j. Based on our analysis in the next section, we would argue that this is very unlikely if *DPA* is valid.

On the other hand, if $w_C = 1$, then the solution to $\tilde{\mathcal{P}}$ will set $\tilde{q}_j = 1$ if marginal cost $MC^j + w_G MC_G^j$ lies everywhere below the demand curve (and in some cases with multiple crossings as well). Hence, the *DPA* solution can be consistent with the presence of an interval of low qualities that are not assigned to anyone.[31] If $MC^j + w_G MC_G^j$ lies everywhere below the demand curve, then when $w_C \to 1$, $\tilde{q}_j \to 1$ as well. A monopolist will thus engage in "more screening" than a social planner, in that the social planner can be either skipping low qualities ($w_C = 1$) or allocating just a few types to low qualities, whereas the monopolist may use these qualities heavily.

[31] That there is nothing non-generic about several of the maximization problems having a corner solution at either boundary $\tilde{q}_j = 0$ or $\tilde{q}_j = 1$.

5.2 Failures of *DPA*

It is possible that in the solution to $\tilde{\mathcal{P}}$, a given consumer has $v_1 < p_1$, so that their payoffs decrease as they move from x_0 to x_1 but has $v_2 > p_2$ so that their payoffs increase as they move from x_1 to x_2. As such, their payoffs fail to be "quasiconcave," in x, and it is unclear which quality they should be allocated to. Further, if they are allocated x_0, then the calculation in $\tilde{\Pi}^2$ will be incorrect in having included them, while if they are allocated x_2 or above, then the calculation in $\tilde{\Pi}^1$ will be incorrect in having excluded them. Hence, for *DPA* to be valid, the first thing that must be true is for each (ψ, θ) in the support of H, $\tilde{p}$ is *quasiconcave consistent (QC)* in that $v_j(\psi, \theta) - \tilde{p}_j$ single-crosses zero from above. When $\tilde{p}$ is *QC* for (ψ, θ) then the allocation to (ψ, θ) is unambiguous, and (ψ, θ) is included in the set of types with $v_j(\psi, \theta) - \tilde{p}_j > 0$ if and only if they are allocated x_j or above.[32] We thus take it as a necessary condition for $\tilde{p}$ to solve $\mathcal{P}$ that $\tilde{p}$ is *QC* for all consumers.[33] Finding conditions for $\tilde{p}$ to be *QC* for all consumers is a notoriously hard problem, which is related to our discussion of sufficiency at the end of Section 3.

Another thing that can go wrong with *DPA* is that the principal might be better off to "bundle" two consecutive increments, forcing consumers into an all-or-nothing choice. For a simple example, assume that $J = 2$ and that half of consumers have $v_1 = v_2 = 1$ while the other half have $v_1 = 1.6$ and $v_2 = .4$. Assume that costs are zero, and that the principal cares only about their own profit. Then, the demand-profile approach has $p_1 = p_2 = 1$. This gives profit $3/2$ since the first type of consumer will choose x_2 while the second type will choose x_1. If instead, the principal removes x_1 from the available set while setting $p_2 = 2$, then both types choose x_2, yielding profit 2.[34]

5.3 When is *DPA* Valid?

We now turn to the question of when *DPA* is valid. We begin with a sufficient condition termed *non-crossing demands (NCD)* that is similar to ones in Wilson (1993), Deneckere and Severinov (2017), and Pass and Halim (2025). It states that we can order consumers in such a way that higher-ordered consumers have higher willingness to pay for each increment. Thus, if we graph the "demand curve" for quality of each individual as a function of quality increment j, these functions do not cross. The condition of non-crossing demands is strong, and unlikely to be satisfied in practice. Indeed, when *NCD* holds, then as we will heavily exploit, the consumer's values are

[32] Since $H \in \mathcal{H}$, ties occur for a zero-measure subset of types.

[33] Going back to the discussion after Proposition 6, if *DPA* is valid then it must be that $\tilde{q}_j \geq \tilde{q}_{j+1}$ for all j, since by definition of *QC*, all types with $v_{j+1} > \tilde{p}_{j+1}$ also have $v_j > \tilde{p}_j$. Thus, if $\tilde{q}_j$ is interior and no consumers are allocated x_j, we must have $\tilde{q}_j = \tilde{q}_{j+1}$. For this to happen with *QC* satisfied, first, the set of types with $v_{j+1} > \tilde{p}_{j+1}$ must be identical to the set of types with $v_j > \tilde{p}_j$, and second, (8) happens to reach zero at exactly the same quantity for j and $j+1$. This would be a remarkable coincidence, and will be non-generic under any reasonable notion of genericity. We conclude that if $w_C < 1$ and x_{j^h} is the highest quality allocated to some types, then except in knife-edge cases, every quality below x_{j^h} is also allocated to some consumers.

[34] We thank Michael Whinston for this example, which builds on the standard logic of bundling.

inherently one dimensional (although their costs need not be). Because of this, *NCD* is unlikely to be satisfied in many examples where types are multi-dimensional in economically significant ways. That said, our empirical investigation in the health insurance setting reveals an interesting path forward. In that setting, we find that even if *NCD* fails to hold exactly, it may still do so to a high degree of approximation, and moreover, the solution to $\mathcal{P}$ may be very well approximated by the solution to $\tilde{\mathcal{P}}$.

Linking these empirical observations, we provide two results showing that if the population "nearly" satisfies *NCD*, then *DPA* continues to be either valid or nearly valid (depending on the notion of "nearly"), provided that the allocation is *decreasing*. We also provide an explicit expression for how close is "close enough." Finally, we provide primitive conditions on the distribution on types and on the value and cost functions that ensure both *NCD* and that the allocation is decreasing. Together, these results provide a strong basis to the applied researcher using *DPA*, as well as novel theoretical underpinnings for the approach. The results also provide a set of primitives under which the necessary conditions from Section 3 are also sufficient (and substantially simplify).

Exactly Non-Crossing Demands Say that H satisfies *non-crossing demands (NCD)* if for all (ψ', θ') and (ψ'', θ'') in the support of H, $v_1(\psi', \theta') > v_1(\psi'', \theta'')$ if and only if $v_j(\psi', \theta') > v_j(\psi'', \theta'')$ for all j. The following proposition tells us that if H satisfies *NCD* , then *DPA* holds whenever it yields a decreasing solution $\tilde{q}$. And, this holds even if one largely dispenses with the conditions inherent in $\mathcal{H}$. Let $\mathcal{H}_A$ be the set of distributions with the property that for each $x' < x''$, $v(x'', \cdot) - v(x', \cdot)$ has a continuous density which has support an interval. It is easy to see that $\mathcal{H} \subseteq \mathcal{H}_A$. See Appendix A.3 for details.

Proposition 7 (*DPA* with *NCD*) *Let $w_C \leq 1$, and let $H \in \mathcal{H}_A$ satisfy NCD. If $\tilde{q}_j$ is decreasing in j, then DPA is valid.*

The idea of the proof (see Appendix A.3) is that if H satisfies *NCD*, then consumers can be ordered by the fraction of types with higher incremental values than themselves. That is to say, their "percentile rank" in the distribution of willingness to pay for incremental quality. But then, by the definition of *NCD*, the solution to $\tilde{\mathcal{P}}$ will satisfy *QC* if and only if $\tilde{q}_j$ is decreasing in j. Similarly, in any *IC* mechanism, if a mass q_j of types are allocated x_j or above, then the types in question will be those with the highest incremental values. Incentive compatibility then ties down that each incremental price is equal to the incremental value of the boundary type, which is to say $P^j(q_j)$, and this allows us to recast $\mathcal{P}$ as

$$\max_{\{q|q_j \geq q_{j+1}\}} \sum_{j=1}^{J} \tilde{\Pi}^j(q_j). \tag{9}$$

But then, $\tilde{\mathcal{P}}$ is a relaxation of $\mathcal{P}$ (it differs only in not imposing that q is decreasing). So, if $\tilde{q}_j$ is decreasing, then it solves $\mathcal{P}$.

An immediate corollary of this result is that if there is just one increment of quality available, *DPA* holds generically. When $J = 1$, *NCD* holds vacuously and $\tilde{q}$ is trivially decreasing, so Proposition 7 applies. This result is of practical relevance because many empirical settings feature a binary choice between an outside option and a single inside option, in which case *DPA* imposes no restrictions beyond the standard assumptions. With two or more increments, *NCD* requires that consumers can be consistently ranked across all margins simultaneously.[35]

Approximately Non-Crossing Demands Let us now explore conditions under which *DPA* holds when *NCD* is approximately satisfied. Say that H is *regular* if, for each j, $\tilde{\Pi}^j$ is strictly quasiconcave.[36] Say that q is *strictly decreasing* if q_j is strictly decreasing in j where it is interior. Say that *DPA* is ε-valid for H if the prices and profits in $\tilde{\mathcal{P}}$ are within ε of those from $\mathcal{P}$ and the fraction of consumers for whom *QC* holds at both p and $\tilde{p}$ is at least $1 - \varepsilon$.

Theorem 2 (*DPA* Near *NCD*) *If $H_0 \in \mathcal{H}_A$ is regular and satisfies NCD, and if DPA yields a strictly decreasing solution at H_0, then DPA is ε-valid for all $H \in \mathcal{H}_A$ sufficiently close to H_0 in the weak topology and (exactly) valid if in addition the support of H is within a sufficiently small neighborhood of that of H_0.*[37]

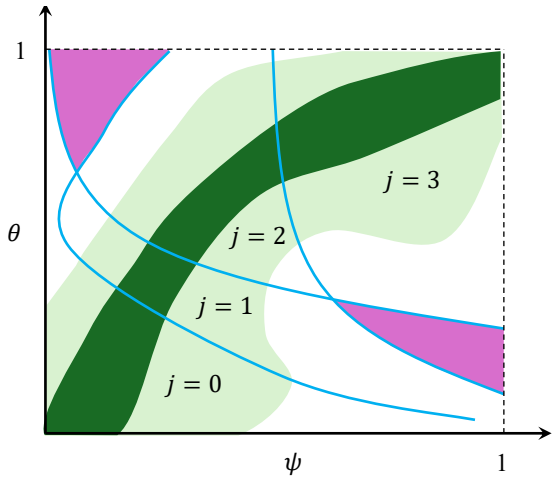
To see the idea of the proof, consider Figure 3. The dark-green region is the support of an H_0 satisfying *NCD* and with successive quantities allocated strictly fewer consumers. The blue lines delineate those types who optimally accept each increment of quality and are thus strictly apart on the dark-green region, and so the dark-green region is strictly apart from the pink regions where *QC* fails and the allocation is incoherent. For example, in the left pink region, consumers are told to purchase the increment from x_1 to x_2, but not the increment from x_0 to x_1. As the figure suggests, the pink regions can be substantially away from the dark-green region, so that one can choose a light-green region substantially bigger than the dark-green region but substantially away from the pink regions. If H is not too different from H_0 then the optimal prices, and hence the blue lines, will not have shifted too much, and so the pink regions will not have expanded so far as to cross into the light-green region. But then, H, which has most of its weight in the light-green region will satisfy *DPA* almost exactly. If in addition the support of H is within the light-green region, then *DPA* is exact. See Appendix A.3 for details.

[35] Geruso et al. (2023) impose precisely this condition (their Assumption 2) in a setting with two increments, where it ensures that consumers sort only between adjacent generosity levels. Our *NCD* condition generalizes this adjacent-margin sorting property to an arbitrary number of quality levels.

[36] That is, the marginal benefit side of (7) crosses the marginal cost side at most once and from above, so that $\tilde{q}_j$ is uniquely defined by (7) or by a corner solution.

[37] The weak topology, recall, is that in which the expectations of bounded continuous functions converge.

Figure 3. Illustration of Theorem 2



Notes: The dark-green region is the support of an $H_0 \in \mathcal{H}_A$ with *NCD* and strictly decreasing solution. The blue lines indicate the allocation. They are strictly separated on the dark-green region, which is thus distinct from the pink regions, where the blue lines cross and the allocation is incoherent. When H is close to H_0, then the pink regions will not have grown much, so that *QC* holds for all types in the light-green region. Thus, since H puts most of its weight in the light-green region, *DPA* is arbitrarily close to valid, and exactly so if the support of H is in the light-green region.

In Section 5.4 we provide primitives for a distribution to be regular and for *DPA* to yield a strictly decreasing solution. There are many such distributions. The two conditions are also trivial to verify numerically in applications. Thus, if one starts from *all* those regular distributions in $\mathcal{H}_A$ which have *NCD* and where the solution is strictly decreasing, then *DPA* holds anywhere in $\mathcal{H}_A$ that is close to any one of these distributions in both the weak topology and support. The ε version of the theorem is highly relevant in empirical settings where there may be substantial correlation of values but a small number of outliers.

Together, the message of this section for both the applied researcher and the theorist is that if the type distribution has multidimensional support but values for different quality increments are sufficiently correlated, then *DPA* will perform well.

5.4 The Necessary Conditions are Sufficient

Finally, let us make the connection to the necessary conditions of Section 3. Recall that our necessary conditions allowed for the possibility that an agent jumped past several quality levels at a time. Let the *simplified necessary conditions* be those that obtain when one simply assumes that $\underline{x}_j(\rho,\theta) = x_{j-1}$ and $\bar{x}_j(\rho,\theta) = x_j$, and let ψ_j be determined as the moment at which the agent is willing to move from x_{j-1} to x_j. Then, the simplified necessary conditions are equivalent

to (7) holding for each j. Thus, under regularity solving the simplified necessary conditions is equivalent to solving $\tilde{\mathcal{P}}$ and so if *DPA* is valid then the simplified necessary conditions in $\mathcal{P}$ are also sufficient.[38] The following is thus immediate.

Theorem 3 (Sufficiency) *Let $H \in \mathcal{H}$ be regular and let DPA be valid at H. Then, the simplified necessary conditions of problem $\mathcal{P}$ are also sufficient.*

Thus, in particular, the simplified necessary conditions are sufficient if the predicates to Theorem 2 hold. To tie those predicates tightly to primitives, Appendix A.3 provides a definition for a distribution to be *aligned*, and shows (Theorem 6) that when $H_0 \in \mathcal{H}_A$ is aligned, then it is regular and satisfies *NCD*, and *DPA* yields a strictly decreasing solution. The aligned distributions have support on structured one-dimensional paths through $\Psi \times \Theta$, which we index by $\tau \in [0, 1]$. Three things are needed. First, v_x increases as we increase τ, which means that higher τ have a higher willingness to pay for quality. Given that $v_{x\psi} > 0$, one way for this to hold is if the path is headed sufficiently in the direction of higher ψ compared to other components. Second, if we let G be the cumulative for τ, then we need that the product of $\mathcal{V}^j/(v^j)_\tau$ and the hazard rate $g/(1-G)$ increases in τ, where by $(v^j)_\tau$ we mean the derivative of the composite function where τ determines a point in $\Psi \times \Theta$ and then v is evaluated at that point. In Online Appendix B.5 we provide conditions under which the first component is increasing regardless of the choice of $\{x_j\}$. But, even if the first term is not increasing, the hazard rate increases under very mild conditions, and by choice of G can be made to do so arbitrarily quickly so as to ensure this condition. The condition implies that H_0 is regular. Finally, we require that $\mathcal{V}_x/(v_x)_\tau$ is strictly decreasing in x, which will imply that q is strictly decreasing in j. Any aligned distribution fails Assumption 1 and so is not an element of $\mathcal{H}$. But, it is an element of $\mathcal{H}_A$, and this suffices for our purposes.

6 Some Extensions

We have derived our necessary conditions for optimality under the assumption of a finite set of quality levels. This is mainly for technical convenience and clarity, since the analysis can be extended to an interval of quality levels. We now summarize two extensions of our results, both related to large sets of quality levels. The first extension provides the necessary conditions for optimality when there is a continuum of quality levels. The second extension connects the finite and continuum cases: under mild assumptions, the optimal solution in the finite case converges to that in the continuum case as the number of quality levels grows without bound.

[38] The proviso that we use the simplified necessary conditions is that there might be other solutions to the necessary conditions that involve complicated jump behavior. When *DPA* is valid, those solutions are not optimal.

A Continuum of Quality Levels While the finite case unlocks many directions of development, it is also interesting to consider the case where the principal can offer all qualities in $[0,1]$. Here we describe our results. See Online Appendix Sections B.3–B.4 for formalities. To begin, fix ρ and then suppress it in the notation, and let $x^0(\psi,\theta)$ be the set of best responses for (ψ,θ). Fix θ and x and let $\acute{\psi}$ be the point at which $x^0(\cdot,\theta)$ passes from below x to above x. If $x^0(\cdot,\theta)$ is continuous at $\acute{\psi}$ then define

$$r(x,\theta) \equiv \frac{\mathcal{V}_x(x,\acute{\psi},\theta)}{v_{x\psi}(x,\acute{\psi},\theta)}.$$

If $x^0(\cdot,\theta)$ jumps at $\psi^0(x,\theta)$ from x_l to x_h then define

$$r(x,\theta) \equiv \frac{\mathcal{V}(x_h,\acute{\psi},\theta) - \mathcal{V}(x_l,\acute{\psi},\theta)}{v_\psi(x_h,\acute{\psi},\theta) - v_\psi(x_l,\acute{\psi},\theta)}.$$

These are natural generalizations of the finite case. Say that x is a *bunching point* if for a positive H-measure set of θ, $x^0(\cdot,\theta) = x$ on an interval $(\psi_l(x,\theta),\psi_h(x,\theta))$ of positive length.

Theorem 4 (Optimality Condition One) *Let $H \in \mathcal{H}$ and let (ρ,χ) be optimal. If $x \in [0,1]$ is not a bunching point then,*

$$\int_\Theta [(1-w_C)\,(1-H(\psi(x,\theta)|\theta)) - r(x,\theta)h(\psi(x,\theta)|\theta)]h(\theta)d\theta = 0, \tag{10}$$

while if x is a bunching point, then (10) holds if one consistently uses either $\psi_l(x,\theta)$ or $\psi_h(x,\theta)$.

For each θ, the term in square brackets is exactly the optimality condition in the "regular case" if H was degenerate at θ so we are in a one dimensional setting. Thus, (10) requires that we optimally average across θ the standard efficiency versus rent extraction tradeoff at that θ.

The proof is in the spirit of the main perturbation we used before, but to deal with the continuum, we first show that without loss ρ can be taken to be absolutely continuous with a slope bounded away from zero, and then make ρ either a little steeper or a little shallower near x. Note that this theorem is a natural generalization of Theorem 1. Here however, there is another optimality condition to consider. Let x be a bunching point, and, in the spirit of ironing, let us consider moving the agents who use x to $x+\varepsilon$ or $x-\varepsilon$ instead.[39] Say that a bunching point x is a *point of continuous allocation* (*POCA*) if for almost all θ, $x^0(\cdot,\theta)$ does not jump at either $\psi_l(x,\theta)$ or $\psi_h(x,\theta)$.

[39] In the finite case, we could also have considered the effect of moving one of the qualities up or down a bit, but it seems economically less interesting to have a finite number of qualities mandated but have substantial control over what those qualites are.

Theorem 5 (Optimality Condition Two) *Let $H \in \mathcal{H}$ and let (ρ, χ) be optimal. If $x \in [0,1]$ is a bunching point that is POCA, then*

$$\int \left[\int_{\psi_l(\theta)}^{\psi_h(\theta)} \left(\mathcal{V}_x(\hat{x}, \psi, \theta) - (1 - w_C) v_{x\psi}(\hat{x}, \psi, \theta) \frac{1 - H(\psi|\theta)}{h(\psi|\theta)} \right) h(\psi|\theta) d\psi \right] h(\theta) d\theta = 0 \tag{11}$$

Here, for each θ, the square-bracketed term is the standard ironing condition when H is degenerate at θ. So, once again, the condition asks us to average a well-understood set of tradeoffs across different θ. The perturbations that drive this result either modify ρ to very steep from a point $x' < x$ up to x, incentivizing agents who chose x to choose x' instead, or modify ρ to flat from x to a point $x'' > x$ up to x, incentivizing agents who chose x to choose x'' instead. These perturbation necessarily also modify the choices of agents who are not choosing x. The key to the proof is to use Theorem 4 and *POCA* to show that these effects net out to zero impact.

Connecting the Finite and Continuum Cases The following convergence result will prove useful in practical application of our analysis. Under some technical assumptions, the optimal solution under a fixed set of potential qualities converges to the optimal solution under a continuum of potential qualities as the number of qualities in the fixed set grows large. Formally, this result requires a suitable notion of convergence of sets of price schedules and a metric on those sets, and then use these notions to show that both the consumer's indirect utility and the principal's maximized payoff are continuous functions. We relegate the technical details to Appendix A.4.

Intuitive as it may be, the convergence result has two very useful implications. First, for numerical purposes, the modeler can use any reasonable set of fixed potential qualities, and be confident that they get a result that approximates what the principal can achieve with a continuum thereof. The details of how the sequence of sets of price functions is constructed simply do not matter, as long as the set of qualities grows dense. We will return to this point in particular in our numerical application below. Second, this result provides theoretical flexibility. If the principal can offer a sufficiently rich set of fixed quality levels, then there is a vanishing amount of value added by also allowing her to optimally vary these quality levels. We can therefore work in the case of a (large) fixed set of qualities or in the continuum, whichever is more convenient.

7 Application to Health Insurance

As an application of this framework, we study the problem facing a health insurance menu designer pricing a set of vertically differentiated insurance contracts. Within this application, we can study the insurer's problem numerically using empirical estimates of the key objects. This numerical analysis allows us to evaluate quantitatively the theoretical predictions derived in Section 4,

assess the empirical plausibility of the conditions under which *DPA* is valid, and to demonstrate the usefulness of *DPA* for gaining intuition about this problem.

7.1 Applied Model

Consider a model of a health insurance market in which a monopolist insurer (principal) chooses a set of vertically ordered contracts to offer and their associated premiums. Heterogeneous consumers then select a single contract, incur health shocks, and choose their subsequent healthcare utilization. Consumers have multidimensional private information at the time they choose an insurance contract. Realized health is also private information, allowing for ex post moral hazard.

Consumers. There is a continuum of consumers. Each consumer is an expected-utility maximizer and is privately informed about her taste for healthcare ω, her risk aversion parameter ψ, and her distribution F over potential health states l, which has density f on bounded support $[0, \bar{l}]$. As above, we denote the consumer's type by (ψ, θ), where now $\theta = (\omega, F)$. The distribution of (ψ, θ) is given by a joint cdf H on $[0, \bar{\psi}] \times [0, \bar{\omega}] \times \Delta([0, \bar{l}])$.[40] To make it a finite-dimensional type space as assumed in previous sections, we will parametrize F by a finite-dimensional vector with values in a compact rectangle in $\mathbb{R}^n$.

If the consumer chooses a dollar amount $a \in [0, \bar{a}]$ of healthcare utilization when her health state is l and her taste for healthcare is ω, she enjoys a utility level which in dollar terms is given by $b(a, l, \omega)$, where b is strictly increasing in a, with $b_{aa} < 0$ and $b_{al} > 0$. That is, the consumer has declining marginal utility for healthcare, but marginal utility is higher when she has worse health. Her risk preferences are then captured by a CARA utility function over money-metric payoffs. Absent insurance, she has expected utility $\int -e^{-\psi[b(a,l,\omega)-a]} dF(l)$.

Insurance Contracts. An insurance contract consists of an out-of-pocket cost function $c(a, x) \in [0, a]$ that specifies how much the consumer pays for different levels of healthcare utilization. Contracts are indexed by a scalar $x \in [0, 1]$, with $c(0, x) = 0$ for all x, $0 \leq c_a \leq 1$, $c_{aa} \leq 0$, and $c_{ax} \leq 0$. Contracts are thus vertically differentiated, with higher x corresponding to higher coverage.

Given a contract x, a health state realization l, and taste for healthcare ω, the consumer chooses an optimal level of healthcare utilization $a^*(l, x, \omega) \equiv \arg\max_a (b(a, l, \omega) - c(a, x))$. Let $z(l, x, \omega) \equiv b(a^*(l, x, \omega), l, \omega) - c(a^*(l, x, \omega), x)$ be the consumer's money-equivalent payoff from optimal healthcare utilization given (l, x, ω).

Consumer Valuation. Since premiums are deterministic and CARA preferences are translation invariant, we can represent the consumer (ψ, θ) choice as maximizing the money-metric certainty

[40] Here $\Delta(B)$ denotes the set of distributions over the set B.

equivalent $v(x, \psi, \theta) - t$, where v is given by

$$v(x, \psi, \theta) \equiv -\frac{1}{\psi} \log \int e^{-\psi z(l,x,\omega)} dF(l), \tag{12}$$

if $\psi > 0$ and $\int z(l, x, \omega) dF(l)$ if $\psi = 0$. Her willingness to pay for contract x relative to the base contract x_0 is then $v(x, \psi, \theta) - v(x_0, \psi, \theta)$, and more generally, her willingness to pay for the jump from any contract x to x' is $v(x', \psi, \theta) - v(x, \psi, \theta)$. Faced with a menu of (x, t) pairs, the consumer chooses the contract that maximizes $v(x, \psi, \theta) - t$.[41] This valuation function provides the mapping from the general model of the preceding sections to the health insurance setting. It is easy to verify that Willingness to pay for insurance is strictly increasing in contract quality x and increasing faster at higher levels of risk aversion. That is, $v_x > 0$ and $v_{x\psi} > 0$.

The Government and the Insurer. The government exogenously provides a base level of insurance x_0. The insurer is risk neutral and is a price-setter. As in the general model, the insurer's objective is governed by weights $w = (w_C, w_P, w_G)$ on consumer surplus, insurer profits, and government spending, respectively. The insurer's financial cost of providing contract x to type (ψ, θ), net of government reimbursement for base coverage, is $\gamma(x, \theta)$, and the government's financial cost is $\eta(x, \theta)$. Note that in this application, costs depend on θ but not on ψ. The insurer's problem is

$$\max_{\rho \in \mathbb{P}, \chi} \int S\left(\rho(\chi(\psi, \theta)), \chi(\psi, \theta), \psi, \theta\right) dH(\psi, \theta) \quad \text{s.t.} \quad IC,$$

where $S(t, x, \psi, \theta) = w_C(v(x, \psi, \theta) - t) + t - \gamma(x, \theta) - w_G \eta(x, \theta)$. As above, w_P is normalized to one. We focus our attention in the application on three insurers: a utilitarian social planner with no excess cost of public funds, corresponding to $w = (1, 1, 1)$; a planner with a 25 percent excess cost of funds, corresponding to $w = (0.8, 1, 1)$; and a monopolist, corresponding to $w = (0, 1, 0)$.

7.2 Numerical Analysis

Consumers. We simulate a population of 10,000 households using a distribution of demographics chosen to match the under-65 US population and parameter estimates reported in Marone and Sabety (2022). The health state distribution F is assumed to have a shifted log-normal distribution such that $\log(l + \kappa) \sim N(\mu, \sigma^2)$. The money-metric payoff from healthcare utilization in a given

[41] Note that the certainty equivalent in equation (12) reflects the consumer's risk aversion, so the welfare gain from insurance includes the reduction in the consumer's risk premium. Aggregating these money-metric payoffs across consumers values a marginal dollar equally for all consumers. This approach abstracts from redistributive concerns, consistent with the standard efficiency criterion in the health insurance literature (e.g., Einav et al., 2010; Azevedo and Gottlieb, 2017).

health state is parameterized such that $b(a, l, \omega) = (a - l) - \frac{1}{2\omega}(a - l)^2$.[42]

Appendix Table B.1 summarizes the characteristics of the simulated population. The average household would have total healthcare spending equal to \$11,164 under a full insurance contract, but only \$9,810 under a null contract, reflecting moral hazard. Facing an equal odds gamble between \$0 and \$100, the average household would have a certainty equivalent of \$48.9, reflecting risk aversion.

Insurance Contracts. We consider a set of contracts that are piecewise linear, with a deductible, coinsurance region, and out-of-pocket maximum design. We suppose that the base level of coverage x_0 is a "Catastrophic" contract with a deductible and out-of-pocket maximum of \$10,000. Our baseline set of potential contracts is depicted in Figure B.2. Because they roughly correspond to the levels of coverage available on the Affordable Care Act exchanges, we refer to the contracts between Catastrophic and full insurance as Bronze, Silver, and Gold.[43] As will become clear, the returns to allowing an increasingly "dense" contract space are economically small.

Convergence. Proposition 9 in Appendix A.4 states that an insurer's payoff when restricted to a finite set of potential contracts will converge to its unrestricted counterpart as the number of contracts grows. It is silent, however, on how quickly this will occur. We illustrate and investigate this result by computing optimal menus on an increasingly dense set of allowable contracts. Figure B.2 depicts a set of five allowable contracts, spaced at \$2,500 out-of-pocket maximum intervals between the minimum and maximum levels of coverage. We increase (and decrease) the density of this potential contract space by varying the number of contracts used to span this range. We move from just two contracts (in which case there is just Catastrophic and full insurance) to 65 contracts (in which case 15 contracts are added between each of the five original contracts).[44]

For each set of potential contracts, we solve for the optimal menu that would be offered by our three focal principals: a social planner with no excess cost of funds, a planner with a 25 percent excess cost of funds, and a monopolist.[45] Figure B.3 plots insurer payoffs as a function of the number of contracts in the potential contract space. While insurer payoffs are of course increasing in contract density, in practice the returns to additional density are small. We find that after five contracts, the gains from moving to 65 contracts do not exceed \$22 per household per year for

[42] Although it is not strictly increasing in a, it is so when restrictive to the relevant domain $0 \leq a \leq l + \omega$.

[43] The contracts' deductibles, coinsurance rates, and out-of-pocket maximums are: \$5,846, 40%, \$7,500 for Bronze; \$3,182, 27%, \$5,000 for Silver; and \$1,125, 15%, \$2,500 for Gold. In our population of consumers, the actuarial value of the five contracts are: 0.38, 0.47, 0.60, 0.78, and 1.00.

[44] We increase the set of allowable contracts by successively adding a contract between each pair of adjacent contracts. We proceed in this iterative manner so that under successively "dense" contract spaces, all previously allowable contracts remain allowable.

[45] Optimal menus are calculated using a numerical algorithm that relies on the necessary condition derived in Section 3. The algorithm is described in detail in Online Appendix B.7.

any insurer. After nine contracts, gains do not exceed \$6.[46]

There are, however, economically meaningful gains from moving between two and five contracts. Over this range, the social planner facing an excess cost of funds can increase social surplus by \$174 per household per year, and a monopolist can increase its profits by \$279. For the social planner, these gains reflect the ability to find a plan that more closely matches the tastes of consumers in the population. For the monopolist, these gains reflect this same increase in potential gains from trade, as well as the ability to more effectively screen consumers and thereby extract greater rents from the market. Our results suggest that while only a modest number of contracts are needed to closely approximate the limiting environment, there are potentially meaningful consequences of over-restricting the contract space. Of course, the precise number of contracts at which payoffs flatten out may vary across settings, in particular with the size of the range between base coverage and full insurance.

Consistent with Proposition 9, we also find that the optimal premium schedules and therefore the optimal allocations themselves converge as the density of the contract space increases. In the case of the monopolist insurer, consumer surplus also converges alongside producer surplus. Online Appendix Figure B.5 depicts the convergence of allocations. As the density of the contract space increases, the insurers "fill in" in the neighborhood of their desired allocation under a sparser contract space. All numerical results are thus quite robust to the density of the contract space.

Optimal Menu and Performance of *DPA*. Proceeding with five potential contracts, Table 1 now reports the optimal price schedule and associated allocations for each of our three focal insurers. These results are reported in the first row in each section, labeled *Solution to* $\mathcal{P}$. These price schedules achieve the maximal insurer payoff $\Pi \equiv \int S\, dH$ over admissible price vectors $\rho \in \mathbb{P}$. Consistent with the implications derived from the necessary conditions in Section 4, both the monopolist and the planner with excess cost of funds (both having $w_C < 1$) allocate a strictly positive mass of consumers to the Catastrophic contract (Proposition 2), excluding these consumers from the market for additional coverage. All three insurers allocate a strictly positive mass to the Gold contract (consistent with Proposition 3).[47] Finally, consistent with Proposition 4, each of the two insurers with $w_C < 1$ uses every contract between Catastrophic and Gold, excluding a positive mass of consumers at every level between x_0 and x^h.

The second row in each section, labeled *Solution to* $\tilde{\mathcal{P}}$ then reports the solution to problem $\tilde{\mathcal{P}}$, using the *DPA* as described in Section 5. Even without relying on the analytical results developed in Section 5.3, it is immediately clear from the numerical results that *DPA* provides a

[46] These results are consistent with both Marone and Sabety (2022) and Ho and Lee (2023), who find that only a limited number of contracts are sufficient to capture almost all the available surplus in their settings.

[47] However, no consumers are allocated to Full insurance, indicating that in this population, the threshold x^h at which moral hazard overwhelms risk protection occurs at the Gold contract.

Table 1. Performance of the Reformulated Problem

	Premiums *$000s*					Allocations *Pct. of households*					Insurer Payoff *$000s*
Insurer	x_0 Cstr.	x_1 Brnz.	x_2 Slvr.	x_3 Gold	x_4 Full	x_0 Cstr.	x_1 Brnz.	x_2 Slvr.	x_3 Gold	x_4 Full	Π
Social planner, $w = (1, 1, 1)$											
Solution to $\mathcal{P}$	–	0.25	0.37	0.63	3.12	<0.01	–	–	1.00	–	1.897
Solution to $\tilde{\mathcal{P}}$	–	0.10	0.28	0.73	3.23	–	<0.01	<0.01	1.00	–	1.897
Social planner, $w = (0.8, 1, 1)$											
Solution to $\mathcal{P}$	–	1.42	2.80	4.64	7.13	0.14	<0.01	0.13	0.72	–	1.750
Solution to $\tilde{\mathcal{P}}$	–	1.30	2.78	4.62	7.13	0.13	0.03	0.12	0.72	–	1.746
Monopolist, $w = (0, 1, 0)$											
Solution to $\mathcal{P}$	–	1.93	4.05	6.38	8.90	0.37	0.09	0.26	0.29	–	0.798
Solution to $\tilde{\mathcal{P}}$	–	1.93	4.05	6.38	8.89	0.37	0.09	0.26	0.29	–	0.798

Notes: The table reports the premium schedules ρ and $\tilde{\rho}$ chosen by insurers with different objective functions $w = (w_C, w_P, w_G)$ when solving the two formulations of the menu design problem: the true problem $\mathcal{P}$ and the reformulated problem $\tilde{\mathcal{P}}$. The table also reports the associated allocations and insurer payoffs. The insurer payoff Π is the objective of problem $\mathcal{P}$, meaning consumers globally optimize with respect to prevailing premiums. Payoffs are expressed on a per household per year basis, and are measured relative to the allocation of all consumers to the Catastrophic contract x_0.

close approximation to the true solution. The natural question is of course *why*, and it is here that the analytical results help shed light.

Theorem 2 establishes that *DPA* is approximately valid when the population distribution of demand for insurance nearly satisfies non-crossing demands and the solution to $\tilde{\mathcal{P}}$ exhibits decreasing quantities. We can provide direct numerical support for both conditions in our data. The first, exact non-crossing demands, cannot be confirmed in a finite population, so we instead measure how close our population lies to it. *NCD* requires an ordering of consumers under which higher-indexed consumers have higher incremental willingness to pay at every margin simultaneously. We measure our population's distance from this ideal by computing the smallest perturbation of consumers' incremental willingness to pay (*WTP*) that admits such an ordering. The mean absolute perturbation is 4.3 percent of each margin's range. In other words, the typical consumer's incremental *WTP* at any given increment needs to shift by only about 4 percent to achieve exact *NCD*.[48] The second premise—decreasing quantities, $q^j \geq q^{j+1}$—can also easily be verified by examination of the solution to problem $\tilde{\mathcal{P}}$, and in our setting holds for each of our focal insurers. Together, these two conditions support that consumers allocated to higher increments

[48] For each candidate ordering of consumers, the projection decomposes into independent isotonic regressions, one per increment. We search over orderings via alternating minimization from multiple starting points, minimizing the root mean squared error. Other measures of near-*NCD* tell a similar story: only 15 percent of all pairwise consumer comparisons exhibit a crossing, and Kendall's coefficient of concordance for consumers' incremental willingness to pay across the four margins is 0.889. This strong correlation across margins suggests that consumer types are approximately ordered along a one-dimensional path through the type space.

of coverage are a subset of those allocated to lower increments of coverage, such that the resulting allocations are coherent.

In this population and for this set of potential contracts, *DPA* thus provides an excellent approximation to the solution to problem $\mathcal{P}$. All of the analysis in Section 5 is therefore applicable in thinking about the solution. Moreover, our convergence analysis in Section 7.2 establishes that this set of five potential contracts is a good approximation of the solution under as many as 65 potential contracts, suggesting that even if *DPA* were to perform more poorly if we added more contracts, doing so is not necessary to capture the first-order economic forces at play.

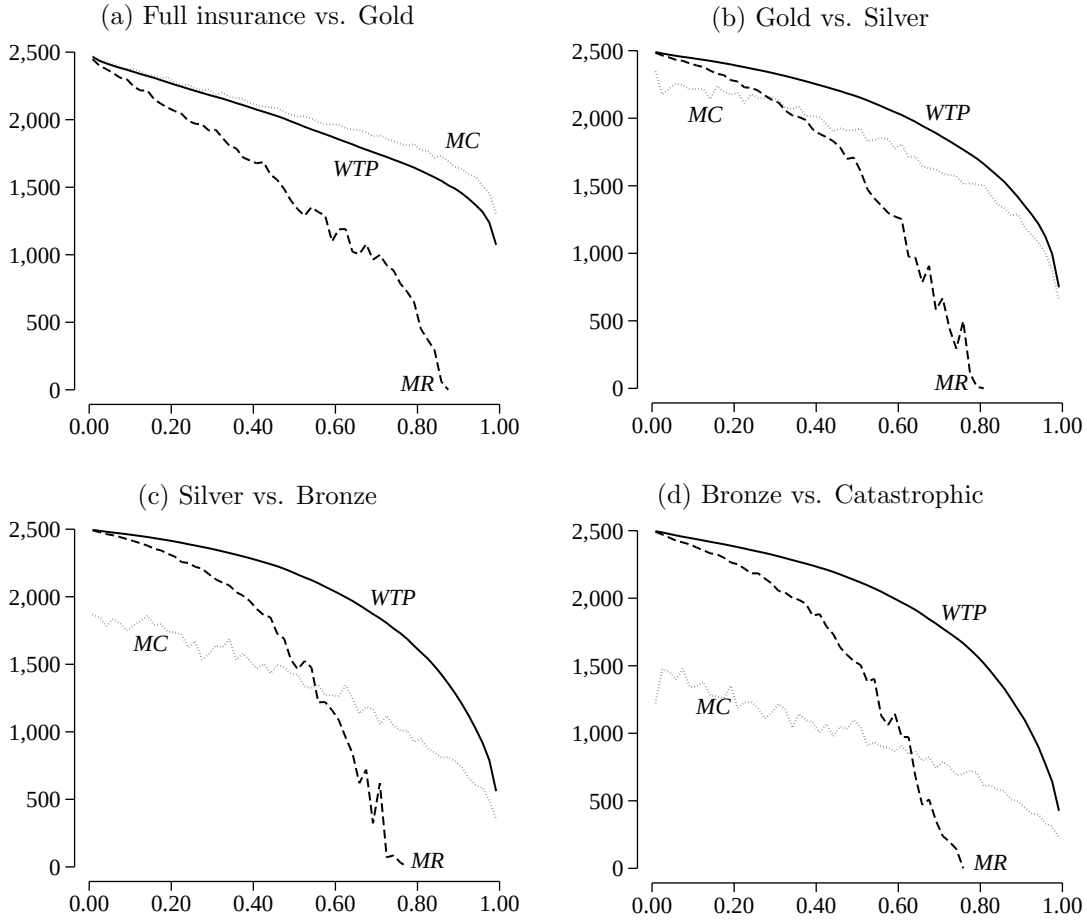
7.3 Graphical Analysis

Because the problem can be solved separately on each increment of coverage, the demand profile approach lends itself directly to graphical analysis. Intuitively, equilibrium outcomes at each increment depend on the demand curves for incremental coverage, the associated marginal revenue curves, and the marginal cost of providing incremental coverage. A monopolist sets marginal cost equal to marginal revenue, while a utilitarian planner sets marginal cost equal to price. A planner with an excess cost of public funds sets marginal cost equal to a weighted average of marginal revenue and price, extracting some surplus from consumers, but not so much as a monopolist.

Figure 4 demonstrates how to carry out the graphical analysis on all increments simultaneously, in order to solve visually for the optimal menu across the full set of potential contracts. Each panel depicts consumers' marginal willingness to pay curve WTP for the given coverage level increment, the associated marginal revenue curve MR, and the marginal cost curve MC associated with providing that coverage level increment.

To solve the insurer's problem, one finds the intersection of marginal benefit and marginal cost in each panel. All types of insurers face the same marginal cost curves, but marginal benefit depends on the objective. For a monopolist, marginal benefit is marginal revenue. The quantities at which MR intersects MC in each panel therefore reveal the fraction of consumers to whom the monopolist wishes to provide that coverage level increment. For example, at the increment between Bronze and Catastrophic coverage, marginal revenue exceeds marginal cost for about the first 63 percent of consumers, consistent with the fact that we see the monopolist optimally allocating 63 percent of consumers to coverage above the Catastrophic contract (cf. Table 1). The associated optimal incremental premium ($1,930) can then be read from the value of the willingness to pay curve at this quantity. At the increment between Gold and Silver coverage, marginal revenue exceeds marginal cost for about the first 30 percent of consumers, consistent with the fact that the monopolist optimally allocates roughly this fraction of consumers to Gold coverage or above. Since the monopolist optimally allocates 63 percent of consumers to coverage above

Figure 4. Illustration of Graphical Analysis for Monopolist and Social Planner



Notes: The figure illustrates the graphical analysis of the reformulated problem. Each panel represents the "market for incremental coverage" between each pair of adjacent contracts. The vertical axes are measured in dollars. The horizontal axes report the percentage of consumers choosing a given incremental level of coverage. Consumers are ordered on the horizontal axes according to their marginal willingness to pay for each coverage level increment. The solid line (WTP) represents consumers' willingness to pay, the dotted line (MC) represents the marginal cost curve, and the dashed line (MR) represents a monopolist's marginal revenue curve. The MC and MR curves are constructed as connected binned scatter plots using 100 points.

Catastrophic, it excludes the remaining 37 percent from the market for incremental coverage. For a social planner with zero cost of funds, the relevant intersection is instead between WTP and MC. Because marginal benefit shifts upward from the marginal revenue curve toward the demand curve as w_C increases, the planner's optimal quantity at each increment lies to the right of the monopolist's, consistent with the comparative statics result in Proposition 5.

Finally, the graphical analysis provides a direct way to verify the second premise of Theorem 2: that the solution to $\tilde{\mathcal{P}}$ exhibits decreasing quantities. In Figure 4, the intersection of marginal

benefit and marginal cost occurs at a progressively higher quantity as one moves from the top coverage increment (Panel (a)) to the bottom (Panel (d)). That is, $\tilde{q}_1 > \tilde{q}_2 > \tilde{q}_3 > \tilde{q}_4$ for each of our focal insurers. Together with the near-*NCD* result documented above, this supports that both premises of Theorem 2 hold in this population, consistent with *DPA* being approximately valid.

7.4 Welfare and Policy Implications

Table 2 reports welfare outcomes, spending outcomes, and allocations for our three focal price-setting insurers as well as under the perfectly competitive equilibrium of Azevedo and Gottlieb (2017). In each case, welfare is measured relative to the allocation of all consumers to the Catastrophic contract.

Table 2. Welfare Outcomes

	Welfare outcomes *$000 per household*			Spending outcomes *$000 per household*			Allocations *Pct. of households*				
Scenario	*SS*†	*CS*†	*PS*	*Gov*	*Prem*	*OOP*	Cstr.	Brnz.	Slvr.	Gold	Full
Panel A. Benchmarks											
* First best	1.93	1.93	–	–	9.05	1.79	<0.01	0.01	0.22	0.57	0.20
Full insurance for all	1.81	1.81	–	11.16	–	–	–	–	–	–	1.00
Minimum coverage for all	–	–	–	5.06	–	5.17	1.00	–	–	–	–
Competitive equilibrium	1.21	1.21	–	5.06	1.55	3.80	0.04	0.65	0.28	0.03	–
Panel B. Optimal menus											
Social planner, $w = (1, 1, 1)$	1.90	1.90	–	8.31	0.83	1.76	<0.01	–	–	1.00	–
Social planner, $w = (0.8, 1, 1)$	1.75	1.75	–	4.91	3.72	2.10	0.14	<0.01	0.13	0.72	–
Monopolist, $w = (0, 1, 0)$	1.17	0.37	0.80	5.06	3.06	3.15	0.37	0.09	0.26	0.29	–

Notes: The table shows welfare outcomes, spending outcomes, and allocations under various scenarios. Social surplus is evaluated using a zero excess cost of public funds. The first set of columns reports social surplus (*SS*), consumer surplus (*CS*), and producer surplus (*PS*) in thousands of dollars per household per year. Note that consumer welfare is normalized to zero at the Catastrophic contract, and accounts for the tax burden associated with government spending. The second set of columns reports expected government spending (*Gov*), premium spending (*Prem*), and expected out-of-pocket spending (*OOP*), again in thousands of dollars per household per year. The final set of columns reports the percentage of households allocated to each contract. †Relative to allocating all consumers to the Catastrophic contract when there is no excess cost of public funds.

Social surplus under a planner facing no excess cost of public funds is $1,897 per household per year. Under a monopolist, social surplus falls to $1,169 and consumer surplus falls to $371. When the insurer is a planner facing an excess cost of public funds, it begins behaving more like the monopolist, placing more weight on profits than consumer surplus. As the excess cost of funds increases from 0 to 0.25, the planner begins to both screen consumers across contracts and to exclude some consumers from the market.

Examining Figure 4, across all four panels simultaneously, provides insight into why the mo-

nopolist destroys so much surplus. While competitive insurers must break even on every contract, the monopolist can internalize pricing externalities that one contract has on the profitability of others. Its markup falls steadily as coverage rises, from 369 percent on the Bronze increment to 186 percent on Silver and 112 percent on Gold (see Online Appendix Table B.2). The lowest increment of coverage is overpriced most aggressively, given that it is relied upon most directly to divert consumers toward higher coverage. With only a single contract, the monopolist has no such tool, as there are no other increments over which to internalize.

It is interesting to note that the monopolist delivers nearly as much social surplus as the competitive equilibrium: \$1,169 versus \$1,211 per household per year. The competitive outcome features substantial unraveling at higher coverage levels—adverse selection drives up the average cost of Gold and Silver coverage, squeezing out lower willingness-to-pay consumers—so even though the monopolist restricts quantity at the bottom, the competitive market restricts it even more at the top.[49] As emphasized by Gottlieb and Moreira (2023), the two market structures cannot be ranked a priori. Each can deliver inefficiently low quantities, but for different reasons. Under competition, adverse selection unravels coverage, preventing trades that would have occurred in a full information environment. Under monopoly, the insurer has the standard rent-seeking incentive to restrict quantity. Which distortion dominates depends on the empirical distribution of types; in our case, the two forces nearly offset one another, with competition delivering only modestly higher social surplus.[50] Of course, the monopolist captures the majority of the surplus it generates, so consumers are considerably better off under competition.

The contrast between the two sources of distortion has direct implications for policy (Mahoney and Weyl, 2017). While linear taxes or subsidies are sufficient for restoring the optimal feasible allocation in a competitive market (Azevedo and Gottlieb, 2017), a price-setting insurer re-optimizes its entire menu in response. An instrument calibrated to a competitive market thus need not have its intended effect under imperfect competition. The graphical analysis in Figure 4, permitted by *DPA*, again makes this point concrete. A linear tax or subsidy shifts the MC curve (or, equivalently, the WTP and MR curves) vertically by a constant in each panel, and the mo-

[49] Note that the competitive equilibrium is not given by the intersection of demand and average cost in the increment panels of Figure 4. Under competition, average costs are endogenous to the allocation, and the equilibrium arises from a fixed-point condition in which each contract is priced at average cost for the consumers who select it. Online Appendix Figure B.4 reproduces the incremental coverage panels with the monopoly and competitive allocations marked, and the resulting deadweight loss shaded. The figure highlights graphically that under monopoly, the welfare loss is concentrated at the bottom increment of coverage, while under competition it appears at the higher increments, where adverse selection unravels coverage.

[50] Indeed, Gottlieb and Moreira (2023) study an analogous decomposition and likewise find that competition delivers higher social surplus in their calibration. Even so, both outcomes are consistent with the theory. That a monopolist can increase total welfare in a selection market is in the spirit of Diamond (1992), who argues that auctioning off the right to serve such a market as a monopolist may dominate free entry, and consistent with Veiga and Weyl (2016) and Ryan (2025), who show that market power can increase efficiency relative to perfect competition.

nopolist's new allocation can be read off where marginal revenue meets the shifted marginal cost. Because marginal revenue is steeper than demand, the monopolist's quantity is less responsive to such a shift than the competitive or efficient quantity. A subsidy is therefore absorbed into premiums and profit more than it is passed through as additional coverage.

Finally, the demand-profile approach also clarifies why the contract space is itself a policy instrument. Because the monopolist prices increment by increment, the set of increments it is permitted to offer determines the pricing externalities it can exploit across coverage levels. Restricting that set, whether by banning contracts or by mandating a higher floor of base coverage, removes screening instruments from the monopolist and weakly lowers its profit. The effect on consumers, however, is ambiguous.[51] A coarser menu extracts less surplus from the consumers the monopolist continues to serve, but it worsens the matching of coverage to type and deepens exclusion at the bottom (cf. Proposition 2), potentially lowering gains from trade overall. The same regulation that softens the distortion at the top can deepen the distortion at the bottom, consistent with Gottlieb and Moreira (2023). Our results highlight the challenges of regulating a non-competitive insurance market, echoing Starc (2014), Jaffe and Shepard (2020), Tebaldi (2024), and Ryan (2025). Strategic insurer responses and endogenous contract characteristics are central to evaluating policy in these markets.

8 Conclusion

We analyze a multidimensional screening model in which a principal offers a menu of vertically differentiated quality-price pairs to agents with multidimensional private information and quasi-linear preferences. Making only a few assumptions on primitives, we derive necessary conditions for optimality that generalize the standard conditions in the one-dimensional case. Despite not being sufficient in general, these conditions have substantive economic content: they establish optimal exclusion, positive trade at all intermediate quality levels, and strict incentives to screen. We apply the framework to health insurance, where the framework is well-suited to studying the problem of insurance contract menu design for a monopolist or a social planner.

We also further develop the demand-profile approach for this class of multidimensional screening problems. We show how this approach can be applied, as well as how its application sheds light on when multidimensional screening problems can be approximately interpreted as a series of one-dimensional problems. Our results provide the conditions for how and when the familiar graphical analysis of health insurance markets introduced by Einav et al. (2010) can be extended to an arbitrary number of vertically-ordered contracts. In the spirit of the original analysis, we view this

[51] Levy and Veiga (2023) show that in competitive insurance markets, by contrast, restricting the maximum allowed coverage is unambiguously welfare-decreasing, even in the presence of moral hazard.

development as a useful tool with which one can understand the basic theory of multidimensional screening in selection markets with endogenous product quality and the associated implications for welfare and public policy.

Finally, we quantify the magnitudes of theoretically identified effects and use a numerical model to evaluate and illustrate a number of our key results. We find that only a small number of contracts are needed to capture the key economic features of the problem, but that over-restricting the contract space (for example to only two contracts) does have material implications. We also find that in our setting, the demand profile approach provides an excellent approximation of the true problem, and therefore represents a powerful tool for understanding its solution. We view further exploration of the validity of the demand-profile approach in other empirical settings to be of central interest. Finally, our analysis assumes a single principal; extending these tools to oligopolistic settings is a natural direction for future work.

References

Araújo, A. P. and V. Perez (2025, February). Single crossing in multidimensional screening: Implementability and optimality. SSRN Working Paper No. 5146614.

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica 64*(1), 51–75.

Azevedo, E. M. and D. Gottlieb (2017). Perfect competition in markets with adverse selection. *Econometrica 85*(1), 67–105.

Barelli, P., S. Basov, M. Bugarin, and I. King (2014). On the optimality of exclusion in multi-dimensional screening. *Journal of Mathematical Economics 54*, 74–83.

Billingsley, P. (1995). *Probability and Measure*. New York: Wiley.

Bulow, J. and J. Roberts (1989). The simple economics of optimal auctions. *Journal of Political Economy 97*(5), 1060–1090.

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

Deneckere, R. and S. Severinov (2017). A solution to a class of multi-dimensional screening problems: Isoquants and clustering. University of British Columbia Working Paper.

Diamond, P. (1992). Organizing the health insurance market. *Econometrica: Journal of the Econometric Society*, 1233–1254.

Einav, L. and A. Finkelstein (2011). Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic Perspectives 25*(1), 115–38.

Einav, L., A. Finkelstein, and M. R. Cullen (2010). Estimating welfare in insurance markets using variation in prices. *The quarterly journal of economics 125*(3), 877–921.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on moral hazard in health insurance. *American Economic Review 103*(1), 178–219.

Farinha Luz, V., P. Gottardi, and H. Moreira (2023). Risk classification in insurance markets with risk and preference heterogeneity. *Review of Economic Studies 90*(6), 3022–3082.

Figalli, A., Y.-H. Kim, and R. McCann (2011). When is multidimensional screening a convex program? *Journal of Economic Theory 146*(2), 454–478.

Gaynor, M., N. Mehta, and S. Richards-Shubik (2023). Optimal contracting with altruistic agents: Medicare payments for dialysis drugs. *American Economic Review 113*(6), 1530–1571.

Geruso, M., T. J. Layton, G. McCormack, and M. Shepard (2023). The two-margin problem in insurance markets. *The Review of Economics and Statistics 105*(2), 237–257.

Gottlieb, D. and H. Moreira (2023, November). Market power and insurance coverage. Discussion Paper 890, Financial Markets Group, London School of Economics.

Ho, K. and R. S. Lee (2023). Health insurance menu design for large employers. *The RAND Journal of Economics 54*(4), 598–637.

Jaffe, S. and M. Shepard (2020). Price-linked subsidies and imperfect competition in health insurance. *American Economic Journal: Economic Policy 12*(3), 279–311.

Kadan, O., P. J. Reny, and J. M. Swinkels (2017). Existence of optimal mechanisms in principal-agent problems. *Econometrica 85*(3), 769–823.

Karlin, S. and Y. Rinott (1980). Classes of orderings of measures and related correlation inequalities. I. multivariate totally positive distributions. *Journal of Multivariate Analysis 10*(4), 467–498.

Laffont, J.-J., E. Maskin, and J.-C. Rochet (1987). Optimal nonlinear pricing with two-dimensional characteristics. In T. Groves, R. Radner, and S. Reiter (Eds.), *Information, Incentives, and Economic Mechanisms: Essays in Honor of Leonid Hurwicz*, pp. 256–266. Minneapolis: University of Minnesota Press.

Levy, Y. J. and A. Veiga (2023). Optimal contract regulation in selection markets. SSRN Working Paper No. 4029945.

Mahoney, N. and E. G. Weyl (2017). Imperfect competition in selection markets. *Review of Economics and Statistics 99*(4), 637–651.

Manelli, A. M. and D. R. Vincent (2006). Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory 127*(1), 1–35.

Marone, V. R. and A. Sabety (2022, January). When should there be vertical choice in health insurance markets? *American Economic Review 112*(1), 304–42.

Maskin, E. and J. Riley (1984). Monopoly with incomplete information. *The RAND Journal of Economics 15*(2), 171–196.

McCann, R. and K. Zhang (2019). On concavity of the monopolist's problem facing consumers with nonlinear price preferences. *Communications on Pure and Applied Mathematics 72*(7), 1386–1423.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*(2), 301–317.

Pass, B. and O. Halim (2025). Multi-to-one dimensional and semi-discrete screening. University of Alberta Working Paper.

Rochet, J.-C. and P. Choné (1998). Ironing, sweeping, and multidimensional screening. *Econometrica 66*(4), 783–826.

Rochet, J.-C. and L. A. Stole (2003). The economics of multidimensional screening. In M. Dewatripont, L. P. Hansen, and S. J. Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume I*, Econometric Society Monographs, pp. 150–197. Cambridge: Cambridge University Press.

Rothschild, M. and J. Stiglitz (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics 90*(4), 629–649.

Rudin, W. (1976). *Principles of Mathematical Analysis* (3rd ed.). McGraw-Hill.

Ryan, C. (2025). Mergers in the Presence of Adverse Selection. *RAND Journal of Economics*. Forthcoming.

Shannon, C. (1995). Weak and strong monotone comparative statics. *Economic Theory 5*, 209–227.

Starc, A. (2014). Insurer pricing and consumer welfare: Evidence from Medigap. *RAND Journal of Economics 45*(1), 198–220.

Stiglitz, J. (1977). Monopoly, non-linear pricing and imperfect information: The insurance market. *Review of Economic Studies 44*(3), 407–430.

Tebaldi, P. (2024). Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the ACA. *The Review of Economic Studies 92*(1), 586–620.

Veiga, A. and E. G. Weyl (2016). Product design in selection markets. *The Quarterly Journal of Economics 131*(2), 1007–1056.

Wilson, R. B. (1993). *Nonlinear Pricing*. Oxford University Press on Demand.

Yang, F. (2025). Costly multidimensional screening. *Review of Economic Studies*. Forthcoming.

Appendices

Appendix A Proofs of Main Results

A.1 Proofs for Section 4

Proof of Proposition 2 Assume $w_C < 1$ and consider any price vector ρ that does not exclude. Thus, for a measure one of θ, $\psi_1(\rho, \theta) = \underline{\psi}$. Now, consider any x_j, and let Θ_j be the subset of Θ where $(\underline{\psi}, \theta)$ chooses x_j facing ρ, noting that since there is no exclusion, $\sum_{j=1}^{J} H(\Theta_j) = 1$. By Assumption 3,

$$H(\{\theta | v(x_j, \underline{\psi}, \theta) - v(x_0, \underline{\psi}, \theta) = \rho(x_j)\}) = 0$$

and by *IC*,

$$H(\{\theta \in \Theta_j | v(x_j, \underline{\psi}, \theta) - v(x_0, \underline{\psi}, \theta) < \rho(x_j)\}) = 0.$$

Thus,

$$H(\{\theta \in \Theta_j | v(x_j, \underline{\psi}, \theta) - v(x_0, \underline{\psi}, \theta) > \rho(x_j)\}) = H(\Theta_j).$$

But then, a measure one subset of Θ will be unresponsive to a small change in the price of all qualities x_1 and above, and so it was appropriate to set $r_1 = 0$ in this case, and we have

$$\int \left((1 - w_C)\left(1 - H(\psi_1(\rho, \theta)|\theta)\right) - r_1(\rho, \theta) h(\psi_1(\rho, \theta)|\theta)\right) h(\theta) d\theta = 1 - w_C > 0,$$

and thus the perturbation in which the price of all qualities x_1 and above is raised by a little is strictly profitable, a contradiction. □

Proof of Proposition 4 Let x_{j^ℓ} be the lowest intermediate quality and x_{j^h} the highest. Let $x_{j^*} \in \{x_{j^\ell}, \ldots, x_{j^h}\}$ be assigned with zero probability. Lower the price of x_{j^*} until it just begins to attract switchers, and then lower it $\varepsilon > 0$ further so that for a positive mass of types θ some values of ψ switch. The relevant θ are of three possible types, depending on whether $\psi_{j^*}(\rho, \theta)$ is interior, at the lower boundary, or at the upper boundary. Assume first that a positive mass of the switching θ have $\psi_{j^*}(\rho, \theta) = \underline{\psi}$ so that for θ in this set, all ψ choose strictly above x_{j^*}. Then, since $\mathcal{V}(\cdot, \underline{\psi}, \theta)$ has an optimum at or below x_{j^ℓ} it follows from the strict concavity of $\mathcal{V}$ in x and Assumption 3 that on a full-measure subset of these types, $\mathcal{V}(\chi(\psi, \theta), \underline{\psi}, \theta) - \mathcal{V}(x_{j^*}, \underline{\psi}, \theta) < 0$. When ε is small, if (ψ, θ) is the type switching, then ψ is close to $\underline{\psi}$ and so the principal strictly benefits from the switch down to x_{j^*}. Similarly, the principal strictly benefits if a positive mass of types with $\psi_{j^*}(\rho, \theta) = \bar{\psi}$ switch their choice up to x_{j^*}.

Finally, for θ where $\psi_j(\rho, \theta)$ is interior and where switching to x_{j^*} is occuring, note that for

ε positive but small, some types with ψ just above $\psi_j(\rho,\theta)$ switch down from $\bar{x}_{j^*}$ to x_{j^*} and some types with ψ just below $\psi_j(\rho,\theta)$ switch from $\underline{x}_{j^*}$ up to x_{j^*}. As in the proof of Theorem 1, the rate of change of the principal's payoffs from those switching down is arbitrarily close to $h(\psi_{j^*}(\rho,\theta)|\theta)$ times

$$-\frac{\mathcal{V}(\bar{x}_{j^*}(\rho,\theta),\psi_{j^*}(\rho,\theta),\theta)-\mathcal{V}(x_{j^*},\psi_{j^*}(\rho,\theta),\theta)}{v_\psi(\bar{x}_{j^*}(\rho,\theta),\psi_{j^*}(\rho,\theta),\theta)-v_\psi(x_{j^*},\psi_{j^*}(\rho,\theta),\theta)},$$

where as usual, the numerator measures the impact of the change in allocation, and one over the denominator is the rate at which the boundary ψ is moving. The approximation is because for $\varepsilon>0$ the boundary is slightly above $\psi_{j^*}(\rho,\theta)$, and the minus sign reflect that the movement is downward. By Cauchy's mean value theorem this fraction equals $-\mathcal{V}_x/v_{\psi x}$ for some $x''\in(x_{j^*},\bar{x}_{j^*})$. Similarly, the impact on the principal from types moving upward is arbitrarily close to $h(\psi_{j^*}(\rho,\theta)|\theta)$ times $\mathcal{V}_x/v_{\psi x}$ evaluated at some $x'\in(\underline{x}_{j^*},x_{j^*})$. By Assumption 5, the difference between the two expressions is strictly positive and so the principal strictly benefits.[52] Since the principal strictly benefits in any of the three cases, we have arrived at a contradiction. □

A.2 Details for Section 5

Proof of Proposition 6 Assume $\tilde{q}_j=1$, so that $\tilde{p}_j=\underline{p}_j\equiv P^j(1)$. We will show that for small ε, raising p_j to $\underline{p}_j+\varepsilon$ strictly profits the principal. To do so, note that $\underline{p}_j=\min_\theta v_j(\underline{\psi},\theta)$, and let $\Theta_j(\varepsilon)=\{\theta|v_j(\underline{\psi},\theta)\in[\underline{p}_j,\underline{p}_j+\varepsilon]\}$. By Assumption 3, $H(\Theta_j(0))=0$, and thus $H\left(\Theta_j(\varepsilon)\right)\to 0$.[53] Let $\underline{v}_{j\psi}=\min_{(\psi,\theta)}(v_j)_\psi>0$. Then, $\{(\psi,\theta)|v_j(\psi,\theta)\in[\underline{p}_j,\underline{p}_j+\varepsilon]\}\subseteq[\underline{\psi},\underline{\psi}+\varepsilon/\underline{v}_{j\psi}]\times\Theta_j(\varepsilon)$. Thus, letting $\bar{h}=\max_{(\psi,\theta)}h(\psi|\theta)$, raising the price to $\underline{p}_j+\varepsilon$ loses at most $(\bar{h}\varepsilon/\underline{v}_{j\psi})H\left(\Theta_j(\varepsilon)\right)$ consumers but raises ε in revenue on $Q^j(\underline{p}_j+\varepsilon)$ consumers. Thus, (dividing by ε) the profit of raising the price to $\underline{p}_j+\varepsilon$ has the same sign as

$$(1-w_C)Q^j(\underline{p}_j+\varepsilon)-H\left(\Theta_j(\varepsilon)\right)\frac{\bar{h}}{\underline{v}_{j\psi}}\max_{(\psi,\theta)}(|\mathcal{V}_j(\psi,\theta)|+(1-w_C)\varepsilon)$$

where $\mathcal{V}_j(\psi,\theta)\equiv\mathcal{V}(x_j,\psi,\theta)-\mathcal{V}(x_{j-1},\psi,\theta)$ and where $\max_{(\psi,\theta)}(|\mathcal{V}_j(\psi,\theta)|+(1-w_C)\varepsilon)$ is the most that can be lost by the principal on a consumer who leaves, since the consumer is within ε of being indifferent. Since $H\left(\Theta_j(\varepsilon)\right)\to 0$ and $Q^j(\underline{p}_j+\varepsilon)\to 1$, this is positive for $\varepsilon>0$ small enough. □

[52] As before, we could have used that the perturbation can be thought of as the sum of two perturbations of the form underlying (*NC*), simply noting that the terms for those who are switching past x_{j^*} cancel, as do the terms $1-H$ for those who switch to x_j^*.

[53] This follows from Theorem 2.1 in Billingsley (1995) since $\Theta_j(0)=\cap_k\Theta_j(1/2^k)$.

A.3 Details for Section 5.3

Let us first show that $\mathcal{H} \subseteq \mathcal{H}_A$. To see this, fix x' and x'', and define $\grave{\psi}(\tau,\theta)$ implicitly by $v(x'',\grave{\psi}(\tau,\theta),\theta) - v(x',\grave{\psi}(\tau,\theta),\theta) = \tau$ when such a solution exists, and by $\underline{\psi}$ and $\bar{\psi}$ otherwise as appropriate. Then using Assumption 3, for each τ, for almost all θ, $\grave{\psi}(\cdot,\theta)$ is continuously differentiable with measurable and bounded derivative (given by the implicit function theorem where $\grave{\psi}$ is interior and given by zero for H-almost all θ where $\grave{\psi}$ is a corner solution). Thus, if we define $Y(\tau) = H(\{v(x'',\psi,\theta) - v(x',\psi,\theta) \leq \tau\})$, then $Y(\tau) = \int_\Theta H(\grave{\psi}(\tau,\theta)|\theta)h(\theta)d\theta$ and so by Lebesgue's dominated convergence theorem, Y has continuous density given by $y(\tau) = \int_\Theta \grave{\psi}_\tau(\tau,\theta)h(\grave{\psi}(\tau,\theta)|\theta)dH(\theta)$. □

Proof of Proposition 7 For each (ψ,θ), let $\iota(\psi,\theta)$ be the fraction of types who have higher incremental values than (ψ,θ). Let (p,χ) be *IC*. If $\chi(\psi',\theta') \geq x_j$ then the same is true of any consumer with $\iota(\psi'',\theta'') < \iota(\psi',\theta')$, since (ψ'',θ'') values all increments by strictly more than (ψ',θ'). Hence, if we let q_j be the fraction of types allocated x_j or above, then

$$\{(\psi,\theta)|x_j \leq \chi(\psi,\theta)\} = \{(\psi,\theta)|\iota(\psi,\theta) \leq q_j\}. \tag{13}$$

Let $q_{j'+1} < q_{j'} = q_{j''} < q_{j''-1}$ for some $j' \geq j''$, so that consumers with $\iota \in (q_{j'+1}, q_{j'})$ are allocated $x_{j'}$ and consumers with $\iota \in (q_{j''}, q_{j''-1})$ are allocated $x_{j''-1}$ (and no types are allocated qualities from $x_{j''}$ to $x_{j'-1}$). Note that for any (ψ,θ) with $\iota(\psi,\theta) = q_j$, we have that $v_j(\psi,\theta) = P^j(q_j)$, and so this type's willingness to pay to move from $x_{j''-1}$ to $x_{j'}$ is $\sum_{j''}^{j'} P^j(q_j)$. The role of *NCD* here is that $\sum_{j''}^{j'} P^j(q_j) = v(x_{j''},\cdot) - v(x_{j'},\cdot)$ for all types on the margin of switching. This would not be true without *NCD*, which opens the door to the optimality of bundling. Thus, if $\sum_{j''}^{j'} p_j > \sum_{j''}^{j'} P^j(q_j)$ then types with ι just below $q_{j'}$ will deviate from $x_{j'}$ down to $x_{j''-1}$.[54] Similarly, it cannot be that $\sum_{j''}^{j'} p_j < \sum_{j''}^{j'} P^j(q_j)$. Thus, we can take $p_j = P^j(q_j)$ for all j where q_j is interior. It is also without loss to take $p_j = P^j(0)$ when $q_j = 0$. Finally, let $\tilde{j}$ be the highest index for which $q_j = 1$, and assume $\tilde{j} \geq 1$. Then, by *IC*, $\sum_{j=1}^{\tilde{j}} p_j \leq \sum_{j=1}^{\tilde{j}} P^j(1)$, otherwise types with ι just below 1 will deviate from $x_{\tilde{j}}$ down to x_0. If the inequality is strict then one can raise $\sum_{j=1}^{\tilde{j}} p_j$ without changing the allocation and hence strictly increase the principal's payoff if $w_C < 1$ (it does not change the principal's payoff if $w_C = 1$). Hence, it is without loss that $p_j = P^j(q_j)$ for all j. Thus, for each j,

$$\int_{\{(\psi,\theta)|x_j \leq \chi(\psi,\theta)\}} \left(\mathcal{V}^j(\psi,\theta) - (1-w_C)(v_j(\psi,\theta) - p_j)\right) dH(\psi,\theta) = \tilde{\Pi}^j(q_j)$$

[54] The assumption that the support of v_j is an interval ensures the existence of a positive measure of such types.

and so, $\mathcal{P}$ can written as $\max_{\{q|q_j \geq q_{j+1}\}} \sum_{j=1}^{J} \tilde{\Pi}^j(q_j)$. But then, $\tilde{\mathcal{P}}$ is indeed a relaxation of $\mathcal{P}$ and so if the solution to $\tilde{\mathcal{P}}$ satisfies that q_j is decreasing, then it solves $\mathcal{P}$. □

We next present a lemma showing that the payoff to the principal in both $\mathcal{P}$ and $\tilde{\mathcal{P}}$ are continuous in prices and distributions when restricted to $\mathcal{H}_A$. Formally, fix $\{x_j\}_{j=0}^{J}$, and let $\Pi(p, H)$ be the payoff to the principal when they set prices p facing distribution H and then allocate the consumers subject to incentive compatibility given p. Let $\tilde{\Pi}(p, H)$ be the principal's payoff in the *DPA* problem. Let $\bar{p} = 1 + \max_{x,\psi,\theta} (v(x, \psi, \theta) - v(x_0, \psi, \theta))$ which is finite since x and θ come from compact spaces and v is continuous, and note that any price above $\bar{p}$ on any increment of demand is equivalent to $\bar{p}$ (it has the effect of x_j being refused by all participants in *DPA*, and of all quality levels at or above x_j being refused in the original problem).

Lemma 1 (Continuity) *The functions* Π *and* $\tilde{\Pi}$ *are continuous on* $[0, \bar{p}]^J \times \mathcal{H}_A$. *The set of optimal price vectors in* $\mathcal{P}$ *(respectively* $\tilde{\mathcal{P}}$*) is upper-hemicontinuous on* $\mathcal{H}_A$.

See Online Appendix B.5 for the details. The hard part is continuity. Upper-hemicontinuity then follows immediately from the theorem of the maximum.

Proof of Theorem 2 Let $H_k \to H_0$ be a sequence in $\mathcal{H}_A$, where S_k is the support of H_k and S_0 is the support of H_0. Write $\mathcal{P}_k$ and $\tilde{\mathcal{P}}_k$ for the problems $\mathcal{P}$ and $\tilde{\mathcal{P}}$ facing H_k and $\mathcal{P}_0$ and $\tilde{\mathcal{P}}_0$ for the associated problems facing H_0.

Step 1: Let $\mathcal{Y}(\psi, \theta, p_j) \equiv \mathcal{V}^j(\psi, \theta) - (1 - w_C)(v_j(\psi, \theta) - p_j))$. Then using (5) and (6) for any given $H \in \mathcal{H}_A$, $\tilde{\mathcal{P}}$ can be written as

$$\max_{p} \sum_{j=1}^{J} \int_{\{(\psi,\theta)|v_j(\psi,\theta) \geq p_j\}} \mathcal{Y}(\psi, \theta, p_j) dH(\psi, \theta) \tag{14}$$

while $\mathcal{P}$ can be written as

$$\max_{p} \sum_{j=1}^{J} \int_{\{(\psi,\theta)|\chi(\psi,\theta) \geq x_j\}} \mathcal{Y}(\psi, \theta, p_j) dH(\psi, \theta) \text{ s.t. } IC, \tag{15}$$

where in each case, we use that since $H \in \mathcal{H}_A$, the measure of types where the agent is indifferent between two qualities for any given p is zero.

Step 2: The solution to $\tilde{\mathcal{P}}_0$ is unique. To see this, note that each summand in (14) depends only on p_j, and is strictly quasiconcave in p_j by regularity. Let the solution in $\tilde{\mathcal{P}}_0$ be p^0 and for each j, let $q_j^0 = H_0(\{(\psi, \theta)|v_j(\psi, \theta) \geq p_j^0\})$. By assumption q_j^0 is strictly decreasing in j. But then, by Proposition 7, p^0 is also the unique solution to $\mathcal{P}_0$.

Step 3: For any given p, define

$$F(p) = \{(\psi,\theta)| \text{ for some } j,\, v_j(\psi,\theta) \leq p_j \text{ and } v_{j+1}(\psi,\theta) \geq p_{j+1}\},$$

where F is mnemonic for "failing", since on the interior of F, QC is violated. On the complement of $F(p)$, QC holds strictly in that if $v_j(\psi,\theta) \geq p_j$ then $v_{j'}(\psi,\theta) > p_{j'}$ for all $j' < j$. Because it is defined in terms of weak inequalities and continuous functions, F is an upper-hemicontinuous correspondence. Note also that on S_0,

$$v_j(\psi,\theta) \leq p_j^0 \Rightarrow \iota(\psi,\theta) \geq q_j^0 \Rightarrow \iota(\psi,\theta) > q_{j+1}^0 \Rightarrow v_{j+1}(\psi,\theta) < p_{j+1}^0.$$

Thus, the closed set $F(p^0)$ is disjoint from the closed set S_0.

Step 4: Choose a closed neighborhood $\hat{S}$ of S_0 where $\hat{S}$ and $F(p^0)$ are disjoint. This can be done since S_0 and $F(p^0)$ are closed and disjoint subsets of a compact metric space. Choose a closed neighborhood P^0 of p^0 such that for $p \in P^0$, $F(p)$ and $\hat{S}$ are disjoint. This can be done since F is upper-hemicontinuous and $\hat{S}$ and $F(p^0)$ are disjoint.

Step 5: Let k be large enough that any optimal price vector in either $\mathcal{P}_k$ or $\tilde{\mathcal{P}}_k$ is contained in P^0. This can be done using the upper-hemicontinuity part of Lemma 1, since $\mathcal{P}_0$ and $\tilde{\mathcal{P}}_0$ both have unique solution p^0. Thus, for any such k, it is without loss of generality to add the constraint that $p^0 \in P^0$ to either (14) or (15). But then, since for $p \in P^0$, $F(p)$ and $\hat{S}$ are disjoint, it follows that for $(\psi,\theta) \in \hat{S}$, and for the associated IC allocation function χ given p (and continuing to ignore the zero probability set of (ψ,θ) where there are indifferences),

$$v_j(\psi,\theta) \geq p_j \Rightarrow v_{j'}(\psi,\theta) > p_{j'} \text{ for all } j' < j \Rightarrow x_j \leq \chi(\psi,\theta) \Rightarrow v_j(\psi,\theta) \geq p_j,$$

where the last implication follows since if for given (ψ,θ), $\tilde{j} \geq j$ is the index associated with $\chi(\psi,\theta)$ then by IC, $v_{\tilde{j}}(\psi,\theta) \geq p_{\tilde{j}}$ and so, since we are not in $F(p)$, $v_j(\psi,\theta) \geq p_j$. Thus, $\mathcal{P}_k$ can be written as

$$\max_{p \in P^0} \left(\sum_{j=1}^{J} \int_{T_1} \mathcal{Y}(\psi,\theta,p_j) dH_k(\psi,\theta) + \sum_{j=1}^{J} \int_{T_2} \mathcal{Y}(\psi,\theta,p_j) dH_k(\psi,\theta) \right) \tag{16}$$
$$\text{s.t. } IC \text{ on } \{(\psi,\theta) \notin \hat{S}\},$$

where $T_1 = \{(\psi,\theta) \in \hat{S}\} \cap \{(\psi,\theta)|v_j(\psi,\theta) \geq p_j\}$ and $T_2 = \{(\psi,\theta) \notin \hat{S}\} \cap \{(\psi,\theta)|x_j \leq \chi(\psi,\theta)\}$.

Step 6: Let k be large enough that any optimal solution to either $\mathcal{P}$ or $\tilde{\mathcal{P}}$ is in P^0 and such that $S_k \in \hat{S}$. Then, since H_k puts no weight outside of $\hat{S}$, the second term in (16) disappears as does

IC, and we can broaden the domain of integration in the first term to simply $\{(\psi,\theta)|v_j(\psi,\theta)\geq p_j\}$. Thus $\mathcal{P}$ and $\tilde{\mathcal{P}}$ coincide and hence *DPA* is valid.

Step 7: As k diverges, $H_k(\hat{S})\to 1$ by the definition of weak convergence. Thus, since $\mathcal{Y}$ is bounded, the second term vanishes asymptotically, and so the payoff to any given $p\in P^0$ in $\mathcal{P}$ converges to that in (14). And, since $H_k(\hat{S})\to 1$, and since QC is automatic on $\hat{S}$ for $p\in P^0$, for a set of types with probability going to one, the allocation in $\mathcal{P}$ is the same as the one in $\tilde{\mathcal{P}}$. Hence, *DPA* is ε-valid, completing the proof of the theorem. □

How Close is Close Enough? Theorem 2 says that on a *DPA* is valid on a neighborhood of H_0. But, it does not say exactly how big that neighborhood is. Here, we address this issue. Let H_0 satisfy the conditions of Theorem 2 and let p^0 be the optimal price schedule. Then (see Step 3 of the proof), there is $\delta>0$ such that $F(p)\cap B_\delta(S_0)$ is empty for all $p\in B_\delta(p^0)$ (we continue to work with $\|p-\hat{p}\|=\max_j|p^j-\hat{p}^j|$). That is, there is a neighborhood of the support of H_0 (the green region in the figure) and a ball in price space such that for p in the ball, the green region and pink regions remain separated. Finding a viable δ is a numerically trivial task, as one simply checks, for each j and $j+1$, how far one can increase p^j and decrease p^{j+1} before a conflict arises, takes half this distance, and then observes that the blue lines can only converge so fast. In numerical examples, δ is often large.

Tracing the proof of Theorem 2, if one knows that for given H with support in $B_\delta(S_0)$ the solutions to both $\mathcal{P}$ and $\tilde{\mathcal{P}}$ are in $B_\delta(p^0)$ then *DPA* is valid. The problem is that the most direct way of checking this is to solve $\mathcal{P}$, and avoiding doing this is part of the point of *DPA*. To get around this, let τ_0 be the minimum amount by which the principal is harmed in $\tilde{\mathcal{P}}$ facing H_0 by being forced to change one of the prices by δ away from its optimum.[55] Since $\tilde{\mathcal{P}}$ has a unique optimum, τ_0 is strictly positive. It is also trivial to calculate in examples.[56] We then have the following result, which provides a bound for H being close enough to H_0 that does not involve solving $\mathcal{P}$. Let λ be a Lipschitz constant for γ^j for all j.

Proposition 8 *DPA is valid for all H with support in $B_\delta(suppH_0)$ and within Wasserstein-1 distance $\delta^2/3\lambda$ from H_0.*[57]

Proof By Kantorovich-Rubinstein duality, when H is within $\frac{\delta}{3\lambda}$ of H_0, $\Pi(p,H)$ is within $\frac{\delta}{3}$ of

[55] That is, $\tau_0=\min_j\min\left\{\tilde{\Pi}^j(p_0^j,H_0)-\tilde{\Pi}^j(p_0^j-\delta,H_0),\tilde{\Pi}^j(p_0^j,H_0)-\tilde{\Pi}^j(p_0^j+\delta,H_0)\right\}>0$,

[56] If one wishes, one can impose conditions on the primitives such that profits are strictly concave in prices, so that τ_0 is proportional to δ^2.

[57] The Wasserstein 1-distance is one way of metrizing the weak topology. It is often referred to as the "earthmover metric" in that if one takes two densities, then the Wasserstein 1-distance is the cost of the most efficient way to take mass from one density and move it so at to achieve the other, where transport costs are linear in weight times distance.

$\Pi(p, H_0)$. Thus, for any $p \notin B_\delta(p^0)$,

$$\Pi(p,H) \leq \Pi(p,H_0) + \frac{\delta}{3} \leq \max_{p \notin B_\delta(p^0)} \Pi(p,H_0) + \frac{\delta}{3} = \Pi(p^0,H_0) - \frac{2\delta}{3} \leq \Pi(p^0,H) - \frac{\delta}{3},$$

and so p is not optimal in $\mathcal{P}$ facing H, since p^0 performs strictly better. Arguing similarly for $\tilde{\mathcal{P}}$, it follows that the solutions to both P and $\tilde{P}$ facing H are within $B_\delta(p^0)$, and hence they coincide and *DPA* is valid. □

Primitives for *NCD* and a Strictly Decreasing Solution A *path* is a twice-continuously differentiable function $\kappa : [0,1] \to \Psi \times \Theta$, where we write $\kappa = (\kappa_\psi, \kappa_1, \ldots, \kappa_N)$. That is, κ traces out a one-dimensional subset of $\Psi \times \Theta$, so that types are in a sense perfectly correlated along the image of κ. Let G be a distribution on $[0,1]$ with continuous density g, and say that H is *generated* by G and κ if H results by drawing τ from $[0,1]$ according to G, and then associating the point $\kappa(\tau) \in \Psi \times \Theta$. That is, for any given measurable set $S \subseteq \Psi \times \Theta$, $H(S) = G(\{\tau | \kappa(\tau) \in S\})$. Let $\hat{\mathcal{V}}(x,\tau) = \mathcal{V}(x,\kappa(\tau))$ and similarly for $\hat{v}$, $\hat{\gamma}$, and $\hat{\eta}$. For given $\{x_i\}_{i=0}^N$, let $\hat{\mathcal{V}}^j(\tau) = \hat{\mathcal{V}}(x_j,\tau) - \hat{\mathcal{V}}(x_{j-1},\tau)$ and similarly for $\hat{v}^j$, $\hat{\gamma}^j$, and $\hat{\eta}^j$.

Definition 1 (Alignment) *Distribution $H \in \mathcal{H}_A$ is aligned given $\{x_j\}_{j=0}^N$ if it is generated by a path κ and distribution G on $[0,1]$ satisfying*

1. $\hat{v}_{x\tau} > 0$,
2. $\frac{\hat{\mathcal{V}}^j}{\hat{v}^j_\tau} \frac{g}{1-G}$ is strictly increasing in τ, and
3. $\frac{\hat{\mathcal{V}}_x(\cdot,\tau)}{\hat{v}_{x\tau}(\cdot,\tau)}$ is strictly decreasing in x.

Condition 1 says that the demand for quality rises along the path. Condition 2 will be seen to be equivalent to the marginal benefit side of (7) crossing the marginal cost side at most once and from above. In Online Appendix B.5 we provide conditions under which Condition 2 holds regardless of the choice of $\{x_j\}$. Condition 3 will imply that q is strictly decreasing in j. It should be compared to Assumption 5, with τ playing the role of ψ. Note that when H is aligned, then it fails, for example, Assumption 1 and so is not an element of $\mathcal{H}$. But, $H \in \mathcal{H}_A$, and this suffices.

Theorem 6 (Primitives) *Let $w_C < 1$ and let $H \in \mathcal{H}_A$ be aligned given $\{x_i\}_{i=0}^N$. Then H has NCD, is regular, and yields a strictly decreasing solution.*

Proof That H has *NCD* is immediate from Condition 1 of Definition 1, since types in the support of H are ordered by τ. Write $\hat{\Pi}^j(\tau)$ for the principal's payoff on increment j when types with

index above τ are served. Then, since the associated price will be $\hat{v}(x_j,\tau)-\hat{v}(x_{j-1},\tau)$, we have

$$\hat{\Pi}^j(\tau)=\int_\tau^1\left(\hat{\mathcal{V}}^j(\tau')-(1-w_C)(\hat{v}^j(\tau')-\hat{v}^j(\tau))\right)g(\tau')d\tau'$$

and hence,

$$\hat{\Pi}^j_\tau(\tau)=-g(\tau)\hat{\mathcal{V}}^j(\tau)+(1-w_C)\hat{v}^j_\tau(\tau)(1-G(\tau)).$$

Note that $\tilde{\Pi}^j$ is strictly quasiconcave in q (i.e., is regular) if and only if $\hat{\Pi}^j$ is strictly quasiconcave in τ. But,

$$\hat{\Pi}^j_{\tau\tau}(\tau)=-g(\tau)\hat{\mathcal{V}}^j_\tau(\tau)-g_\tau(\tau)\hat{\mathcal{V}}^j(\tau)+(1-w_C)\left(\hat{v}^j_{\tau\tau}(\tau)(1-G(\tau))-\hat{v}^j_\tau(\tau)g(\tau)\right),$$

and so, where $\hat{\Pi}^j_\tau=0$, we have that $g(\tau)\hat{\mathcal{V}}^j(\tau)=(1-w_C)\hat{v}^j_\tau(\tau)(1-G(\tau))>0$ and so,

$$\hat{\Pi}^j_{\tau\tau}(\tau)=_s-\frac{\hat{\mathcal{V}}^j_\tau(\tau)}{\hat{\mathcal{V}}^j(\tau)}-\frac{g_\tau(\tau)}{g(\tau)}+\frac{\hat{v}^j_{\tau\tau}(\tau)}{\hat{v}^j_\tau(\tau)}-\frac{g(\tau)}{1-G(\tau)}=_s-\left(\frac{\hat{\mathcal{V}}^j}{\hat{v}^j_\tau}\frac{g}{1-G}\right)_\tau,$$

and so $\hat{\Pi}^j_{\tau\tau}(\tau)<0$ by Condition 2 of Definition 1.

Finally, if we let τ_j solve $\hat{\Pi}^j_\tau(\tau_j)=0$, so that q_j is interior, then $\hat{\Pi}^{j+1}_\tau(\tau_j)>0$, so that $\tau_{j+1}>\tau_j$, and hence $q_{j+1}<q_j$. To see this, note that

$$\hat{\Pi}^j_\tau(\tau)=_s-\frac{\hat{\mathcal{V}}^j(\tau_j)}{\hat{v}^j_\tau(\tau_j)}+(1-w_C)\frac{1-G(\tau_j)}{g(\tau_j)},$$

where by Cauchy's mean-value theorem,

$$\frac{\hat{\mathcal{V}}^j(\tau_j)}{\hat{v}^j_\tau(\tau_j)}=\frac{\int_{x_{j-1}}^{x_j}\hat{\mathcal{V}}_x(x,\tau_j)dx}{\int_{x_{j-1}}^{x_j}\hat{v}_{\tau x}(x,\tau_j)dx}=\frac{\hat{\mathcal{V}}_x(x',\tau_j)}{\hat{v}_{\tau x}(x',\tau_j)},$$

for some $x'\in(x_{j-1},x_j)$, and similarly,

$$\frac{\hat{\mathcal{V}}^{j+1}(\tau_j)}{\hat{v}^{j+1}_\tau(\tau_j)}=\frac{\hat{\mathcal{V}}_x(x'',\tau_j)}{\hat{v}_{\tau x}(x'',\tau_j)}$$

for some $x''\in(x_j,x_{j+1})$. Hence by Condition 3 of Definition 1, $\hat{\Pi}^{j+1}_\tau(\tau_j)>0$. □

A.4 Convergence to the Continuum of Qualities

Recall that $\mathbb{P}$ is the set of all price functions ρ with $\rho(x_0)=0$ and $\rho\leq\bar{\rho}$, where $\bar{\rho}$ is bigger than the maximum willingness to pay for the highest quality relative to x_0. Wlog, $\rho\in\mathbb{P}$ is increasing

in x and left-continuous. We begin by defining a suitable notion of convergence of sets of price functions.

Definition 2 *Say that a sequence $(\mathbb{P}^n)$ of closed subsets of the closed subset $\mathbb{P}^0 \subseteq \mathbb{P}$ converges to $\mathbb{P}^0$ if for all $\rho \in \mathbb{P}^0$, there is a sequence (ρ^n) with each $\rho^n \in \mathbb{P}^n$ such that $\rho^n \to \rho$.*

Next, let us define a metric on the space of price functions ρ. Given ρ' and ρ'', the distance $d(\rho', \rho'')$ is the smallest number such that for each x, there is $\hat{x}$ within $d(\rho', \rho'')$ to the left of x with $\rho''(\hat{x}) \leq \rho'(x) + d(\rho', \rho'')$, and vice versa. Formally,

$$d(\rho', \rho'') = \min\{\delta | \rho''(\max(x - \delta, 0)) \leq \rho'(x) + \delta \text{ and } \rho'(\max(x - \delta, 0)) \leq \rho''(x) + \delta \text{ for all } x \in [0, 1]\}.$$

The minimum is well-defined since ρ is left-continuous. It is straightforward to check that d is a metric. Indeed, d is the Levy metric (Billingsley (1995); Problem 14.5, p.198) adjusted to take account of the fact that x lies in a compact support, and we will refer to it as such henceforth. Under the assumptions made, the set of price functions ρ is compact in this metric.

Lemma 2 *The consumer's value function $V(\rho, \psi, \theta) \equiv \max_{x \in [x_0, 1]}(v(x, \psi, \theta) - \rho(x))$ is continuous. The best-response correspondence $X(\rho, \psi, \theta)$ is upper hemicontinuous in ρ and θ.*[58]

Proof We first establish that V is continuous. Let $(\psi^n, \theta^n, \rho^n) \to (\psi, \theta, \rho)$. Let us show first that $V(\rho, \psi, \theta) \geq \limsup_n V(\rho^n, \psi^n, \theta^n)$. Choose a subsequence $(\psi^{n_k}, \theta^{n_k}, \rho^{n_k})$ such that $V(\rho^{n_k}, \psi^{n_k}, \theta^{n_k})$ converges to $\limsup_n V(\rho^n, \psi^n, \theta^n)$. For each n_k, choose $x^{n_k} \in X(\rho^{n_k}, \psi^{n_k}, \theta^{n_k})$. Passing to a further subsequence if necessary, assume that x^{n_k} converges to some $\tilde{x} \in [0, 1]$. Let $\hat{x}^{n_k} = \max(x^{n_k} - d(\rho^{n_k}, \rho), 0)$, and note that since $\hat{x}^{n_k}$ is a feasible choice,

$$\begin{aligned} V(\rho, \psi, \theta) &\geq v(\hat{x}^{n_k}, \psi, \theta) - \rho(\hat{x}^{n_k}) \\ &= v(x^{n_k}, \psi^{n_k}, \theta^{n_k}) - \rho^{n_k}(x^{n_k}) + v(\hat{x}^{n_k}, \psi, \theta) - v(x^{n_k}, \psi^{n_k}, \theta^{n_k}) + \rho^{n_k}(x^{n_k}) - \rho(\hat{x}^{n_k}) \\ &\geq V(\rho^{n_k}, \psi^{n_k}, \theta^{n_k}) + v(\hat{x}^{n_k}, \psi, \theta) - v(x^{n_k}, \psi^{n_k}, \theta^{n_k}) - d(\rho^{n_k}, \rho), \end{aligned}$$

where the third inequality uses that $V(\rho^{n_k}, \psi^{n_k}, \theta^{n_k}) = v(x^{n_k}, \psi^{n_k}, \theta^{n_k}) - \rho^{n_k}(x^{n_k})$ and that $\rho(\hat{x}^{n_k}) \leq \rho^{n_k}(x^{n_k}) + d(\rho^{n_k}, \rho)$ by definition of d and by construction of $\hat{x}^{n_k}$. But then, since v is continuous with $\lim \hat{x}^{n_k} = \lim x^{n_k} = \tilde{x}$, and since $d(\rho^{n_k}, \rho) \to 0$, we can take limit on each side to arrive at $V(\rho, \psi, \theta) \geq \limsup_n V(\rho^n, \psi^n, \theta^n)$ as desired.

Showing that $V(\rho, \psi, \theta) \leq \liminf_n V(\rho^n, \psi^n, \theta^n)$ is similar. Following analogous steps, let x^{n_k}

[58] Since $v(x, \psi, \theta) - \rho(x)$ is not continuous in ρ, this is not an immediate consequence of the theorem of the maximum.

converge to $\check{x} \in X(\rho, \psi, \theta)$, and let $\check{x}^{n_k} = \max(\check{x} - d(\rho^{n_k}, \rho), 0)$, and observe that for all n_k,

$$\begin{aligned} V(\rho^{n_k}, \psi^{n_k}, \theta^{n_k}) &\geq v(\check{x}^{n_k}, \psi^{n_k}, \theta^{n_k}) - \rho^{n_k}(\check{x}^{n_k}) \\ &= v(\check{x}, \psi, \theta) - \rho(\check{x}) + v(\check{x}^{n_k}, \psi^{n_k}, \theta^{n_k}) - v(\check{x}, \psi, \theta) + \rho(\check{x}) - \rho^{n_k}(\check{x}^{n_k}) \\ &\geq V(\rho, \psi, \theta) + v(\check{x}^{n_k}, \psi^{n_k}, \theta^{n_k}) - v(\check{x}, \psi, \theta) - d(\rho^{n_k}, \rho). \end{aligned}$$

Thus, since v is continuous, $\rho^{n_k} \to \rho$, $d(\rho^{n_k}, \rho) \to 0$, and $V(\rho^{n_k}, \psi^{n_k}, \theta^{n_k}) \to \liminf_n V(\rho^n, \psi^n, \theta^n)$, it follows from taking limits on each side that $\liminf_n V(\rho^n, \psi^n, \theta^n) \geq V(\rho, \psi, \theta)$. Hence, V is continuous.

Now, let us show that X is upper hemicontinuous. To do so, let $(x^n, \rho^n, \psi^n, \theta^n) \to (x, \rho, \psi, \theta)$ where for each n, $x^n \in X(\rho^n, \psi^n, \theta^n)$. We desire to show $x \in X(\rho, \psi, \theta)$. So, choose any $\grave{x}$, and for each n, let $\grave{x}^n = \max(\grave{x} - d(\rho^n, \rho), 0)$. Since $x^n \in X(\rho^n, \psi^n, \theta^n)$, we have

$$v(x^n, \psi^n, \theta^n) - \rho^n(x^n) \geq v(\grave{x}^n, \psi^n, \theta^n) - \rho^n(\grave{x}^n), \tag{17}$$

for all n. We will show that this implies that $v(x, \psi, \theta) - \rho(x) \geq v(\grave{x}, \psi, \theta) - \rho(\grave{x})$. Since $\grave{x}$ is arbitrary, this would establish that $x \in X(\rho, \psi, \theta)$.

Since ρ is increasing and left-continuous, it is lower-semicontinuous (Rudin (1976) Theorem 4.29, p.95), and thus $-\rho$ is upper-semicontinuous. Hence, for any n, the Levy metric and the lower-semicontinuity of ρ yields $\limsup_n(-\rho^n(x^n)) \leq -\rho(x)$. Thus, $v(x, \psi, \theta) - \rho(x) \geq \limsup_n(v(x^n, \psi^n, \theta^n) - \rho^n(x^n))$, and so, since from (17), $\limsup_n(v(x^n, \psi^n, \theta^n) - \rho^n(x^n)) \geq \limsup_n(v(\grave{x}^n, \psi^n, \theta^n) - \rho^n(\grave{x}^n))$, we are done if $\limsup_n(v(\grave{x}^n, \psi^n, \theta^n) - \rho^n(\grave{x}^n)) \geq v(\grave{x}, \psi, \theta) - \rho(\grave{x})$ or, since v is continuous and $\grave{x}^n \to \grave{x}$, if $\limsup_n(-\rho^n(\grave{x}^n)) \geq -\rho(\grave{x})$. But, $-\rho^n(\grave{x}^n) = -\rho(\grave{x}) + \rho(\grave{x}) - \rho^n(\grave{x}^n) \geq -\rho(\grave{x}) - d(\rho^n, \rho)$, and we the result follows since $d(\rho^n, \rho) \to 0$. □

The next lemma tells us that for any given closed set $\mathbb{P}^0 \subseteq \mathbb{P}$, if we take a sequence $\mathbb{P}^n$ of increasingly fine approximation to $\mathbb{P}^0$ then anything the principal can do in $\mathbb{P}^0$ can come arbitrarily close to what can be done in $\mathbb{P}^n$.

Lemma 3 *The principal's payoff* $\Pi(\rho)$ *is continuous in* ρ.

Proof We claim first that the set of (ψ, θ) where $X(\rho, \psi, \theta)$ is singleton valued has full H-measure. To see this, note that since v is strictly supermodular in x and ψ, for each pair ψ'' and ψ' with $\psi'' > \psi'$, the smallest best response at ψ'' is at least as large as the largest best response at ψ', or formally, $\inf X(\rho, \psi'', \theta) \geq \sup X(\rho, \psi', \theta)$. But then, for each θ there is a countable set of values of ψ such that $X(\rho, \cdot, \theta)$ is unique except on this set (see Shannon, 1995). Since the distribution over ψ conditional on θ is atomless, with probability one conditional on θ, the set $X(\rho, \cdot, \theta)$ is a singleton. Since θ was arbitrary, the claim has been proved.

Fix $\grave{\rho}$ and $\grave{\rho}^n \to \grave{\rho}$, and fix any measurable selection $\grave{\chi}$ from $X(\grave{\rho}, \cdot)$ and $\grave{\chi}^n$ from $X^n(\grave{\rho}^n, \cdot)$, so that

$$\Pi(\grave{\rho}^n) = \int S(\grave{\rho}^n(\grave{\chi}^n(\psi,\theta)), \grave{\chi}^n(\psi,\theta), \psi, \theta) dH(\psi,\theta),$$

and similarly for $\Pi(\grave{\rho})$. Let (ψ, θ) be any type for whom $X(\grave{\rho}, \psi, \theta)$ has a unique element $\grave{x}$. Then, $\grave{\chi}^n(\psi,\theta) \to \grave{\chi}(\psi,\theta)$ by Lemma 2. But, also from Lemma 2, $V(\grave{\rho}^n, \psi, \theta) = v(\grave{\chi}^n(\psi,\theta), \psi, \theta) - \grave{\rho}^n(\grave{\chi}^n(\psi,\theta))$ converges to $V(\grave{\rho}, \psi, \theta)$, and since v is continuous, $v(\grave{\chi}^n(\psi,\theta), \psi, \theta)$ converges to $v(\grave{\chi}(\psi,\theta), \psi, \theta)$. But then, it follows that $\grave{\rho}^n(\grave{\chi}^n(\psi,\theta)) \to \grave{\rho}(\grave{\chi}(\psi,\theta))$. Hence, since S is continuous, $S(\grave{\rho}^n(\grave{\chi}^n(\psi,\theta)), \grave{\chi}^n(\psi,\theta), \psi, \theta) \to S(\grave{\rho}(\grave{\chi}(\psi,\theta)), \grave{\chi}(\psi,\theta), \psi, \theta)$. By the Lebesgue dominated convergence theorem, since S is bounded, and since the set of (ψ, θ) where $X(\rho, \psi, \theta)$ is singleton valued has full H-measure, $\Pi(\grave{\rho}^n) \to \Pi(\grave{\rho})$, and we are done. □

The proof of the following proposition is now immediate from Lemmas 2 and 3:

Proposition 9 (Convergence) *Let $\mathbb{P}^0$ be a closed set, and let $\mathbb{P}^n \to \mathbb{P}^0$. Then, the principal's maximized payoff under $\mathbb{P}^n$ converges to her maximized payoff under $\mathbb{P}^0$. Further, if $\rho^n \to \hat{\rho}$ is any convergent sequence of optimal solutions for the principal given $\mathbb{P}^n$, then $\hat{\rho}$ is optimal for the principal in $\mathbb{P}^0$, and the payoff to the consumer of each type converges to that under $\hat{\rho}$.*

Appendix B Online Appendix

B.1 Justifying Assumptions 2 and 3

In this section, we present mild conditions that guarantee Assumptions 2 and 3. We begin with Assumption 2, and then turn to Assumption 3 at the end of the section. The first condition justifying Assumption 2 is that for some coordinate n of Θ, the ratio $v_{\theta_n x}/v_{\psi x}$ is strictly monotone in x. Equivalent is that, if for given (ψ, θ) we vary x and plot the resulting pairs (v_{θ_n}, v_ψ), then the curve traced out is either strictly concave or strictly convex. With this condition, and some minimal differentiability, we show that Assumption 2 will be satisfied regardless of the quality vector $\{x_j\}$. The condition guarantees that the level sets for different quality increments cross at an angle, and so coincide only on small sets of θ.

While this condition is conceptually simple, it is hard to check in the healthcare model. So, we provide a complementary proposition. It strengthens the differentiability condition, but requires only that there is no (ψ, θ) such that the pairs (v_{θ_n}, v_ψ) form a straight line as x is varied. If so, then Assumption 2 holds unless one was remarkably unlucky in ones choice of the set X of available quality vectors, in a sense made precise. The non-linearity condition is trivial to verify in the health care setting.

To interpret Assumption 2, let $\Theta_1 \subseteq \Theta$ be the set where for some ψ, (ψ, θ) has willingness to pay τ_1 for x_m versus x_ℓ, and let $\psi_1(\theta)$ be that value. That is, on Θ_1,

$$v(x_m, \psi_1(\theta), \theta) - v(x_\ell, \psi_1(\theta), \theta) = \tau_1,$$

where since $v(x_m, \cdot, \theta) - v(x_\ell, \cdot, \theta)$ is strictly increasing, ψ_1 is uniquely defined. Define Θ_2 similarly as the set where for some ψ, (ψ, θ) is willing to pay τ_2 for x_h versus x_m, and let ψ_2 be defined by

$$v(x_h, \psi_2(\theta), \theta) - v(x_m, \psi_2(\theta), \theta) = \tau_2.$$

Note that $\hat{\Theta} \subseteq \Theta_1 \cap \Theta_2$. We will look for conditions under which $H(\hat{\Theta}) = 0$, which is to say that ψ_1 and ψ_2 do not coincide over some positive measure set of types. They translate to conditions under which ψ_1 and ψ_2 "twist" relative to each other as θ varies.

Proposition 10 (Primitives for Assumption 2) *Assume that v is $\mathcal{C}^2$ and that for some n, either $\frac{v_{\theta_n x}(\cdot,\psi,\theta)}{v_{\psi x}(\cdot,\psi,\theta)}$ is strictly monotone, or for some $k \neq n$, $v_{\theta_k}(\cdot, \psi, \theta)$ and $\frac{v_{\theta_n x}(\cdot,\psi,\theta)}{v_{\theta_k x}(\cdot,\psi,\theta)}$ are strictly monotone. Then, Assumption 2 is satisfied.*

Proof The proof consists of several steps.

Step 1 Let I be an open interval containing $[\underline{\psi}, \bar{\psi}]$. Let us extend v to $I \times \Theta$. For $\psi > \bar{\psi}$, define

$$v(x, \psi, \theta) = v(x, \bar{\psi}, \theta) + \left(\psi - \bar{\psi}\right) v_\psi(x, \bar{\psi}, \theta),$$

and for $\psi < \underline{\psi}$ define

$$v(x, \psi, \theta) = v(x, \underline{\psi}, \theta) + \left(\psi - \underline{\psi}\right) v_\psi(x, \underline{\psi}, \theta).$$

Since v is $\mathcal{C}^2$ on T, its extension is $\mathcal{C}^1$ on $I \times \Theta$.

Step 2 Everywhere on $\Psi \times \Theta$, $v_{x\psi} > 0$, and by definition, for $\theta \in \Theta_1$, $\psi_1(\theta)$ exists. Hence by the implicit function theorem, for any $\theta \in \Theta_1$, $\psi_1(\cdot)$ is uniquely defined on a neighborhood in $I \times \Theta$ of $\hat{\theta}$, with

$$\psi_{1\theta_n}(\theta) = -\frac{v_{\theta_n}(x_m, \psi_1(\theta), \theta) - v_{\theta_n}(x_\ell, \psi_1(\theta), \theta)}{v_\psi(x_m, \psi_1(\theta), \theta) - v_\psi(x_\ell, \psi_1(\theta), \theta)}, \tag{18}$$

and similarly for each $\hat{\theta} \in \Theta_2$ there is a neighborhood in $I \times \Theta$ of $\hat{\theta}$ on which

$$\psi_{2\theta_n}(\theta) = -\frac{v_{\theta_n}(x_h, \psi_2(\theta), \theta) - v_{\theta_n}(x_m, \psi_2(\theta), \theta)}{v_\psi(x_h, \psi_2(\theta), \theta) - v_\psi(x_m, \psi_2(\theta), \theta)}. \tag{19}$$

Since v is $\mathcal{C}^1$, so are ψ_1 and ψ_2.

Step 3 Let $\tilde{\psi} = \psi_2 - \psi_1$, so that $\hat{\Theta} \subseteq \Theta$ is the set where $\tilde{\psi} = 0$. We claim that if $H(\hat{\Theta}) > 0$ then there exists $\theta \in \hat{\Theta}$ with $\tilde{\psi}_{\theta_n}(\theta) = 0$ for all n. Observe first that for each $\theta \in \Theta \backslash \hat{\Theta}$, continuity of $\tilde{\psi}$ implies that there is a neighborhood $N(\theta)$ of θ such that $N(\theta) \cap \hat{\Theta} = \varnothing$ and hence $H(\hat{\Theta}|N(\theta)) = 0$. So, assume the claim is false, and consider any $\theta \in \hat{\Theta}$. Since the claim is false, there is n such that $\tilde{\psi}_{\theta_n}(\theta) \neq 0$, and so, since $\tilde{\psi}$ is $\mathcal{C}^1$, there is a neighborhood $\hat{N} \subseteq \Theta$ of θ, which can be taken to be a cube, where $\tilde{\psi}_{\theta_n} \neq 0$. Consider $\tilde{\psi}$ restricted to $\hat{N}$. By the implicit function theorem, if we let $\hat{N}_{-n}$ be the projection of $\hat{N}$ onto dimensions other than n, then for all $\grave{\theta}_{-n}$ on a neighborhood $\tilde{N}_{-n}$ of θ_{-n} which we can take to be contained within $\hat{N}_{-n}$, there is a unique $\hat{\theta}_n(\grave{\theta}_{-n}) \in \hat{N}_n$ such that $\tilde{\psi}(\hat{\theta}_n(\grave{\theta}_{-n}), \grave{\theta}_{-n}) = 0$. Let $N(\theta)$ be $\tilde{N}_{-n} \times \hat{N}_n$, observing that this is a neighborhood of θ. But then, recalling that $h(\theta_n, \theta_{-n})$ can be written as $h(\theta_n|\theta_{-n})h(\theta_{-n})$, so that the conditional probability of any given value of θ_n given θ_{-n} is 0, we have that $H(\hat{\Theta}|N(\theta)) = 0$. Thus, for each $\theta \in \Theta$ we have $H(\hat{\Theta}|N(\theta)) = 0$. But, $\hat{\Theta}$ is closed and hence compact and so the collection $N(\cdot)$, which covers $\hat{\Theta}$, has a finite subcover of $\hat{\Theta}$. Since on each element $\grave{N}$ of that subcover, $H(\hat{\Theta}|\grave{N}) = 0$, it follows that $H(\hat{\Theta}) = 0$.

Step 4 Assume that $H(\hat{\Theta}) > 0$, and, using Step 3, choose a point θ where $\psi_1(\theta) = \psi_2(\theta)$ and $\psi_{1\theta_n}(\theta) = \psi_{2\theta_n}(\theta)$. Then, by (18) and (19) we have

$$\frac{v_{\theta_n}(x_m, \psi_1(\theta), \theta) - v_{\theta_n}(x_\ell, \psi_1(\theta), \theta)}{v_\psi(x_m, \psi_1(\theta), \theta) - v_\psi(x_\ell, \psi_1(\theta), \theta)} + \psi_{1\theta_n} = 0$$

and

$$\frac{v_{\theta_n}(x_h, \psi_1(\theta), \theta) - v_{\theta_n}(x_m, \psi_1(\theta), \theta)}{v_\psi(x_h, \psi_1(\theta), \theta)) - v_\psi(x_m, \psi_1(\theta), \theta)} + \psi_{1\theta_n} = 0$$

where the second equation uses $\psi_1(\theta) = \psi_2(\theta)$ and $\psi_{1\theta_n}(\theta) = \psi_{2\theta_n}(\theta)$. Thus,

$$\frac{v_{\theta_n}(x_m, \psi_1(\theta), \theta) - v_{\theta_n}(x_\ell, \psi_1(\theta), \theta)}{v_\psi(x_m, \psi_1(\theta), \theta) - v_\psi(x_\ell, \psi_1(\theta), \theta)} = \frac{v_{\theta_n}(x_h, \psi_1(\theta), \theta) - v_{\theta_n}(x_m, \psi_1(\theta), \theta)}{v_\psi(x_h, \psi_1(\theta), \theta) - v_\psi(x_m, \psi_1(\theta), \theta)}. \tag{20}$$

Step 5 Assume that $\frac{v_{\theta_n x}(\cdot, \psi, \theta)}{v_{\psi x}(\cdot, \psi, \theta)}$ is strictly monotone. By Cauchy's mean-value theorem, the *lhs* of (20) is equal to $\frac{v_{\theta_n x}(x_1, \psi_1(\theta), \theta)}{v_{\psi x}(x_1, \psi_1(\theta), \theta)}$ for some $x_1 \in (x_\ell, x_m)$ while the *rhs* of (20) is equal to $\frac{v_{\theta_n x}(x_1, \psi_1(\theta), \theta)}{v_{\psi x}(x_1, \psi_1(\theta), \theta)}$ for some $x_2 \in (x_m, x_h)$. Since $\frac{v_{\theta_n x}(\cdot, \psi, \theta)}{v_{\psi x}(\cdot, \psi, \theta)}$ is strictly monotone, this is a contradiction.

Step 6 Assume that for some $k \neq n$, $v_{\theta_k}(\cdot, \psi, \theta)$ is strictly monotone, and $\frac{v_{\theta_n x}(\cdot, \psi, \theta)}{v_{\theta_k x}(\cdot, \psi, \theta)}$ is strictly monotone. Then,

$$\frac{v_{\theta_k}(x_m, \psi_1(\theta), \theta) - v_{\theta_k}(x_\ell, \psi_1(\theta), \theta)}{v_\psi(x_m, \psi_1(\theta), \theta) - v_\psi(x_\ell, \psi_1(\theta), \theta)} = \frac{v_{\theta_k}(x_h, \psi_1(\theta), \theta) - v_{\theta_k}(x_m, \psi_1(\theta), \theta)}{v_\psi(x_h, \psi_1(\theta), \theta) - v_\psi(x_m, \psi_1(\theta), \theta)},$$

where all numerators and denominators in this expression are non-zero. But then, dividing each

side of (20) by the same side of this expression and canceling common terms, we have

$$\frac{v_{\theta_n}(x_h,\psi_1(\theta),\theta)-v_{\theta_n}(x_m,\psi_1(\theta),\theta)}{v_{\theta_k}(x_h,\psi_1(\theta),\theta)-v_{\theta_k}(x_m,\psi_1(\theta),\theta)}=\frac{v_{\theta_n}(x_m,\psi_1(\theta),\theta)-v_{\theta_n}(x_\ell,\psi_1(\theta),\theta)}{v_{\theta_k}(x_m,\psi_1(\theta),\theta)-v_{\theta_k}(x_\ell,\psi_1(\theta),\theta)}.$$

Since $\frac{v_{\theta_n x}(\cdot,\psi,\theta)}{v_{\theta_k x}(\cdot,\psi,\theta)}$ is strictly monotone, this is a contradiction as in Step 5. □

The following provides a complementary take on Assumption 2. It says that if v is analytic, and if $v_{\theta_n}(\cdot,\psi,\theta)$ is never linear in $v_\psi(\cdot,\psi,\theta)$ then for "almost all" specifications of the vector $\{x_j\}$ of quality levels, Assumption 2 will hold, which is to say that only for a remarkably unlucky choice of $\{x_j\}$ might there be a problem.

Proposition 11 *Assume that v is analytic and that for some n, $\frac{v_{\theta_n x}(\cdot,\psi,\theta)}{v_{\psi x}(\cdot,\psi,\theta)}$ is not constant for any (ψ,θ), or equivalently, that $v_{\theta_n}(\cdot,\psi,\theta)$ is not affine in $v_\psi(\cdot,\psi,\theta)$. Then, for each j and for each specification x_{-j} of $J-1$ contracts, there are at most a countable set of specifications of x_j where Assumption 2 can fail.*[59]

To prove this proposition, we begin with a Lemma.

Lemma 4 *Fix a density r on $[0,1]^M$, an open subset D of $[0,1]^M$, and a $\mathcal{C}^1$ function $\lambda: D\to\mathbb{R}$. Then, for all $\tau\in\mathbb{R}$,*

$$\int_{y\in D}\mathbf{1}_{\lambda(y)=\tau}r(y)dy\leq\int_{y\in D}\mathbf{1}_{(\lambda_{y_1}(y),\ldots,\lambda_{y_M}(y))=\mathbf{0}}r(y)dy.$$

Proof Fix m, and let $Y(m)\equiv\{y\in D|\lambda_{y_m}\neq 0\}$. Then, $\int_{Y(m)}\mathbf{1}_{\lambda(y)=\tau}r(y)dy$ is equal to

$$\int_{y_{-m}}\left(\int_{y_m\in\{y_m|(y_m,y_{-m})\in Y(m)\}}\mathbf{1}_{\lambda(y)=\tau}r(y_m|y_{-m})dy_m\right)r(y_{-m})dy_{-m}. \tag{21}$$

But, for any given y_{-m}, $\{y_m|(y_m,y_{-m})\in Y(m)\}$ can be divided into a countable number of disjoint non-empty sets. To see this, note first that the set $\{y_m|(y_m,y_{-m})\in D\}$ is an open subset of $[0,1]$ and hence consists of a countable union of disjoint open intervals. Within any such interval $\mathcal{I}$, and for any point $y_m\in\mathcal{I}$ where $\lambda_{y_m}(y_m,y_{-m})\neq 0$ take the largest interval around y_m that is contained within $\mathcal{I}$ and such that $\lambda_{y_m}(y_m',y_{-m})\neq 0$. This interval has strictly positive length since $\mathcal{I}$ is open and since λ is $\mathcal{C}^1$. Hence there are at most a countable set of such maximal intervals. But, on each such interval, $\lambda(\cdot,y_{-m})$ is strictly monotone, and so can cross τ at most

[59] Thus, if one specifies v and then follows any stochastic process of choosing the x such that the conditional distribution of x_j given x_{-j} is absolutely continuous with respect to Lebesgue measure, then the probability that Assumption 2 is violated is zero.

once. Hence, on the set $\{y_m|(y_m, y_{-m}) \in Y(m)\}$, λ can equal τ at most a countable number of times, which, given that $r(y_m|y_{-m})$ is a density, implies that the inner integral in (21) is equal to zero. Hence the outer integral is zero as well, and we have $\int_{Y(m)} \mathbf{1}_{\lambda(y)=\tau} r(y)dy = 0$.

To complete the argument, note that $\{(\lambda_{y_1}(y), \ldots, \lambda_{y_M}(y)) = \mathbf{0}\} = D \backslash \cup_m Y(m)$. Hence,

$$\begin{aligned}
\int_{y\in D} \mathbf{1}_{\lambda(y)=\tau} r(y)dy &= \int_{D\backslash\cup_m Y(m)} \mathbf{1}_{\lambda(y)=\tau} r(y)dy + \int_{\cup_m Y(m)} \mathbf{1}_{\lambda(y)=\tau} r(y)dy \\
&= \int_{\{(\lambda_{y_1}(y),\ldots,\lambda_{y_M}(y))=\mathbf{0}\}} \mathbf{1}_{\lambda(y)=\tau} r(y)dy + \int_{\cup_m Y(m)} \mathbf{1}_{\lambda(y)=\tau} r(y)dy \\
&\leq \int_{\{(\lambda_{y_1}(y),\ldots,\lambda_{y_M}(y))=\mathbf{0}\}} r(y)dy + \sum_m \int_{Y(m)} \mathbf{1}_{\lambda(y)=\tau} r(y)dy \\
&= \int_{y\in D} \mathbf{1}_{(\lambda_{y_1}(y),\ldots,\lambda_{y_M}(y))=\mathbf{0}} r(y)dy
\end{aligned}$$

and we are done. □

Proof of Proposition 11 Using Whitney's extension theorem, extend v to an analytic function on a set $\Psi^+ \times \Theta$, where Ψ^+ is an open interval containing Ψ, and where Ψ^+ is defined small enough that $v_{x\psi}$ remains strictly positive on $\Psi^+ \times \Theta$. Fix $j \neq 0$, and x_{-j}, and consider the situation where in the notation of Assumption 2, x_j is in the role of x_h, so that x_l and $x_m > x_\ell$ correspond to two indices below j. Fix also $\tau_1 > 0$. Let ψ_1 be defined as usual by

$$v(x_m, \psi_1(\theta), \theta) - v(x_\ell, \psi_1(\theta), \theta) = \tau_1$$

wherever a solution to this equation exists with $\psi_1 \in \Psi^+$. Since $v_{x\psi} > 0$ on $\Psi^+ \times \Theta$, ψ_1 is analytic on its domain by the analytic version of the implicit function theorem. Note that by the implicit function theorem, for any $\theta \in \Theta$, if ψ_1 exists (i.e., is contained within Ψ^+), then ψ_1 is defined on an open neighborhood around θ, and hence the domain of ψ_1 is open.

Now, consider any given value for $x_j \in (x_{j-1}, x_{j+1})$ if $j < J$ and $x_j \in (x_{j-1}, 1]$ if $j = J$. Define

$$\Delta(x_j, \theta) = v(x_j, \psi_1(\theta), \theta) - v(x_m, \psi_1(\theta), \theta)$$

for all θ in the domain of ψ_1. Then, Assumption 2 holds as long as for every $\tau > 0$, $G(\{\theta|\Delta(x_j, \theta) = \tau\}) = 0$. But, since the domain of ψ_1 is open, so is the domain of Δ, and thus by Lemma 4 $\max_\tau G(\{\theta|\Delta(x_j, \theta) = \tau\}) \leq G(\tilde{\Theta}_{x_j})$, where $\tilde{\Theta}_{x_j} \equiv \{\theta|\Delta_{\theta_1}(x_j, \theta), \ldots, \Delta_{\theta_N}(x_j, \theta)) = \mathbf{0}\}$.

Now, fix $\delta > 0$, and assume that there is a countably infinite collection $\left\{x_j^k\right\}_{k=1}^\infty$ of values of x_j where for each k, $G(\tilde{\Theta}_{x_j^k}) \geq \delta$. Let $\Theta_k^s = \cup_{k' \geq k} \tilde{\Theta}_{x_j^{k'}}$ be the set of types who appear in $\tilde{\Theta}_{x_j^{k'}}$ at some point on or after k. By construction $\Theta_{k+1}^s \subseteq \Theta_k^s$. Let $\Theta^* = \cap_{k=1}^\infty \Theta_k^s$ be the set of types who

are in $\tilde{\Theta}_{x_j^k}$ an infinite number of times. Billingsley Thm 2.1 tells us that $G(\Theta^*) = \lim G(\Theta_k^s) \geq \delta$.

Choose any $\theta \in \Theta^*$. Then, for an infinite number of distinct x_j^k, we have $(\Delta_{\theta_1}(x_j^k, \theta), \ldots, \Delta_{\theta_N}(x_j^k, \theta)) = \mathbf{0}$, and so, since each Δ_{θ_j} is analytic $(\Delta_{\theta_1}(\cdot, \theta), \ldots, \Delta_{\theta_N}(\cdot, \theta)) \equiv \mathbf{0}$. But,

$$\Delta_{\theta_j}(\cdot, \theta) = (v_\psi(\cdot, \psi_1(\theta), \theta) - v_\psi(x_m, \psi_1(\theta), \theta))\, \psi_{1\theta_j}(\theta) + v_{\theta_j}(\cdot, \psi_1(\theta), \theta) - v_{\theta_j}(x_m, \psi_1(\theta), \theta) \equiv 0,$$

and so, rearranging,

$$-\frac{v_{\theta_j}(\cdot, \psi_1(\theta), \theta) - v_{\theta_j}(x_m, \psi_1(\theta), \theta)}{v_\psi(\cdot, \psi_1(\theta), \theta) - v_\psi(x_m, \psi_1(\theta), \theta)} \equiv \psi_{1\theta_j}(\theta),$$

which is to say that the *lhs* is independent of x_j. We thus have

$$\begin{aligned}
0 &\equiv \left(\frac{v_{\theta_j}(\cdot, \psi_1(\theta), \theta) - v_{\theta_j}(x_m, \psi_1(\theta), \theta)}{v_\psi(\cdot, \psi_1(\theta), \theta) - v_\psi(x_m, \psi_1(\theta), \theta)} \right)_{x_j} \\
&=_s v_{\theta_j x}(\cdot, \psi_1(\theta), \theta)\, (v_\psi(x_j, \psi_1(\theta), \theta) - v_\psi(x_m, \psi_1(\theta), \theta)) \\
&- \left(v_{\theta_j}(\cdot, \psi_1(\theta), \theta) - v_{\theta_j}(x_m, \psi_1(\theta), \theta)\right) v_{\psi x}(x_j, \psi_1(\theta), \theta) \\
&=_s \frac{v_{\theta_j x}(\cdot, \psi_1(\theta), \theta)}{v_{\psi x}(\cdot, \psi_1(\theta), \theta)} - \frac{v_{\theta_j}(\cdot, \psi_1(\theta), \theta) - v_{\theta_j}(x_m, \psi_1(\theta), \theta)}{v_\psi(\cdot, \psi_1(\theta), \theta) - v_\psi(x_m, \psi_1(\theta), \theta)} \\
&= \frac{v_{\theta_j x}(\cdot, \psi_1(\theta), \theta)}{v_{\psi x}(\cdot, \psi_1(\theta), \theta)} - \psi_{1\theta_j}(\theta),
\end{aligned}$$

where the divisions are valid since $v_{\psi x} > 0$ and $x_j > x_m$. But then, $\frac{v_{\theta_j x}(\cdot, \psi_1(\theta), \theta)}{v_{\psi x}(\cdot, \psi_1(\theta), \theta)}$ is everywhere constant, violating the premise.[60] Hence, there is at most a finite set of x_j where $\max_\tau G(\{\theta | \Delta(x_j^k, \theta) = \tau\}) \geq \delta$. But then, by considering a sequence $\delta^k \to 0$, and taking the union of points where $\max_\tau G(\{\theta | \Delta(x_j^k, \theta) = \tau\}) \geq \delta^k$, we have that Assumption 2 holds when x_j plays the role of x_h for all but a countable set of values of x_j for any given specification of x_{-j}. Minor variations of the proof hold when x_j is in the role x_m or x_ℓ, and the set of j under consideration is finite, and so we are done. □

Finally, let us turn to Assumption 3. Note that by Lemma 4, the probability of θ such that $v(x'', \bar{\psi}, \theta) - v(x', \bar{\psi}, \theta)$ takes on any given value is at most the probability that $v_{\theta_j}(x'', \bar{\psi}, \theta) - v_{\theta_j}(x', \bar{\psi}, \theta) = 0$ for all j. Further, if v is analytic, then this probability is positive only if in fact $v(x'', \bar{\psi}, \cdot) - v(x', \bar{\psi}, \cdot)$ is a constant.

[60] Indeed, this holds everywhere on Θ^*, and so the premise can be weakened to say that there is no positive measure set of θ for which there is an associated ψ such that $\frac{v_{\theta_j x}(\cdot, \psi, \theta)}{v_{\psi x}(\cdot, \psi, \theta)}$ is constant.

B.2 Justifying Assumption 5

Let us interpret the condition that $\mathcal{V}_x/v_{\psi x}$ is strictly decreasing. For simplicity, set $w_G = 0$. Then,

$$\left(\frac{\mathcal{V}_x}{v_{\psi x}}\right)_x = \left(\frac{v_x - \gamma_x}{v_{x\psi}}\right)_x =_s (v_{xx} - \gamma_{xx})\, v_{x\psi} - (v_x - \gamma_x)\, v_{xx\psi}$$
$$=_s v_{xx} - v_x \frac{v_{xx\psi}}{v_{x\psi}} - \left(\gamma_{xx} - \gamma_x \frac{v_{xx\psi}}{v_{x\psi}}\right).$$

It is easy to show that $v_{xx} - v_x \frac{v_{xx\psi}}{v_{x\psi}}$ is strictly negative when v_x is strictly log-supermodular. So, it suffices that the bracketed term is positive. This is automatic if γ is increasing and convex and $v_{xx\psi} \leq 0$. When $v_{xx\psi} > 0$, it is equivalent to the condition that $\frac{\gamma_{xx}}{\gamma_x} - \frac{v_{xx\psi}}{v_{x\psi}} \geq 0$ or $\frac{\gamma_x}{v_{x\psi}}$ is increasing. This will follow if γ is sufficiently convex.

These are far from the only conditions that will give us that $\mathcal{V}_x/v_{\psi x}$ is strictly decreasing. Consider in a Mussa-Rosen fashion

$$v = \alpha(\psi, \theta)x - \beta(\psi, \theta)x^2$$

with $\alpha > 0$, $\beta > 0$ and $\beta_\psi < 0$ (so that $v_{x\psi} > 0$), and $w_G = 0$. Then, since $\mathcal{V}_x = v_x - \gamma_x$,

$$\frac{\mathcal{V}_x}{v_{x\psi}} = \frac{1}{-\beta_\psi}\frac{\alpha - \beta x - \gamma_x}{x} =_s \frac{\alpha - \gamma_x}{x} - \beta$$

which is trivially decreasing in x as long as either $\gamma_x < \alpha$ or γ_x/x is increasing in x.

B.3 *NC* in the Continuum: Proof of Theorem 4

Our derivation of *NC* in the finite case is elementary, as the principal's problem essentially reduces to a tractable optimization problem in J incremental prices. Deriving the necessary conditions for optimality is more subtle in the continuum case, since, for example, taking into account jumps in the allocation and differentiability issues call for some extra care.

In this section, we analyze the implications of what we call the first perturbation, which provides the analog of *NC* in the continuum case. In this perturbation, one chooses a base $\hat{x}$, a distance δ and a small ε. One then raises the slope of ρ by ε/δ over the interval $[\hat{x}, \hat{x} + \delta]$, but otherwise, they run parallel to each other. Formally, $\rho^\varepsilon(x) = \rho(x)$ for $x \leq \hat{x}$, $\rho^\varepsilon(x) = \rho(x) + \frac{\varepsilon}{\delta}(x - \hat{x})$ for $x \in (\hat{x}, \hat{x} + \delta)$, and $\rho^\varepsilon(x) = \rho(x) + \varepsilon$ for $x \geq \hat{x} + \delta$. We will analyze a fixed $\hat{x}$ throughout, and so will suppress it in various parts of the notation and, similarly, we will suppress as many of δ and θ as we can without confusion.

We will use the following notation which is a natural extension of the one we used in the finite case. Let $x^\varepsilon(\psi)$ be the set of optimal choices by ψ (given θ) to ρ^ε, and note that x^0 is unique at all but a countable number of values for ψ and on a measure-one subset of Θ x^0 never has more than two elements. Let $\psi^\varepsilon(x)$ be the set of ψ such that $x^\varepsilon(\psi) \geq x$ for all $\psi > \psi^\varepsilon(x)$ and $x^\varepsilon(\psi) \leq x$ for all $\psi < \psi^\varepsilon(x)$. Both ψ^ε and ψ^0 are unique except at the countable set of x where x^0 or x^ε is constant for an interval.

Throughout the analysis of the continuum case, besides our standing assumptions we will add the following one:

Assumption 6 (v_x Bounded away from Zero) *The derivative of v with respect to x is bounded away from zero on $[0,1] \times \Psi \times \Theta$.*

B.3.1 Replacing ρ by a Nicer Price Function

In principle, ρ can be a very complicated object, with jumps, and no absolute continuity. Tracking the effect of perturbations to ρ is therefore forbidding. We begin by replacing ρ by a price function $\tilde{\rho}$ which agrees with ρ at any x that is actually chosen, and implements the same allocation, but is well-behaved. We will then use that $\tilde{\rho}$ as the basis for simple perturbations and consider their limits. Critically, the limit expressions that arise will only depend on the value of $\tilde{\rho}$ at qualities that are chosen facing ρ. Formally, we have the following result whose proof is in Section B.3.11.

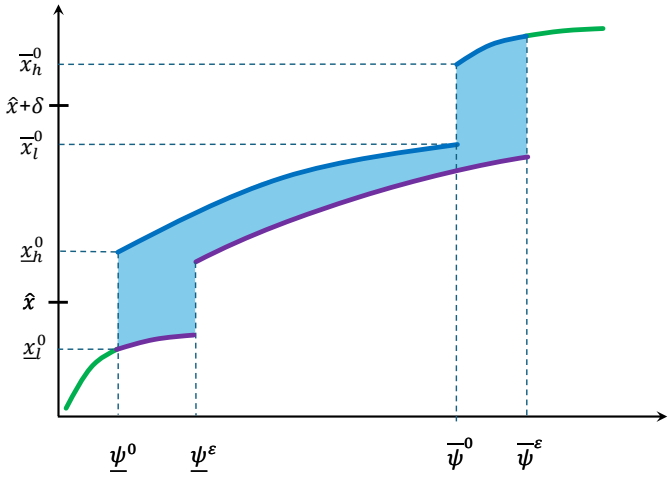
Lemma 5 (Tidying up ρ) *Assume that v_x is bounded away from zero on $[0,1] \times \Psi \times \Theta$. Then for every (ρ, χ) that is incentive compatible, there is a $\tilde{\rho}$ such that $(\tilde{\rho}, \chi)$ is incentive compatible and such that (i) $\tilde{\rho}$ agrees with ρ at any x that is in the range of χ, (ii) $\tilde{\rho}$ is strictly increasing, absolutely continuous, and (iii) $\tilde{\rho}$ has left-derivative that is continuous from the left and right-derivative that is continuous from the right.*

In the rest of the analysis of the continuum case, we shall assume that the premise on v_x is satisfied and that ρ has been replaced by a suitable $\tilde{\rho}$, and based the perturbations on $\tilde{\rho}$ instead of on ρ. To avoid carrying a new symbol, below we will continue to denote the price function by ρ instead of $\tilde{\rho}$, with the understanding that we are dealing with the tidied-up price function with the properties stated in Lemma 5.

B.3.2 A Useful Picture

Consider $\hat{x}$ and any $\delta > 0$. Through most of the proof we consider $\varepsilon > 0$, since the proof for the case of $\varepsilon < 0$ is similar and thus omitted. We will first consider what happens for any given θ, and then to integrate across θ. Let $\rho^\varepsilon(\cdot) \equiv \rho(\cdot, \hat{x}, \delta, \varepsilon)$, and recall the definition of $x^\varepsilon(\psi)$ and x^0.

Figure B.1. Optimal Choice



Notes: The consumer's optimal choice as a function of ψ for given θ when facing ρ and ρ^ε.

So, fix any such θ and consider Figure B.1. Let $\underline{\psi}^0$ be the highest ψ at which x^0 has an element at or below $\hat{x}$, and let $\underline{x}_l^0$ and $\underline{x}_h^0$ be the lowest and highest elements of $x^0(\underline{\psi}^0)$. These will agree if x^0 is continuous at $\underline{\psi}^0$, but in this example, $\underline{x}_l^0 < \hat{x} < \underline{x}_h^0$. By choice of θ, these are the only elements of $x^0(\underline{\psi}^0)$. Below $\underline{\psi}^0$, any ψ is choosing qualities facing ρ at or below $\hat{x}$, and so a fortiori, ψ's choice remains optimal when qualities above $\hat{x}$ become more expensive. Thus, x^0 and x^ε agree below $\underline{\psi}^0$, as denoted by the green line segment on this range. Similarly, let $\bar{\psi}^\varepsilon$ be the first point at which x^ε has an element at or above $\hat{x}+\delta$, and let $\bar{x}_l^0 \equiv x_h^0(\bar{\psi}^0)$ and $\bar{x}_l^\varepsilon \equiv x_h^\varepsilon(\bar{\psi}^\varepsilon)$. Since beyond $\bar{\psi}^\varepsilon$ any ψ is choosing qualities at or above $\hat{x}+\delta$, and since the relative prices of such contracts are the same under ρ and ρ^ε, it follows that x^0 and x^ε agree above $\bar{\psi}^\varepsilon$, as denoted by the green line segment on that range. Between $\underline{\psi}^0$ and $\bar{\psi}^\varepsilon$, the dark blue line denotes x^0 and the purple line denotes x^ε. Since ρ^ε is steeper than ρ, $x^\varepsilon \leq x^0$ on this range. There may be both jumps and flat spots in x^0 or x^ε, but none of this will turn out to matter.

B.3.3 Revenue Effects

As ρ is perturbed, two things are happening. First, some types are paying more. Second, some types are changing their choice. Let us deal with the first effect first. Fix a θ and consider the effect on each ψ when ε is raised by a small amount near zero. For ψ's that are choosing $\hat{x}$ or below, they take the same allocation as before and they pay the same. For ψ's that are choosing $x \in (\hat{x}, \hat{x}+\delta)$, they pay an extra amount $\varepsilon(x-\hat{x})/\delta$. Finally, for each $\psi > \psi^0(\hat{x}+\delta)$ when ε is sufficiently small, they choose their original choice but pay ε more. Hence, the rate of change

from revenue effects is

$$\int_{\hat{x}}^{\hat{x}+\delta} \frac{1}{\delta}\,(x-\hat{x})\,h(\psi^0(x,\theta)|\theta)dx + (1 - H(\psi^0(\hat{x}+\delta,\theta)|\theta)).$$

Note that $0 \leq x - \hat{x} \leq \delta$ in the integrand in the first term, and hence the first term will eventually vanish when we make δ go to zero, and thus for simplicity we will ignore it from now on. Thus, wlog the rate of change from revenue effects is, for each θ and $\hat{x}$, given by

$$1 - H(\psi^0(\hat{x}+\delta,\theta)|\theta). \tag{22}$$

B.3.4 A Change of Perspective

The second effect is the loss to the principal from those types who switch to a different choice. Let $L(\theta,\delta,\varepsilon)$ be the loss to the principal from ψ's for whom $x^0 > x^\varepsilon$. Suppressing θ and δ, it is given by

$$L(\varepsilon) = \int_{\underline{\psi}^0}^{\bar{\psi}^\varepsilon} \left(\mathcal{V}(x^0(\psi),\psi) - \mathcal{V}(x^\varepsilon(\psi),\psi)\right) h(\psi)d\psi = \int_{\underline{\psi}^0}^{\bar{\psi}^\varepsilon} \left(\int_{x^\varepsilon(\psi)}^{x^0(\psi)} \mathcal{V}_x(x,\psi)h(\psi)dx\right) d\psi. \tag{23}$$

That is, $L(\varepsilon)$ is the integral of $\mathcal{V}_x(x,\psi)h(\psi)$ over the shaded areas in Figure B.1. In this form, $L(\varepsilon)$ is very hard to calculate, because of the behavior of $x^0(\psi) - x^\varepsilon(\psi)$ as ψ and ε vary.

Recall the definition of $\psi^\varepsilon(x)$ and $\psi^0(x)$, and note that both ψ^ε and ψ^0 are unique except for the countable set of x where x^0 or x^ε is constant for an interval. For $x > \bar{x}_h^\varepsilon$, $\psi^0(x)$ is choosing over elements above $\hat{x}+\delta$, and so since ρ and ρ^ε are parallel over such x, the optimal choices facing ρ and ρ^ε agree, and so $\psi^\varepsilon(x) - \psi^0(x) = 0$, and similarly when $x \leq x_l^0(\psi^0(\hat{x}))$. Between $\underline{x}_l^0$ and $\bar{x}_h^\varepsilon$, $\psi^0(x)$ is the left-hand boundary of the shaded regions, and $\psi^\varepsilon(x)$ is the right-hand boundary. Thus, an equivalent expression for $L(\varepsilon)$ is

$$L(\varepsilon) = \int \left(\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi\right) dx. \tag{24}$$

Our first job is to understand the limiting behavior of $L(\varepsilon)/\varepsilon$. We will see that this is much simpler, because $\psi^\varepsilon(x) - \psi^0(x)$ is vastly more tractable than $x^0(\psi) - x^\varepsilon(\psi)$.

B.3.5 A Roadmap

To see where we are going, note that for each $\varepsilon > 0$, since $\mathcal{V}_x(x,\psi)h(\psi)$ is continuous, and since ψ^0 and ψ^ε are increasing and hence measurable,

$$\frac{1}{\varepsilon}\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi dx$$

is measurable in x. We will show shortly that there is $\kappa < \infty$ such that

$$\left|\frac{1}{\varepsilon}\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi\right| < \kappa$$

for all ε. But then, by the extended version of Fatou's Lemma to handle a negative but bounded integrand,

$$\begin{aligned}
&\int_{\underline{x}_l^0}^{\bar{x}_h^\varepsilon}\left(\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi\right)dx \\
&\leq \liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_{\underline{x}_l^0}^{\bar{x}_h^\varepsilon}\left(\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi\right)dx \\
&= \liminf_{\varepsilon\downarrow 0}\frac{L(\varepsilon)}{\varepsilon} \\
&\leq \limsup_{\varepsilon\downarrow 0}\frac{L(\varepsilon)}{\varepsilon} \\
&= \limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_{\underline{x}_l^0}^{\bar{x}_h^\varepsilon}\left(\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi\right)dx \\
&\leq \int_{\underline{x}_l^0}^{\bar{x}_h^\varepsilon}\left(\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi\right)dx.
\end{aligned}$$

and similarly when $\varepsilon \uparrow 0$. We will show that for all but a countable set of x, $\frac{1}{\varepsilon}\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi$ in fact converges, with its limit also measurable. But then, making similar arguments for $\varepsilon < 0$, it follows from the squeeze theorem that L is differentiable at $\varepsilon = 0$ with

$$L_\varepsilon(0) = \int\left(\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi\right)dx$$

where we will show that the integrand is also measurable in x, and provide a closed-form expression for $L_\varepsilon(0)$, which we can later integrate with respect to θ to complete the derivation of the principal's loss from those types who switch after prices are perturbed.

B.3.6 Some Properties of $(\psi^\varepsilon(x) - \psi^0(x))/\varepsilon$

To understand $\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi)h(\psi)d\psi$ and show its desired properties, we first study $(\psi^\varepsilon(x) - \psi^0(x))/\varepsilon$. We begin with a simple lemma.

Lemma 6 *Fix $x' < x''$ and $\psi' < \psi''$. Let $0 \leq I(x', x'') \leq \delta$ be the length of the intersection of (x', x'') with $(\hat{x}, \hat{x} + \delta)$. Then*

(i) If ψ' weakly prefers x'' to x' facing ρ and

$$\frac{1}{\varepsilon}\int_{\psi'}^{\psi''} \left[v_\psi(x'', \psi) - v_\psi(x', \psi)\right] d\psi > \frac{I(x', x'')}{\delta}$$

then ψ'' strictly prefers x'' to x' facing ρ^ε;

(ii) If ψ' weakly prefers x' to x'' facing ρ and

$$\frac{1}{\varepsilon}\int_{\psi'}^{\psi''} \left[v_\psi(x'', \psi) - v_\psi(x', \psi)\right] d\psi < \frac{I(x', x'')}{\delta}$$

then ψ'' strictly prefers x' to x'' facing ρ^ε.

That is if ψ' weakly prefers x'' to x' facing ρ and ψ'' is sufficiently far above ψ' then ψ'' strictly prefers x'' to x' facing ρ^ε, while if ψ' weakly prefers x' to x'' facing ρ and ψ'' is not very far above ψ' then ψ'' strictly prefers x' to x'' facing ρ^ε.

Proof Immediate since

$$v(x'', \psi'') - v(x', \psi'') = v(x'', \psi') - v(x', \psi') + \int_{\psi'}^{\psi''} \left[v_\psi(x'', \psi) - v_\psi(x', \psi)\right] d\psi$$

while

$$\rho^\varepsilon(x'') - \rho^\varepsilon(x') = \rho(x'') - \rho(x') + \frac{I(x', x'')}{\delta}\varepsilon,$$

since $I(x', x'')$ measures the length of the part of (x', x'') that is in $(\hat{x}, \hat{x}+\delta)$ and hence over which ρ^ε is steeper (by ε/δ) than ρ. □

Lemma 6 will allow us to derive bounds on $(\psi^\varepsilon(x) - \psi^0(x))/\varepsilon$. To this end, we will define the following set:

Definition 3 (Excluding a Countable Set of Qualities) *Let X be the set of x where (a) $\psi^0(x)$ is unique, (b) x is not the beginning or end of a jump in x^0, and (c) $x \notin \{\hat{x}, \hat{x} + \delta\}$.*

Since ψ^0 is an increasing function, and since a point where $\psi^0(x)$ is non-unique is a point where ψ^0 jumps there are at most a countable set of points where $\psi^0(x)$ is non-unique. Similarly there are at most a countable set of points where x^0 jumps, and hence at most a countable set of values of x that are the beginning or end of a jump. Hence, $[0,1]\backslash X$ is countable.

Fix $x \in X$ (and suppress it as possible in what follows). We seek to provide asymptotically tight upper and lower bounds on $(\psi^\varepsilon(x) - \psi^0(x))/\varepsilon$. If $\psi^\varepsilon(x) - \psi^0(x) = 0$, we are done. Otherwise, for any $\varepsilon < 1/2$, let $\psi^0_{u,\varepsilon} \equiv \psi^0(x) + \varepsilon(\psi^\varepsilon(x) - \psi^0(x))$, which is to say that we nudge $\psi^0(x)$ upwards by ε of the distance towards $\psi^\varepsilon(x)$. Let $x^0_{u,\varepsilon} \equiv x^0_l(\psi^0_{u,\varepsilon}(x))$ be the lower optimal choice facing ρ when $\psi^0(x)$ is nudged upwards. Note that $x^0_{u,\varepsilon} \geq x^0_h(\psi^0(x)) \geq x$, and hence $\lim_{\varepsilon \downarrow 0} x^0_{u,\varepsilon} = x^0_h(\psi^0(x))$ because any cluster point of $x^0_{u,\varepsilon}$ belongs to $x^0(\psi^0(x))$ and so can be no higher than $x^0_h(\psi^0(x))$. Indeed, since $\psi^0(x)$ is unique, for each $\varepsilon > 0$, $x^0_{u,\varepsilon} > x$. This is immediate if $x^0_h(\psi^0(x)) > x$, since $x^0_{u,\varepsilon} \geq x^0_h(\psi^0(x))$. If $x^0_h(\psi^0(x)) = x$ then $x^0_{u,\varepsilon} > x$ because $\psi^0(x)$ is unique.

Since $x^0_{u,\varepsilon} \in x^0(\psi^0_{u,\varepsilon})$, $\psi^0_{u,\varepsilon}$ weakly prefers $x^0_{u,\varepsilon} > x$ to $\underline{x}^\varepsilon \equiv x^\varepsilon_l(\psi^\varepsilon(x)) \leq x$ facing ρ, and so, since $\psi^0_{u,\varepsilon} < \psi^\varepsilon(x)$, we have by Lemma 6 that

$$\int_{\psi^0_{u,\varepsilon}}^{\psi^\varepsilon(x)} \left[v_\psi(x^0_{u,\varepsilon}, \psi) - v_\psi(\underline{x}^\varepsilon, \psi)\right] d\psi \leq \frac{I(\underline{x}^\varepsilon, x^0_{u,\varepsilon})}{\delta}\varepsilon, \tag{25}$$

since otherwise, $\psi^\varepsilon(x)$ would strictly prefer $x^0_{u,\varepsilon}$ to $\underline{x}^\varepsilon$ facing ρ^ε, contradicting that $\underline{x}^\varepsilon \in x^\varepsilon(\psi^\varepsilon(x))$.

Let

$$\underline{\Delta}(\varepsilon) \equiv \min_{\psi \in [\psi^0_{u,\varepsilon}, \psi^\varepsilon(x)]} \left[v_\psi(x^0_{u,\varepsilon}, \psi) - v_\psi(\underline{x}^\varepsilon, \psi)\right].$$

Then,

$$\int_{\psi^0_{u,\varepsilon}}^{\psi^\varepsilon} \left[v_\psi(x^0_{u,\varepsilon}, \psi) - v_\psi(\underline{x}^\varepsilon, \psi)\right] d\psi \geq (\psi^\varepsilon(x) - \psi^\varepsilon_u)\underline{\Delta}(\varepsilon),$$

and so from (25),

$$\frac{I(\underline{x}^\varepsilon, x^0_{u,\varepsilon})}{\delta}\varepsilon \geq (\psi^\varepsilon(x) - \psi^0_{u,\varepsilon})\underline{\Delta}(\varepsilon)$$

or, rearranging and using that $\psi^\varepsilon(x) - \psi^0_{u,\varepsilon} = (1-\varepsilon)(\psi^\varepsilon(x) - \psi^0(x))$, we have the upper bound

$$\frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \leq \frac{I(\underline{x}^\varepsilon, x^0_{u,\varepsilon})}{\delta(1-\varepsilon)} \frac{1}{\underline{\Delta}(\varepsilon)}. \tag{26}$$

To derive a lower bound, let $\psi^0_{d,\varepsilon} \equiv \psi^0(x) - \varepsilon(\psi^\varepsilon(x) - \psi^0(x))$ by nudging $\psi^0(x)$ so that (arguing as above) $x^0_{d,\varepsilon} \equiv x^0_l(\psi^0_{d,\varepsilon}) < x$ and $\bar{x}^\varepsilon \equiv x^\varepsilon_h(\psi^\varepsilon(x)) \geq x$. Then letting

$$\bar{\Delta}(\varepsilon) \equiv \max_{\psi \in [\psi^0_{d,\varepsilon}, \psi^\varepsilon(x)]} \left[v_\psi(\bar{x}^\varepsilon, \psi) - v_\psi(x^0_{d,\varepsilon}, \psi)\right],$$

and following the same steps as above,

$$\frac{\psi^\varepsilon(x)-\psi^0(x)}{\varepsilon} \geq \frac{I(x^0_{d,\varepsilon},\bar{x}^\varepsilon)}{(1+\varepsilon)\delta}\frac{1}{\bar{\Delta}(\varepsilon)}. \tag{27}$$

Let us first derive a global upper bound on $(\psi^\varepsilon(x)-\psi^0(x))/\varepsilon$ (for δ as fixed above, but letting $x, \hat{x}$ and θ vary). Let $v^{\min}_{x\psi} = \min_{[0,1]\times\Psi\times\Theta} v_{x\psi} > 0$. Then, for each ψ, and $x' \leq x''$,

$$v_\psi(x'',\psi) - v_\psi(x',\psi) = \int_{x'}^{x''} v_{\psi x}(x,\psi)dx \geq (x''-x')v^{\min}_{x\psi},$$

and so, regardless of x, $\hat{x}$ and θ

$$\underline{\Delta}(\varepsilon) \geq (x^0_{u,\varepsilon} - \underline{x}^\varepsilon)v^{\min}_{x\psi}.$$

Thus, since $I(\underline{x}^\varepsilon, x^0_{u,\varepsilon}) \leq x^0_{u,\varepsilon} - \underline{x}^\varepsilon$, (26) yields

$$\frac{\psi^\varepsilon(x)-\psi^0(x)}{\varepsilon} \leq \frac{I(\underline{x}^\varepsilon, x^0_{u,\varepsilon})}{(1-\varepsilon)\delta}\frac{1}{\underline{\Delta}(\varepsilon)} \leq \frac{x^0_{u,\varepsilon}-\underline{x}^\varepsilon}{(1-\varepsilon)\delta}\frac{1}{(x^0_{u,\varepsilon}-\underline{x}^\varepsilon)v^{\min}_{x\psi}} = \frac{1}{(1-\varepsilon)\delta v^{\min}_{x\psi}},$$

and so for $\varepsilon < 1/2$,

$$0 \leq \frac{\psi^\varepsilon(x)-\psi^0(x)}{\varepsilon} \leq \frac{2}{\delta v^{\min}_{x\psi}}. \tag{28}$$

where an obvious but important implication of this is that $\lim_{\varepsilon\downarrow 0}\psi^\varepsilon(x) = \psi^0(x)$. Note that this bound is independent of θ, $\hat{x}$, and x, and for each $\hat{x}$ and δ holds for all $x \in X$, and hence for all but a countable subset of $[0,1]$.

B.3.7 Tight Bounds and the Derivative of ψ^ε

We will now derive much tighter bounds on $(\psi^\varepsilon(x)-\psi^0(x))/\varepsilon$ for $x \in X$ and indeed show that for all $x \in X$, $\psi^\varepsilon(x)$ is differentiable in ε at $\varepsilon = 0$ and provide a closed-form expression for it. Since $x \in X$, there are only two cases. In the first case, x is interior to a jump in x^0. That is $x^0_h(\psi^0(x)) > x > x^0_l(\psi^0(x))$. In the second case, $x^0_h(\psi^0(x)) = x = x^0_l(\psi^0(x))$ since x is not in the interior of a jump of x^0 since the first case covers that, and is not an endpoint of a jump in x^0 by Definition 3.

Jumps. Fix x interior to a jump, and let $\bar{x} > x > \underline{x}$ be the endpoints of the jump in x^0 at $\psi^0(x)$. Then,

$$\lim_{\varepsilon\downarrow 0} x^0_{d,\varepsilon} = \lim_{\varepsilon\downarrow 0}\underline{x}^\varepsilon = \underline{x} \text{ and } \lim_{\varepsilon\downarrow 0} x^0_{u,\varepsilon} = \lim_{\varepsilon\downarrow 0}\bar{x}^\varepsilon = \bar{x},$$

since by choice of θ, $x^0(\psi^0(x))$ has only $\underline{x}$ and $\bar{x}$ as elements and thus since $x^0_{d,\varepsilon} < x$ and $\underline{x}^\varepsilon \leq x$ are optimal choices facing a price function arbitrarily close to ρ, and so each must converge to $\underline{x}$ and similarly for $x^0_{u,\varepsilon}$ and $\bar{x}^\varepsilon$. Thus,

$$\lim_{\varepsilon \downarrow 0} \underline{\Delta}(\varepsilon) = \lim_{\varepsilon \downarrow 0} \bar{\Delta}(\varepsilon) = v_\psi(\bar{x}, \psi^0(x)) - v_\psi(\underline{x}, \psi^0(x)) > 0.$$

Thus, from (26),

$$\limsup_{\varepsilon \downarrow 0} \frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \leq \limsup_{\varepsilon \downarrow 0} \frac{I(\underline{x}^\varepsilon, x^0_{u,\varepsilon})}{(1-\varepsilon)\delta} \frac{1}{\underline{\Delta}(\varepsilon)} = \frac{1}{\delta} \frac{I(\underline{x}, \bar{x})}{v_\psi(\bar{x}, \psi^0(x)) - v_\psi(\underline{x}, \psi^0(x))}$$

and similarly by (27),

$$\liminf_{\varepsilon \downarrow 0} \frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \geq \frac{1}{\delta} \frac{I(\underline{x}, \bar{x})}{v_\psi(\bar{x}, \psi^0(x)) - v_\psi(\underline{x}, \psi^0(x))}.$$

Using the same arguments for $\varepsilon < 0$, we obtain that $\psi^\varepsilon(x)$ is differentiable in ε at $\varepsilon = 0$, with derivative given by

$$\zeta(x) \equiv \frac{\partial}{\partial \varepsilon} \psi^\varepsilon(x) \Big|_{\varepsilon=0} = \frac{1}{\delta} \frac{I(x^0_l(\psi^0(x)), x^0_h(\psi^0(x)))}{v_\psi(x^0_h(\psi^0(x)), \psi^0(x)) - v_\psi(x^0_l(\psi^0(x)), \psi^0(x))}. \tag{29}$$

Non-Jumps. Since X excludes the beginning or end of jumps in x^0, the remaining case is that $x^0_l(\psi^0(x)) = x^0_h(\psi^0(x)) = x$. It follows that along any sequence where $\psi \to \psi^0(x)$ and $\varepsilon \downarrow 0$ and for any selection from $x^\varepsilon(\psi)$, the resulting sequence goes to x.

Assume first that $x \notin [\hat{x}, \hat{x} + \delta]$. Then, since $\underline{x}^\varepsilon$ and $x^0_{u,\varepsilon}$ converge to x, for all ε sufficient small, $I(\underline{x}^\varepsilon, x^0_{u,\varepsilon}) = 0$, and thus from (26)

$$\lim_{\varepsilon \downarrow 0} \frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} = 0$$

and so $\psi^\varepsilon(x)$ is differentiable in ε at $\varepsilon = 0$ with $\zeta(x) = 0$.

So, recalling that X excludes $\hat{x}$ and $\hat{x} + \delta$, let us turn to $x \in (\hat{x}, \hat{x} + \delta)$. Then, since $\underline{x}^\varepsilon$ and $x^0_{u,\varepsilon}$ converge to x, for small ε, $I(\underline{x}^\varepsilon, x^0_{u,\varepsilon}) = x^0_{u,\varepsilon} - \underline{x}^\varepsilon$. But also,

$$\underline{\Delta}(\varepsilon) \equiv \min_{\psi \in [\psi^0_{u,\varepsilon}, \psi^\varepsilon(x)]} \left[v_\psi(x^0_{u,\varepsilon}, \psi) - v_\psi(\underline{x}^\varepsilon, \psi) \right] = (x^0_{u,\varepsilon} - \underline{x}^\varepsilon) v_{x\psi}(x^*, \psi^*)$$

for some $(x^*, \psi^*) \in [x^0_{u,\varepsilon}, \underline{x}^\varepsilon] \times [\psi^0_{u,\varepsilon}, \psi^\varepsilon(x)]$ where we first pick ψ^* as a minimizer, and then since

$$\underline{\Delta}(\varepsilon) = v_\psi(x^0_{u,\varepsilon}, \psi^*) - v_\psi(\underline{x}^\varepsilon, \psi^*) = \int_{\underline{x}^\varepsilon}^{x^0_{u,\varepsilon}} v_{x\psi}(x', \psi^*)dx'$$

we use the mean value theorem to pick x^*. But then, from (26),

$$\frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \leq \frac{x^0_{u,\varepsilon} - \underline{x}^\varepsilon}{\delta} \frac{1}{(\underline{x}^\varepsilon - x^0_{u,\varepsilon}) v_{x\psi}(x^*, \psi^*)} = \frac{1}{\delta v_{x\psi}(x^*, \psi^*)}.$$

Finally, note that as $\varepsilon \to 0$, the rectangle $[x^0_{u,\varepsilon}, \underline{x}^\varepsilon] \times [\psi^0_{u,\varepsilon}, \psi^\varepsilon(x)]$ converges to the point $(x, \psi^0(x))$, and so

$$\limsup_{\varepsilon \downarrow 0} \frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \leq \frac{1}{\delta v_{x\psi}(x, \psi^0(x))}.$$

But, reasoning in the same manner from (27),

$$\liminf_{\varepsilon \downarrow 0} \frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \geq \frac{1}{\delta v_{x\psi}(x, \psi^0(x))}$$

and so, again appealing to the same arguments for $\varepsilon < 0$, we have that $\psi^\varepsilon(x)$ is differentiable with respect to ε at $\varepsilon = 0$, with

$$\zeta(x) = \frac{1}{\delta} \frac{1}{v_{x\psi}(x, \psi^0(x))}. \tag{30}$$

B.3.8 Closing the Circle on the Roadmap.

Now, let us solidify the argument made in Section B.3.5. Note that for all $\varepsilon < 1/2$, we can use the mean value theorem to establish that for some $\psi^* \in [\psi^0(x), \psi^\varepsilon(x)]$

$$\begin{aligned}
\left| \frac{1}{\varepsilon} \int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x, \psi) h(\psi) d\psi \right| &= \left| \frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \mathcal{V}_x(x, \psi^*) h(\psi^*) \right| \\
&\leq \left| \frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \right| |\mathcal{V}_x(x, \psi^*) h(\psi^*)| \\
&\leq \frac{2}{\delta v^{\min}_{x\psi}} \max_{[0,1] \times \Psi \times \Theta} |\mathcal{V}_x h|,
\end{aligned}$$

using (28) which gives the required uniform upper bound on $\left| \frac{1}{\varepsilon} \int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x, \psi) h(\psi) d\psi \right|$, and hence establishes the validity of applying Fatou's lemma as was done in Section B.3.5.

Note also that for each $x \in X$

$$\begin{aligned}\liminf_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi) h(\psi) d\psi &\geq \liminf_{\varepsilon \to 0} \left(\frac{\psi^\varepsilon(x) - \psi^0(x)}{\varepsilon} \min_{\psi \in [\psi^0(x), \psi^\varepsilon(x)]} (\mathcal{V}_x(x,\psi) h(\psi)) \right) \\ &= \zeta(x) \mathcal{V}_x(x, \psi^0(x)) h(\psi^0(x)),\end{aligned}$$

and similarly when we take a lim sup, and thus for all $x \in X$,

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_{\psi^0(x)}^{\psi^\varepsilon(x)} \mathcal{V}_x(x,\psi) h(\psi) d\psi = \zeta(x) \mathcal{V}_x(x, \psi^0(x)) h(\psi^0(x))$$

and so (once again using similar arguments for $\varepsilon < 0$) we have (reintroducing θ)

$$\begin{aligned}L_\varepsilon(0,\theta) &= \lim_{\varepsilon \to 0} \frac{L(\varepsilon,\theta)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int \left(\int_{\psi^0(x,\theta)}^{\psi^\varepsilon(x,\theta)} \mathcal{V}_x(x,\psi,\theta) h(\psi|\theta) d\psi \right) dx \\ &= \int \left(\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{\psi^0(x,\theta)}^{\psi^\varepsilon(x,\theta)} \mathcal{V}_x(x,\psi,\theta) h(\psi|\theta) d\psi \right) dx \\ &= \int \zeta(x,\theta) \mathcal{V}_x(x, \psi^0(x,\theta), \theta) h(\psi^0(x,\theta)) dx,\end{aligned}$$

where we remind the reader that we continue to suppress $\hat{x}$ and δ. We have thus established the claims in Section B.3.5.

B.3.9 Loss from Switchers

Consider

$$\mathcal{L}(\varepsilon) \equiv \int \frac{1}{\varepsilon} L(\varepsilon,\theta) h(\theta) d\theta.$$

We will show that $L(\varepsilon,\theta)$ is for each $\varepsilon > 0$ continuous in θ (and hence trivially measurable), and has a uniform bound across θ and ε, and so by Fatou's Lemma,

$$\begin{aligned}\int \left(\liminf_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} L(\varepsilon,\theta) \right) h(\theta) d\theta &\leq \liminf_{\varepsilon \downarrow 0} \int \frac{1}{\varepsilon} L(\varepsilon,\theta) h(\theta) d\theta \\ &\leq \limsup_{\varepsilon \downarrow 0} \int \frac{1}{\varepsilon} L(\varepsilon,\theta) h(\theta) d\theta \\ &\leq \int \left(\limsup_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} L(\varepsilon,\theta) \right) h(\theta) d\theta.\end{aligned}$$

But, as established in the previous section,

$$\liminf_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}L(\varepsilon,\theta)=\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}L(\varepsilon,\theta)=\int\zeta(x,\theta)\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))dx$$

and so $\mathcal{L}$ is differentiable at $\varepsilon=0$, with

$$\mathcal{L}_\varepsilon(0)=\int\left(\int\zeta(x,\theta)\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))dx\right)h(\theta)d\theta.$$

Note that via ζ, this expression depends on δ. So, let us consider a limit of $\mathcal{L}_\varepsilon(0)$ as $\delta\downarrow 0$.

Case 1. Suppose that for a given θ, $\hat{x}$ is part of a jump from $\underline{x}\leq\hat{x}$ to $\bar{x}>\hat{x}$ in $x^0(\cdot,\theta)$. Then, for all δ sufficiently small, $[\hat{x},\hat{x}+\delta]\subseteq[\underline{x},\bar{x}]$, and so for every $x\in(\underline{x},\bar{x})$, x is not a beginning or endpoint or a jump and hence on all but a countable subset of $(\underline{x},\bar{x})$ we have by (29) that

$$\begin{aligned}\zeta(x,\theta)&=\frac{1}{\delta}\frac{I(x_l^0(\psi^0(x,\theta),\theta),x_h^0(\psi^0(x,\theta),\theta))}{v_\psi(x_h^0(\psi^0(x,\theta),\theta),\psi^0(x,\theta),\theta)-v_\psi(x_l^0(\psi^0(x,\theta),\theta),\psi^0(x,\theta),\theta)}\\&=\frac{1}{\delta}\frac{\delta}{v_\psi(\bar{x},\psi^0(x,\theta),\theta)-v_\psi(\underline{x},\psi^0(x,\theta),\theta)}\end{aligned}$$

and thus, for $x\in(\underline{x},\bar{x})$

$$\zeta(x,\theta)\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))=\frac{\mathcal{V}_x(x,\psi^0(x,\theta),\theta)}{v_\psi(\bar{x},\psi^0(x,\theta),\theta)-v_\psi(\underline{x},\psi^0(x,\theta),\theta)}h(\psi^0(x,\theta)),$$

while for $x\notin(\underline{x},\bar{x})$, $x\notin[\hat{x},\hat{x}+\delta]$ and hence $\int\zeta(x,\theta)\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))dx=0$. But then

$$\begin{aligned}\int\zeta(x,\theta)\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))dx&=\int_{\underline{x}}^{\bar{x}}\frac{\mathcal{V}_x(x,\psi^0(x,\theta),\theta)}{v_\psi(\bar{x},\psi^0(x,\theta),\theta)-v_\psi(\underline{x},\psi^0(x,\theta),\theta)}h(\psi^0(x,\theta))dx\\&=\frac{\mathcal{V}(\bar{x},\psi^0(x,\theta),\theta)-\mathcal{V}(\underline{x},\psi^0(x,\theta),\theta)}{v_\psi(\bar{x},\psi^0(x,\theta),\theta)-v_\psi(\underline{x},\psi^0(x,\theta),\theta)}h(\psi^0(x,\theta)),\end{aligned}$$

where the second equality uses that $\psi^0(\cdot,\theta)$ is constant in x on $[\underline{x},\bar{x}]$.

Case 2: Assume instead that $\hat{x}$ is not the beginning of a jump.

$$\begin{aligned}&\int_{\hat{x}}^{x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)}\zeta(x,\theta)dx\max_{(x,\psi)\in[\hat{x},x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)]\times[\psi^0(\hat{x},\theta),\psi^0(\hat{x}+\delta,\theta)]}\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))\\&\geq\int_{\hat{x}}^{x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)}\zeta(x,\theta)\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))dx\\&\geq\int_{\hat{x}}^{x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)}\zeta(x,\theta)dx\min_{(x,\psi)\in[\hat{x},x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)]\times[\psi^0(\hat{x},\theta),\psi^0(\hat{x}+\delta,\theta)]}\mathcal{V}_x(x,\psi^0(x,\theta),\theta)h(\psi^0(x,\theta))\end{aligned}$$

But, since $\hat{x}$ is not the beginning of a jump as $\delta \downarrow 0$, $\psi^0(\hat{x}+\delta,\theta) \downarrow \psi^0(\hat{x},\theta)$ and $x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta) \downarrow \hat{x}$, and so the minimum and maximum converge to $\mathcal{V}_x(\hat{x},\psi^0(\hat{x},\theta),\theta)h(\psi^0(\hat{x},\theta))$

So, consider

$$\int_{\hat{x}}^{x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)} \zeta(x,\theta)dx.$$

Let $x \in [\hat{x}, x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)]$ be a point where $x_h^0(\psi(x,\theta),\theta) = x_l^0(\psi(x,\theta),\theta)$. Then

$$\zeta(x,\theta) = \frac{1}{\delta v_{x\psi}(x,\psi^0(x),\theta),\theta} \cong \frac{1}{\delta v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta),\theta)}.$$

Instead let $x < x_l^0(\psi^0(\hat{x}+\delta,\theta))$ be in a jump in x^0 from $\hat{x} \le \underline{x} \le \bar{x} \le \hat{x}+\delta$. Then

$$\zeta(x,\theta) = \frac{1}{\delta}\frac{I(\underline{x},\bar{x})}{v_\psi(\bar{x},\psi,\theta) - v_\psi(\underline{x},\psi,\theta)} = \frac{1}{\delta}\frac{\bar{x}-\underline{x}}{v_{x\psi}(x^*,\psi,\theta)(\bar{x}-\underline{x})} \cong \frac{1}{\delta v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta),\theta)}$$

and so

$$\int_{\underline{x}}^{\bar{x}} \zeta(x,\theta)dx \cong \int_{\underline{x}}^{\bar{x}} \frac{1}{\delta v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta))}dx$$

and hence,

$$\int_{x}^{x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)} \zeta(x,\theta)dx \cong \frac{x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)-\hat{x}}{\delta}\frac{1}{v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta),\theta)}$$

Finally, if $x \in (x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta), x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)$ then

$$\zeta(x,\theta) \cong \frac{\hat{x}+\delta-x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)}{\delta}\frac{1}{v_{x\psi}(x^*,\psi,\theta)(x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)-x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)}$$

and so

$$\int_{x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)}^{x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)} \zeta(x,\theta)dx \cong \frac{\hat{x}+\delta-x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)}{\delta}\frac{1}{v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta),\theta)}$$

and thus

$$\begin{aligned}\int_{\hat{x}}^{x_h^0(\psi^0(\hat{x}+\delta,\theta),\theta)} \zeta(x,\theta)dx &\cong \left(\frac{x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)-\hat{x}}{\delta} + \frac{\hat{x}+\delta-x_l^0(\psi^0(\hat{x}+\delta,\theta),\theta)}{\delta}\right)\frac{1}{v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta),\theta)} \\ &= \frac{1}{v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta),\theta)}\end{aligned}$$

where all of the approximations are arbitrarily good as $\delta \downarrow 0$, and so

$$\lim_{\delta\downarrow 0} L_\varepsilon(0,\theta) = \frac{\mathcal{V}_x(\hat{x},\psi^0(\hat{x},\theta),\theta)}{v_{x\psi}(\hat{x},\psi^0(\hat{x},\theta),\theta)}h(\psi^0(\hat{x},\theta))$$

which is again what the theorem claims.

B.3.10 The Necessary Conditions

Adding the two effects (revenue change and impact from switchers) and integrating over θ we obtain the necessary conditions for optimality. To see this, consider the function r defined in Section 6. Let $H \in \mathcal{H}$ and let (ρ, χ) be an optimal menu. If $x \in [0,1]$ is not a bunching point then,

$$\int_\Theta [(1 - w_C)(1 - H(\psi(x,\theta)|\theta)) - r(x,\theta)h(\psi(x,\theta)|\theta)]h(\theta)d\theta = 0,$$

while if x is a bunching point, then this optimality condition holds holds if one consistently uses either $\psi_l(x,\theta)$ or $\psi_h(x,\theta)$. □

B.3.11 Proof of Lemma 5

From Lemma 2 in Appendix A.4, the function V given by $V(\psi,\theta) \equiv \max_{x\in[0,1]}(v(x,\psi,\theta) - \rho(x))$ is continuous in (ψ,θ). Define $\tilde{\rho}$ by

$$\tilde{\rho}(x) \equiv \max_{(\psi,\theta)} v(x,\psi,\theta) - V(\psi,\theta). \tag{31}$$

We claim that $\tilde{\rho} \leq \rho$, with equality for any x that is an optimal choice for some type. To see that $\tilde{\rho} \leq \rho$, assume that for some x, $\rho(x) < \tilde{\rho}(x)$. Then, for some type and quality $\rho(x) < v(x,\psi,\theta) - V(\psi,\theta)$, or equivalently, $V(\psi,\theta) < v(x,\psi,\theta) - \rho(x)$, contradicting the definition of V. And, if $x = x^*(\psi,\theta)$ then $V(\psi,\theta) = v(x,\psi,\theta) - \rho(x)$, and so

$$\tilde{\rho}(x) \geq v(x,\psi,\theta) - (v(x,\psi,\theta) - \rho(x)) = \rho(x),$$

which implies that $\tilde{\rho}(x) = \rho(x)$. Note that by construction, facing $\tilde{\rho}$, each x is an optimal choice for at least one (ψ,θ). By Milgrom and Segal (2002), $\tilde{\rho}$ is absolutely continuous and differentiable almost everywhere, where at points of differentiability, $\tilde{\rho}$ has slope $\tilde{\rho}_x(x) = v_x(x,\psi,\theta) > 0$ for any (ψ,θ) for whom x is an optimal choice.[61] We have thus proven parts (i) and the absolute continuity claim in part (ii).

Let $T(x)$ be the set of (ψ,θ) that maximizes $v(x,\psi,\theta) - V(\psi,\theta)$. We will now show that $\tilde{\rho}$ is right-differentiable at any $x_0 < 1$ with a strictly positive right-derivative given by $\tilde{\rho}^+(x_0) = \max_{(\psi,\theta)\in T(x)} v_x(x_0,\psi,\theta)$, where x_0 is an optimal choice for type (ψ_0,θ_0). This follows as a

[61] To verify the premises of the cited theorem, note that the objective function is, for each (ψ,θ), absolutely continuous in x since it is $\mathcal{C}^1$, and v_x is bounded on $[0,1]\times\Psi\times\Theta$ by Assumption 6. Finally, the correspondence of maximizers is nonempty for each x. The conclusion of the theorem now follows and shows that $\tilde{\rho}$ is absolutely continuous and thus differentiable almost everywhere with derivative $\tilde{\rho}_x(x) = v_x(x,\psi,\theta)$.

direct application of part (ii) of Corollary 4 in Milgrom and Segal (2002). The same corollary shows that $\tilde{\rho}$ is left-differentiable at any $x_0 > 0$ with a strictly positive left-derivative $\tilde{\rho}^-(x_0) = \min_{(\psi,\theta)\in T(x)} v_x(x_0, \psi, \theta)$. And since v_x is strictly positive, $\tilde{\rho}$ is absolutely continuous (and thus continuous), and has strictly positive right- and left-derivatives, it follows that it is strictly increasing, completing the proof of part (ii).

To complete the proof of the lemma, it remains to prove that the right-derivative is continuous from the right and the left-derivative is continuous from the left. Once again, we prove the claim involving the right-derivative as the other one is analogous. Let $x_k \downarrow x_0$, and consider $\tilde{\rho}_x^+(x_k)$. We desire to show that $\tilde{\rho}_x^+(x_k) \to \tilde{\rho}_x^+(x_0)$. For each x_k construct a sequence $x_{kj} \downarrow x_k$ of points of differentiability of $\tilde{\rho}$, and an associated sequence of types (ψ_{kj}, θ_{kj}) for whom x_{kj} is optimal. Since $\tilde{\rho}$ is differentiable at x_{kj}, we have $v_x(x_{kj}, \psi_{kj}, \theta_{kj}) = \rho_x(x_{kj})$ and also that $v_x(x_{kj}, \psi_{kj}, \theta_{kj}) \to_j \tilde{\rho}_x^+(x_k)$. So, choosing for each k a j_k such that

$$\left|v_x(x_{kj_k}, \psi_{kj_k}, \theta_{kj_k}) - \tilde{\rho}_x^+(x_k)\right| \leq \frac{1}{2^k} \text{ and } x_{kj_k} - x_k \leq \frac{1}{2^k},$$

and taking a subsequence along which $(\psi_{kj_k}, \theta_{kj_k})$ converges to some (ψ_0, θ_0), note that

$$\begin{aligned}\left|\tilde{\rho}_x^+(x_k) - \tilde{\rho}_x^+(x_0)\right| &\leq \left|\tilde{\rho}_x^+(x_k) - v_x(x_{kj_k}, \psi_{kj_k}, \theta_{kj_k})\right| + \left|v_x(x_{kj_k}, \psi_{kj_k}, \theta_{kj_k}) - \tilde{\rho}_x^+(x_0)\right| \\ &\leq \frac{1}{2^k} + \left|v_x(x_{kj_k}, \psi_{kj_k}, \theta_{kj_k}) - \tilde{\rho}_x^+(x_0)\right|.\end{aligned}$$

But, considered as a sequence in k, we have $x_{kj_k} \downarrow x_0$, and x_{kj_k} is a point of differentiability of $\tilde{\rho}$, and hence by the existence of the right-derivative, $|v_x(x_{kj_k}, \psi_{kj_k}, \theta_{kj_k}) - \tilde{\rho}_x^+(x_0)| \to 0$, and so $|\tilde{\rho}_x^+(x_k) - \tilde{\rho}_x^+(x_0)| \to 0$ completing the proof of part (iii). □

B.4 *NC* in the Continuum: Proof of Theorem 5

Consider a setting where some $\hat{x}$ is chosen by a positive mass of types, which is to say that $\psi_l(\hat{x}, \theta) < \psi_h(\hat{x}, \theta)$ for a positive H-measure set of θ. A necessary condition for this is that $\rho_x^-(\hat{x}) < \rho_x^+(\hat{x})$. We will first analyze a perturbation that causes types who chose $\hat{x}$ to choose a somewhat lower x instead, and then one that cause agents to replace $\hat{x}$ by a slightly higher value. Between them, we will derive the stated necessay condition. Along the way, we will see that $\hat{x}$ being *POCA* is not only a useful simplification in proving our result, it in fact essentially necessary. A key component of the proofs is that the perturbations that shift the consumers choosing $\hat{x}$ up or down necessarily have other effects as well. For example, one cannot cause the agents choosing $\hat{x}$ to move to a higher value $\tilde{x}(\varepsilon)$ without some combination of other types switching to $\tilde{x}(\varepsilon)$ and revenue implication for higher types. The key will be to use Theorem 4 along with *POCA* to cancel these effects. Throughout the proof we are going to assume that ρ

has the properties derived in Section B.3.1.

B.4.1 Lowering $\hat{x}$

Fix $\hat{x}$ and θ, and suppress them in the notation. Let $s > \max_{(x,\psi,\theta)} v_x(x,\psi,\theta)$, and let $\tilde{x}(\varepsilon) < \hat{x}$ be defined by

$$\rho(\tilde{x}(\varepsilon)) + s(\hat{x} - \tilde{x}(\varepsilon)) = \rho(\hat{x}) + \varepsilon.$$

Because ρ is absolutely continuous, so is $\tilde{x}$, and so on the measure one subset of ε were $\tilde{x}(\varepsilon)$ is not a kink point of ρ,

$$\infty < \tilde{x}_\varepsilon(\varepsilon) = -\frac{1}{s - \rho_x(\tilde{x}(\varepsilon))} < 0.$$

which has limit as $\varepsilon \downarrow 0$ equal to

$$\tilde{x}_\varepsilon(0) = -\frac{1}{s - \rho_x^-(\hat{x})} < 0$$

where we use $\rho_\varepsilon^-(\hat{x})$ since $\tilde{x}(\varepsilon) \uparrow \hat{x}$ and where $s - \rho_x^-(\hat{x}) > 0$ since $\rho_x^-(\hat{x}) \leq \max v_x$.

Let

$$\rho^\varepsilon(x) = \rho(x) \text{ for } x \leq \tilde{x}(\varepsilon),\ \rho(\tilde{x}(\varepsilon)) + s(x - \tilde{x}(\varepsilon)) \text{ for } x \in [\tilde{x}(\varepsilon), \hat{x}] \text{ and } \rho(x) + \varepsilon \text{ for } x \geq \hat{x}.$$

That is to form ρ^ε, ρ is untouched until shortly before $\hat{x}$, then rises at slope s to hit $\rho(\hat{x}) + \varepsilon$ at $\hat{x}$, and then runs parallel to ρ but ε higher to the right of $\hat{x}$. Facing ρ^ε, all types strictly prefer $\tilde{x}(\varepsilon)$ to anything in $[\tilde{x}(\varepsilon), \hat{x}]$, since $s > \max v_x$. Hence, if $\tilde{\psi}(\varepsilon) > \psi_h$ is defined by

$$V(\tilde{\psi}(\varepsilon)) = \varepsilon + v(\tilde{x}(\varepsilon), \tilde{\psi}(\varepsilon)) - \rho(\tilde{x}(\varepsilon)).$$

then all types in $(\psi^0(\tilde{x}(\varepsilon)), \tilde{\psi}(\varepsilon))$ switch their choice down to $\tilde{x}(\varepsilon)$. Other types do not change their choice, but those above $\tilde{\psi}(\varepsilon)$ pay ε more than they used to. Since $V_\psi(\psi) = v_\psi(x^0(\psi), \psi)$, we have

$$\tilde{\psi}_\varepsilon(\varepsilon) = \frac{1 + \Big(v_x(\tilde{x}(\varepsilon), \tilde{\psi}(\varepsilon)) - \rho_x(\tilde{x}(\varepsilon))\Big)\tilde{x}_\varepsilon(\varepsilon)}{v_\psi(x^0(\tilde{\psi}(\varepsilon)), \tilde{\psi}(\varepsilon)) - v_\psi(\tilde{x}(\varepsilon), \tilde{\psi}(\varepsilon))}.$$

The principal's profit at ρ^ε minus that at ρ^0 is thus

$$(1 - w_C)(1 - H(\tilde{\psi}(\varepsilon))\varepsilon - \int_{\psi^0(\tilde{x}(\varepsilon))}^{\tilde{\psi}(\varepsilon)} \left(\begin{array}{c} \mathcal{V}(x^0(\psi), \psi) - (1 - w_C)\left(v(x^0(\psi), \psi) - \rho(x^0(\psi))\right) \\ -\left(\mathcal{V}(\tilde{x}(\varepsilon), \psi) - (1 - w_C)\left(v(\tilde{x}(\varepsilon), \psi) - \rho(\tilde{x}(\varepsilon))\right)\right) \end{array} \right) h(\psi) d\psi.$$

Now, break up the integral, and subtract $(1 - w_C)\Big(H(\tilde{\psi}(\varepsilon)) - H(\psi_h)\Big)\varepsilon$ from the first integral

and adjust the first term to compensate and arrive at

$$
\begin{aligned}
(32)\qquad & (1-w_C)(1-H(\psi_h))\varepsilon \\
& - \int_{\psi_h}^{\tilde{\psi}(\varepsilon)} \left(\begin{array}{c} \mathcal{V}(x^0(\psi),\psi) - (1-w_C)\left(v(x^0(\psi),\psi) - \rho(x^0(\psi)) - \varepsilon\right) \\ -\left(\mathcal{V}(\tilde{x}(\varepsilon),\psi) - (1-w_C)\left(v(\tilde{x}(\varepsilon),\psi) - \rho(\tilde{x}(\varepsilon))\right)\right) \end{array} \right) h(\psi)d\psi \\
& - \int_{\psi^0(\tilde{x}(\varepsilon))}^{\psi_h} \left(\begin{array}{c} \mathcal{V}(x^0(\psi),\psi) - (1-w_C)\left(v(x^0(\psi),\psi) - \rho(x^0(\psi))\right) \\ -\left(\mathcal{V}(\tilde{x}(\varepsilon),\psi) - (1-w_C)\left(v(\tilde{x}(\varepsilon),\psi) - \rho(\tilde{x}(\varepsilon))\right)\right) \end{array} \right) h(\psi)d\psi,
\end{aligned}
$$

To differentiate this, we begin witht top Leibniz term of the first integral of (32), which is

$$
\tilde{\psi}_\varepsilon(\varepsilon) \left(\begin{array}{c} \mathcal{V}(x^0(\psi),\psi) - (1-w_C)\left(v(x^0(\psi),\psi) - \rho(x^0(\psi) - \varepsilon)\right) \\ -\left(\mathcal{V}(\tilde{x}(\varepsilon),\psi) - (1-w_C)\left(v(\tilde{x}(\varepsilon),\psi) - \rho(\tilde{x}(\varepsilon))\right)\right) \end{array} \right) h(\psi)
$$

evaluated at $\psi = \tilde{\psi}(\varepsilon)$. But, by definition of $\tilde{\psi}$, at $\psi = \tilde{\psi}(\varepsilon)$

$$
v(x^0(\psi),\psi) - \rho(x^0(\psi)) - \varepsilon = (v(\tilde{x}(\varepsilon),\psi) - \rho(\tilde{x}(\varepsilon)))
$$

and so the large bracketed expression reduces to $\mathcal{V}(x^0(\tilde{\psi}(\varepsilon)),\tilde{\psi}(\varepsilon)) - \mathcal{V}(\tilde{x}(\varepsilon),\tilde{\psi}(\varepsilon))$ and thus, using the expression for $\tilde{\psi}_\varepsilon$, we arrive at

$$
\frac{\mathcal{V}(x^0(\tilde{\psi}(\varepsilon)),\tilde{\psi}(\varepsilon)) - \mathcal{V}(\tilde{x}(\varepsilon),\tilde{\psi}(\varepsilon))}{v_\psi(x^0(\tilde{\psi}(\varepsilon)),\tilde{\psi}(\varepsilon)) - v_\psi(\tilde{x}(\varepsilon),\tilde{\psi}(\varepsilon))} \left(1 + \left(v_x(\tilde{x}(\varepsilon),\tilde{\psi}(\varepsilon)) - \rho_x(\tilde{x}(\varepsilon))\right)\tilde{x}_\varepsilon(\varepsilon)\right) h(\tilde{\psi}(\varepsilon)).
$$

Taking a limit as $\varepsilon \downarrow 0$ and using that $\tilde{x}(\varepsilon) \uparrow \hat{x}$, $x^0(\tilde{\psi}(\varepsilon)) \downarrow x_h^0(\psi_h)$, and $\tilde{\psi}(\varepsilon) \downarrow \psi_h$, we have that whether $x_h^0(\psi_h) > \hat{x}$ or $x_h^0(\psi_h) = \hat{x}$, the fraction converges to $r(\hat{x},\theta)$ evaluated at ψ_h, and so the limit of the top Leibniz term is

$$
r\left(1 + \left(v_x(\hat{x},\psi_h) - \rho_x^-(\hat{x})\right)\tilde{x}_\varepsilon(\varepsilon)\right) h(\psi_h).
$$

The bottom Leibniz term in the second integral in (32) is zero, since $x^0(\psi) = x^0(\psi^0(\tilde{x}(\varepsilon))) = \tilde{x}(\varepsilon)$, and so the integrand is identically zero at $\psi^0(\tilde{x}(\varepsilon))$.

Now, the derivative of the integrand is

$$
-\left(\mathcal{V}_x(\tilde{x}(\varepsilon),\psi) - (1-w_C)\left(v_x(\tilde{x}(\varepsilon),\psi) - \rho_x(\tilde{x}(\varepsilon))\right)\right)\tilde{x}_\varepsilon(\varepsilon)
$$

which is finite and measurable, and so we can interchange integration and differentiation by Lebesgue's dominated convergence theorem. Hence, recombining the integrals, the part of the

derivative that reflects the integrand is

$$-\tilde{x}_\varepsilon(\varepsilon)\int_{\psi^0(\tilde{x}(\varepsilon))}^{\tilde{\psi}(\varepsilon)}\left(\mathcal{V}_x(\tilde{x}(\varepsilon),\psi)-(1-w_C)\left(v_x(\tilde{x}(\varepsilon),\psi)-\rho_x(\tilde{x}(\varepsilon))\right)\right)h(\psi)d\psi,$$

where again using that the integrand is bounded, and since $\tilde{\psi}(\varepsilon)\downarrow\psi_h$ and $\psi^0(\tilde{x}(\varepsilon))\uparrow\psi_l$, the limit is

$$-\tilde{x}_\varepsilon(0)\int_{\psi_l}^{\psi_h}\left(\mathcal{V}_x(\hat{x},\psi)-(1-w_C)\left(v_x(\hat{x},\psi)-\rho_x^-(\hat{x})\right)\right)h(\psi)d\psi,$$

where we use $\rho_x^-(\hat{x})$ since $\tilde{x}(\varepsilon)\uparrow\hat{x}$.

Thus, the derivative of the principal's payoff at $\varepsilon=0$ is then (recalling that the integral appeared with a minus sign)

$$\begin{aligned}&(1-w_C)\left(1-H(\psi_h)\right)-r\left(1+\left(v_x(\hat{x},\psi_h)-\rho_x^-(\hat{x})\right)\tilde{x}_\varepsilon(\varepsilon)\right)h(\psi_h)\\&+\tilde{x}_\varepsilon(0)\int_{\psi_l}^{\psi_h}\left(\mathcal{V}_x(\hat{x},\psi)-(1-w_C)\left(v_x(\hat{x},\psi)-\rho_x^-(\hat{x})\right)\right)h(\psi)d\psi.\end{aligned}$$

or, since $\tilde{x}_\varepsilon(0)=\frac{-1}{s-\rho_x^-(\hat{x})}$, we can substitute and then multiply by $s-\rho_x^-(\hat{x})>0$ and break up the integral to arrrive at

$$\begin{aligned}&\left(s-\rho_x^-(\hat{x})\right)(1-w_C)\left(1-H(\psi_h)\right)-r\left(s-\rho_x^-(\hat{x})-\left(v_x(\hat{x},\psi_h)-\rho_x^-(\hat{x})\right)\right)h(\psi_h))\\&-\int_{\psi_l}^{\psi_h}\mathcal{V}_x(\hat{x},\psi)h(\psi)d\psi+(1-w_C)\int_{\psi_l}^{\psi_h}\left(v_x(\hat{x},\psi)-\rho_x^-(\hat{x})\right)h(\psi)d\psi.\end{aligned}$$

Now, integrate by parts to see that

$$\begin{aligned}&\int_{\psi_l}^{\psi_h}\left(v_x(\hat{x},\psi)-\rho_x^-(\hat{x})\right)h(\psi)d\psi\\=\;&\left(1-H(\psi_l)\right)\left(v_x(\hat{x},\psi_l)-\rho_x^-(\hat{x})\right)-\left(1-H(\psi_h)\right)\left(v_x(\hat{x},\psi_h)-\rho_x^-(\hat{x})\right)\\&+\int_{\psi_l}^{\psi_h}v_{x\psi}(\hat{x},\psi)(1-H(\psi))d\psi\end{aligned}$$

and substitute and collect terms to get

$$\begin{aligned}&\left(s-\rho_x^-(\hat{x})\right)(1-w_C)\left(1-H(\psi_h)\right)-r\left(s-\rho_x^-(\hat{x})-\left(v_x(\hat{x},\psi_h)-\rho_x^-(\hat{x})\right)\right)h(\psi_h))\\&+(1-w_C)(1-H(\psi_l))\left(v_x(\hat{x},\psi_l)-\rho_x^-(\hat{x})\right)-(1-w_C)(1-H(\psi_h))\left(v_x(\hat{x},\psi_h)-\rho_x^-(\hat{x})\right)\\&-\int_{\psi_l}^{\psi_h}\left(\mathcal{V}_x(\hat{x},\psi)-(1-w_C)v_{x\psi}(\hat{x},\psi)\frac{1-H(\psi)}{h(\psi)}\right)h(\psi)d\psi.\end{aligned}$$

We need to understand the integral across θ of the top two lines. By Theorem xx $rh(\psi_h)$ and $(1-w_C)(1-H(\psi_h))$ integrate across θ to the same thing and so, cancelling appropriately, we arrive at

$$(1-w_C)(1-H(\psi_l))\left(v_x(\hat{x},\psi_l)-\rho_x^-(\hat{x})\right)$$

so since the exercise here was to lower x, the general necessary condition is that

$$\begin{aligned}&(1-w_C)\int\left((1-H(\psi_l(\theta)))\left(v_x(\hat{x}(\theta),\psi_l(\theta))-\rho_x^-(\hat{x}(\theta))\right)\right)h(\theta)d\theta\\&-\int\left(\int_{\psi_l(\theta)}^{\psi_h(\theta)}\left(\mathcal{V}_x(\hat{x},\psi,\theta)-(1-w_C)v_{x\psi}(\hat{x},\psi,\theta)\frac{1-H(\psi|\theta)}{h(\psi|\theta)}\right)h(\psi|\theta)d\psi\right)h(\theta)d\theta\geq 0\end{aligned}$$

and where under *POCA* $v_x(\hat{x}(\theta),\psi_l(\theta))-\rho_x^-(\hat{x}(\theta))\equiv 0$ and so the first term dissappears and we have

$$\int\left(\int_{\psi_l(\theta)}^{\psi_h(\theta)}\left(\mathcal{V}_x(\hat{x},\psi,\theta)-(1-w_C)v_{x\psi}(\hat{x},\psi,\theta)\frac{1-H(\psi|\theta)}{h(\psi|\theta)}\right)h(\psi|\theta)d\psi\right)h(\theta)d\theta\geq 0. \tag{33}$$

One thing to note is that by incentive compatibility, $v_x(\hat{x}(\theta),\psi_l(\theta))-\rho_x^-(\hat{x}(\theta))$ is always weakly positive, otherwise lowering the choice from $\hat{x}$ strictly benefits the agent. So, the first term is always weakly positive, and disappears only if $v_x(\hat{x}(\theta),\psi_l(\theta))=\rho_x^-(\hat{x}(\theta))$ for all but a zero measure set of θ.

B.4.2 Raising $\hat{x}$

The idea here is similar to before but enough details are different to merit being explicit. We repeat the structure of the previous proof very closely, but now, instead of making ρ very steep to the left of $\hat{x}$, and hence pushing agents from $\hat{x}$ to a lower choice, we flatten ρ for an interval to the right of $\hat{x}$, pushing agents from $\hat{x}$ to a higher choice.

Let $\tilde{x}(\varepsilon)>\hat{x}$ be defined by

$$\rho(\hat{x})=\rho(\tilde{x}(\varepsilon))-\varepsilon,$$

recalling that ρ is absolutely continuous with slope bounded away from zero, and so $\tilde{x}$ is absolutely continuous and strictly increasing, and at the measure one subset of ε were $\tilde{x}(\varepsilon)$ is not a kink point of ρ,

$$0<\tilde{x}_\varepsilon(\varepsilon)=\frac{1}{\rho_x(\tilde{x}(\varepsilon))}<\infty.$$

which has limit as $\varepsilon\downarrow 0$ equal to

$$\tilde{x}_\varepsilon(0)=\frac{1}{\rho_x^+(\hat{x})}>0$$

where we use $\rho_\varepsilon^+(\hat{x})$ since $\tilde{x}(\varepsilon) \downarrow \hat{x}$.

Let

$$\rho^\varepsilon(x) = \rho(x) \text{ for } x \leq \tilde{x}(\varepsilon),\ \rho(\hat{x}) \text{ for } x \in [\hat{x}, \tilde{x}(\varepsilon)] \text{ and } \rho(x) - \varepsilon \text{ for } x \geq \tilde{x}(\varepsilon).$$

That is, ρ is untouched until $\hat{x}$, flat from $\hat{x}$ to $\tilde{x}(\varepsilon)$ and then is parallel to ρ but ε lower beyond that. Facing ρ^ε, all types strictly prefer $\tilde{x}(\varepsilon)$ to anything in $[\hat{x}, \tilde{x}(\varepsilon)]$, since $v_x > 0$. Hence, if $\tilde{\psi}(\varepsilon) < \psi_l$ is defined by

$$V(\tilde{\psi}(\varepsilon)) = v(\tilde{x}(\varepsilon), \tilde{\psi}(\varepsilon)) - \rho(\hat{x}).$$

then all types in $(\tilde{\psi}(\varepsilon), \psi^0(\tilde{x}(\varepsilon)))$ switch their choice up to $\tilde{x}(\varepsilon)$. Other types do not change their choice, but those above $\psi^0(\tilde{x}(\varepsilon))$ pay ε less. Similar to before,

$$\tilde{\psi}_\varepsilon(\varepsilon) = -\frac{v_x(\tilde{x}(\varepsilon), \tilde{\psi}(\varepsilon))\tilde{x}_\varepsilon(\varepsilon)}{v_\psi(\tilde{x}(\varepsilon), \tilde{\psi}(\varepsilon)) - v_\psi(x^0(\tilde{\psi}(\varepsilon)), \tilde{\psi}(\varepsilon))}.$$

The prinicipal's payoff at ε versus 0 is thus

$$\begin{array}{l} -(1-w_C)(1-H(\psi^0(\tilde{x}(\varepsilon))))\varepsilon \\ +\displaystyle\int_{\tilde{\psi}(\varepsilon)}^{\psi^0(\tilde{x}(\varepsilon))} \left(\begin{array}{c} \mathcal{V}(\tilde{x}(\varepsilon), \psi) - (1-w_C)\left(v(\tilde{x}(\varepsilon), \psi) - \rho(\hat{x})\right) \\ -\left(\mathcal{V}(x^0(\psi), \psi) - (1-w_C)\left(v(x^0(\psi), \psi) - \rho(x^0(\psi))\right)\right) \end{array} \right) h(\psi)d\psi. \end{array}$$

Split the integral at ψ_h and add $(1-w_C)\left(H(\psi^0(\tilde{x}(\varepsilon)) - H(\psi_h)\right)\varepsilon$ to the first integral above ψ_h while adjusting the first term to compensate and then use that $\rho(\hat{x}) + \varepsilon = \rho\left(\tilde{x}(\varepsilon)\right)$ and arrive at

$$\begin{array}{l} -(1-w_C)(1-H(\psi_h))\varepsilon \\ \displaystyle\int_{\psi_h}^{\psi^0(\tilde{x}(\varepsilon))} \left(\begin{array}{c} \mathcal{V}(\tilde{x}(\varepsilon), \psi) - (1-w_C)\left(v(\tilde{x}(\varepsilon), \psi) - \rho\left(\tilde{x}(\varepsilon)\right)\right) \\ -\left(\mathcal{V}(x^0(\psi), \psi) - (1-w_C)\left(v(x^0(\psi), \psi) - \rho(x^0(\psi))\right)\right) \end{array} \right) h(\psi)d\psi \\ +\dfrac{\partial}{\partial\varepsilon}\displaystyle\int_{\tilde{\psi}(\varepsilon)}^{\psi_h} \left(\begin{array}{c} \mathcal{V}(\tilde{x}(\varepsilon), \psi) - (1-w_C)\left(v(\tilde{x}(\varepsilon), \psi) - \rho(\hat{x})\right) \\ -\left(\mathcal{V}(x^0(\psi), \psi) - (1-w_C)\left(v(x^0(\psi), \psi) - \rho(x^0(\psi))\right)\right) \end{array} \right) h(\psi)d\psi. \end{array}$$

The top Leibnitz term in the first integral disappears, since the large bracketed term is zero since $x^0(\psi^0(\tilde{x}(\varepsilon))) = \tilde{x}(\varepsilon)$. The derivative of the integrand in the top integral is bounded, and $\psi^0(\tilde{x}(\varepsilon)) \downarrow \psi_h$, so that the top integral vanishes as $\varepsilon \downarrow 0$. Turning to the bottom integral, to evaluate the bottom Leibniz term, note that by definition of $\tilde{\psi}$,

$$v(\tilde{x}(\varepsilon), \tilde{\psi}(\varepsilon)) - \rho(\hat{x}) = v(x^0(\tilde{\psi}(\varepsilon)), \tilde{\psi}(\varepsilon)) - \rho(x^0(\tilde{\psi}(\varepsilon))),$$

and so, using the expression for $\tilde{\psi}_\varepsilon$,

$$\frac{\mathcal{V}(\tilde{x}(\varepsilon),\tilde{\psi}(\varepsilon))-\mathcal{V}(x^0(\tilde{\psi}(\varepsilon)),\tilde{\psi}(\varepsilon))}{v(\tilde{x}(\varepsilon),\tilde{\psi}(\varepsilon))-v_\psi(x^0(\psi),\psi)}v_x(\tilde{x}(\varepsilon),\tilde{\psi}(\varepsilon))\tilde{x}_\varepsilon(\varepsilon)h(\tilde{\psi}(\varepsilon)).$$

Taking $\varepsilon \downarrow 0$ and using that $\tilde{x}(\varepsilon) \downarrow \hat{x}$ and $x^0(\tilde{\psi}(\varepsilon)) \uparrow x_l^0(\psi_l)$, the fraction converges to $r(\hat{x},\theta)$ evaluated at ψ_l, and we arrive at

$$rv_x(\hat{x},\psi_l)\tilde{x}_\varepsilon(0)h(\psi_l).$$

The derivative of the integrand is

$$\left(\mathcal{V}_x(\tilde{x}(\varepsilon),\psi)-(1-w_C)v_x(\tilde{x}(\varepsilon),\psi)\right)\tilde{x}_\varepsilon(\varepsilon)$$

which is finite. Hence, the part of the derivative that reflects the integrand is

$$\tilde{x}_\varepsilon(\varepsilon)\int_{\tilde{\psi}(\varepsilon)}^{\psi^0(\tilde{x}(\varepsilon))}\left(\mathcal{V}_x(\tilde{x}(\varepsilon),\psi)-(1-w_C)v_x(\tilde{x}(\varepsilon),\psi)\right)h(\psi)d\psi,$$

which, using that $\tilde{\psi}(\varepsilon) \uparrow \psi_l$, $\tilde{x}(\varepsilon) \downarrow \hat{x}$ and $\psi^0(\tilde{x}(\varepsilon)) \downarrow \psi_h$ converges to

$$\tilde{x}_\varepsilon(0)\int_{\psi_l}^{\psi_h}\left(\mathcal{V}_x(\hat{x},\psi)-(1-w_C)v_x(\hat{x},\psi)\right)h(\psi)d\psi.$$

Thus, the derivative of the principal's payoff at $\varepsilon = 0$ is

$$-(1-w_C)\left(1-H(\psi_h)\right)+rv_x(\hat{x},\psi_l)\tilde{x}_\varepsilon(0)h(\psi_l)+\tilde{x}_\varepsilon(0)\int_{\psi_l}^{\psi_h}\left(\mathcal{V}_x(\hat{x},\psi)-(1-w_C)v_x(\hat{x},\psi)\right)h(\psi)d\psi,$$

where, since $\tilde{x}_\varepsilon(0)=\frac{1}{\rho_x^+(\hat{x})}>0$, we can multiply by $\rho_x^+(\hat{x})>0$ to arrrive at

$$-(1-w_C)\left(1-H(\psi_h)\right)\rho_x^+(\hat{x})+rv_x(\hat{x},\psi_l)h(\psi_l)+\int_{\psi_l}^{\psi_h}\mathcal{V}_x(\hat{x},\psi)h(\psi)d\psi-(1-w_C)\int_{\psi_l}^{\psi_h}v_x(\hat{x},\psi)h(\psi)d\psi,$$

Now,

$$\int_{\psi_l}^{\psi_h}v_x(\hat{x},\psi)h(\psi)d\psi=(1-H(\psi_l))v_x(\hat{x},\psi_l)-(1-H(\psi_h))v_x(\hat{x},\psi_h)+\int_{\psi_l}^{\psi_h}v_{x\psi}(\hat{x},\psi)(1-H(\psi))d\psi$$

and so we can substitute and collect terms to get

$$\begin{aligned}&-(1-w_C)\left((1-H(\psi_h))\rho_x^+(\hat{x})+rv_x(\hat{x},\psi_l)h(\psi_l)+(1-H(\psi_l))v_x(\hat{x},\psi_l)-(1-H(\psi_h))v_x(\hat{x},\psi_h)\right)\\&+\int_{\psi_l}^{\psi_h}\left(\mathcal{V}_x(\hat{x},\psi)-(1-w_C)v_{x\psi}(\hat{x},\psi)\frac{1-H(\psi)}{h(\psi)}\right)h(\psi)d\psi.\end{aligned}$$

Here, since we evaluated r at ψ_l, by Theorem 4 we can cancel $rh(\psi_l)$ with $-(1-w_C)(1-H(\psi_l))$ to arrive at

$$-(1-w_C)\left((1-H(\psi_h))\,\rho_x^+(\hat{x}) + -(1-H(\psi_h))v_x(\hat{x},\psi_h)\right)$$
$$-(1-w_C)\left((1-H(\psi_h)\left(\rho_x^+(\hat{x}) - v_x(\hat{x},\psi_h)\right)\right),$$

and so since we are raising x, the necessary condition is

$$-(1-w_C)\int\left((1-H(\psi_h(\theta)))\left(\rho_x^+(\hat{x}(\theta) - v_x(\hat{x}(\theta),\psi_h(\theta)))\right)\right)h(\theta)d\theta$$
$$+\int\left(\int_{\psi_l(\theta)}^{\psi_h(\theta)}\left(\mathcal{V}_x(\hat{x},\psi,\theta) - (1-w_C)v_{x\psi}(\hat{x},\psi,\theta)\frac{1-H(\psi|\theta)}{h(\psi|\theta)}\right)h(\psi|\theta)d\psi\right)h(\theta)d\theta \leq 0.$$

The first term is weakly negative and disappears if and only if $v_x(\hat{x}(\theta),\psi_l(\theta)) = \rho_x^-(\hat{x}(\theta))$ is 0 with H-measure one, and thus the condition is that

$$\int\left(\int_{\psi_l(\theta)}^{\psi_h(\theta)}\left(\mathcal{V}_x(\hat{x},\psi,\theta) - (1-w_C)v_{x\psi}(\hat{x},\psi,\theta)\frac{1-H(\psi|\theta)}{h(\psi|\theta)}\right)h(\psi|\theta)d\psi\right)h(\theta)d\theta \leq 0.$$

which together with (33), gives that for any $\hat{x}$ that is $POCA$,

$$\int\left(\int_{\psi_l(\theta)}^{\psi_h(\theta)}\left(\mathcal{V}_x(\hat{x},\psi,\theta) - (1-w_C)v_{x\psi}(\hat{x},\psi,\theta)\frac{1-H(\psi|\theta)}{h(\psi|\theta)}\right)h(\psi|\theta)d\psi\right)h(\theta)d\theta = 0$$

and we are done. □

B.5 Details for Section A.3

Proof of Lemma 1 Let us first show the continuity of Π on $\mathcal{H}_A$. Given our definition of ρ_j, note that it is equal to $\rho_j = \sum_{j'=1}^{j} p_j$. For given ρ, let

$$B_j(\rho) = \{(\psi,\theta)|j \in \arg\max_{j'}(v(x_{j'},\psi,\theta) - \rho_{j'}\}$$

be the set of types for whom buying quality x_j is optimal. For any given $j' < j''$, for $B_{j'}(p)$ and $B_{j''}(p)$ to overlap (so that the consumer has more than one best response), we must have $v(x_{j''},\psi,\theta) - v(x_{j'},\psi,\theta) = \rho_{j''} - \rho_{j'}$, which is a zero probability event for any H in $\mathcal{H}_A$. Hence, we can write $\Pi(\rho,H) = \sum_{j=1}^{J}\pi_j(\rho,H)$, where

$$\pi_j(\rho,H) \equiv \int_{B_j(p)} S(\rho_j,x_j,\psi,\theta)dH(\psi,\theta)$$

is the profit on those types for whom j is an optimal choice. It is thus enough to show that π_j is continuous for each j.

Let (p^k, H^k) be a sequence in $[0,\bar{p}]^J \times \mathcal{H}_A$ that converges to $(p^0, H^0) \in [0,\bar{p}]^J \times \mathcal{H}_A$ with associated ρ^k and ρ^0. For any $\delta > 0$, let

$$E_\delta = \{(\psi,\theta) | v(x_{j''},\psi,\theta) - v(x_{j'},\psi,\theta) \in [\rho^0_{j''} - \rho^0_{j'} - \delta, \rho^0_{j''} - \rho^0_{j'} + \delta] \text{ for some } j' \neq j''\}.$$

That is, E_δ is the event that for some j' and j'', (ψ,θ) has incremental value for the move from $x_{j'}$ to $x_{j''}$ that is within δ of the incremental price under ρ^0. Now, for given k, let $\delta^k = \max_{(j',j'')} |\rho^k_{j''} - \rho^k_{j'} - (\rho^0_{j''} - \rho^0_{j'})|$. We claim that for all j, $B_j(\rho^k) \backslash B_j(\rho^0) \subseteq E_{\delta^k}$. To see this, note that for (ψ,θ) to be in $B_j(\rho^k)\backslash B_j(\rho^0)$ it must be that for some $j' \neq j$,

$$v(x_j,\psi,\theta) - \rho^0_j < v(x_{j'},\psi,\theta) - \rho^0_{j'}$$

because x_j is not an optimal choice facing ρ^0 but

$$v(x_j,\psi,\theta) - \rho^k_j \geq v(x_{j'},\psi,\theta) - \rho^k_{j'}$$

since x_j is an optimal choice facing ρ^k and so, rearranging,

$$\rho^k_j - \rho^k_{j'} \leq v(x_j,\psi,\theta) - v(x_{j'},\psi,\theta) < \rho^0_j - \rho^0_{j'}$$

and thus since by definition, $\rho^k_j - \rho^k_{j'}$ is within δ^k of $\rho^0_j - \rho^0_{j'}$, $(\psi,\theta) \in E_{\delta^k}$. Similarly, $B_j(\rho^0)\backslash B_j(\rho^k) \subseteq E_{\delta^k}$.

Let $s \equiv \max_{(j,\psi,\theta)} |S(\rho^0_j, x_j, \psi, \theta)|$ be the maximum gain or loss on any consumer facing ρ^0. This is bounded since γ is continuous on a compact domain. Now, for all k,

$$\begin{aligned}
|\pi_j(\rho^k, H^k) - \pi_j(\rho^0, H^0)| &= \left| \int_{B_j(\rho^k)} S(\rho^k_j, x_j, \psi, \theta) dH^k(\psi,\theta) - \int_{B_j(\rho^0)} S(\rho^0_j, x_j, \psi, \theta) dH^0(\psi,\theta) \right| \\
&\leq \left| \int_{B_j(\rho^0)} S(\rho^k_j, x_j, \psi, \theta) dH^k(\psi,\theta) - \int_{B_j(\rho^0)} S(\rho^0_j, x_j, \psi, \theta) dH^0(\psi,\theta) \right| \\
&\quad + \int_{B_j(\rho^k)\backslash B_j(\rho^0) \cup B_j(\rho^0)\backslash B_j(\rho^k)} |S(\rho^k_j, x_j, \psi, \theta)| dH^k(\psi,\theta)
\end{aligned}$$

Fix any $\delta > 0$. Then, for k sufficiently large that $\delta^k < \delta$ and such that $H^k(E_\delta) \leq 2H^0(E_\delta)$,

$$\int_{B_j(\rho^k)\backslash B_j(\rho^0) \cup B_j(\rho^0)\backslash B_j(\rho^k)} |\rho^k_j - \gamma(x_j,\psi,\theta)| dH^k(\psi,\theta) \leq (s + |\rho^k_j - \rho^0_j|) 2H^0(E_\delta).$$

But, since H^0 is atomless, $H^0(E_\delta)$ can be made as small as desired by choosing δ sufficiently small and so these terms can be ignored. Finally,

$$\left|\int_{B_j(\rho^0)} S(\rho_j^k, x_j, \psi, \theta) dH^k(\psi,\theta) - \int_{B_j(\rho^0)} S(\rho_j^0, x_j, \psi, \theta) dH^0(\psi,\theta)\right|$$
$$\leq |1 - w_C||\rho_j^k - \rho_j^0| + \left|\int_{B_j(\rho^0)} S(\rho_j^0, x_j, \psi, \theta) dH^k(\psi,\theta) - \int_{B_j(\rho^0)} S(\rho_j^0, x_j, \psi, \theta) dH^0(\psi,\theta)\right|$$

where on the right-hand side the difference in integrals goes to zero in k since $S(\rho_j^0, x_j, \psi, \theta)$ is a continuous function and $H^k \to H^0$, and the first term vanishes in k. We have thus established that $|\pi_j(\rho^k, H^k) - \pi_j(\rho^0, H^0)| \to 0$, and so Π is continuous on $[0, \bar{p}]^J \times \mathcal{H}_A$.

To establish the continuity of $\tilde{\Pi}$, define $\tilde{B}_j(p_j) = \{(\psi,\theta)|v_j(\psi,\theta) > p_j\}$ as the set of types for whom buying quality increment j is optimal. Then, $\tilde{\Pi}(p, H) = \sum_{j=1}^J \tilde{\pi}_j(p, H)$, where $\tilde{\pi}_j(p, H) \equiv \int_{\tilde{B}_j(p)} (\mathcal{V}^j(\psi,\theta) - (1 - w_C)(v_j(\psi,\theta) - p_j)) dH(\psi,\theta)$. The proof then proceeds as before with minor modifications in establishing that $\tilde{B}_j(p^k) \backslash \tilde{B}_j(p^0)$ and $\tilde{B}_j(p^0) \backslash \tilde{B}_j(p^k)$ are subsets of E_{δ^k}. □

Lemma 7 *Assume (as in Definition 1) that $\hat{v}_{x\tau} > 0$ and $\hat{\mathcal{V}}_x(\cdot,\tau)/\hat{v}_{x\tau}(\cdot,\tau)$ is strictly increasing. Also assume that $1 - G$ is strictly log-concave and (1) $\hat{\mathcal{V}}_x/\hat{v}_{x\tau}$ increasing in τ, (2) $\hat{v}_{x\tau\tau} > 0$, (3) $\hat{\mathcal{V}}_{x\tau} \geq 0$, and (4) $\hat{v}_{x\tau}$ is log submodular. Then, $\frac{\hat{\mathcal{V}}^j}{\hat{v}_\tau^j} \frac{g}{1-G}$ is strictly increasing in τ.*

Proof Given that $1 - G$ is strictly log-concave, it is enough that $\hat{\mathcal{V}}^j/\hat{v}_\tau^j$ is decreasing, or

$$\hat{v}_{\tau\tau}^j(\tau)\hat{\mathcal{V}}^j(\tau) < \hat{v}_\tau^j(\tau)\hat{\mathcal{V}}_\tau^j(\tau)$$

or equivalently

$$\int_{x_{j-1}}^{x_j} \hat{v}_{x\tau\tau}(x,\tau)dx \int_{x_{j-1}}^{x_j} \hat{\mathcal{V}}_x(x,\tau)dx \leq \int_{x_{j-1}}^{x_j} \hat{v}_{x\tau}(x,\tau)dx \int_{x_{j-1}}^{x_j} \hat{\mathcal{V}}_{x\tau}(x,\tau)dx.$$

Let $x' \leq x''$. Then, by the Ahlswede-Daykin Inequality (see Karlin and Rinott (1980), Theorem 2.1) it is enough that

$$\hat{v}_{x\tau\tau}(x',\tau)\hat{\mathcal{V}}_x(x'',\tau) \leq \hat{v}_{x\tau}(x'',\tau)\hat{\mathcal{V}}_{x\tau}(x',\tau)$$

and

$$\hat{v}_{x\tau\tau}(x'',\tau)\hat{\mathcal{V}}_x(x',\tau) \leq \hat{v}_{x\tau}(x'',\tau)\hat{\mathcal{V}}_{x\tau}(x',\tau).$$

Each of these holds when $x' = x''$ by (1). Using (2), the first condition is equivalent to

$$\frac{\hat{\mathcal{V}}_x(x'',\tau)}{\hat{v}_{x\tau}(x'',\tau)} \leq \frac{\hat{\mathcal{V}}_{x\tau}(x',\tau)}{\hat{v}_{x\tau\tau}(x',\tau)},$$

and this is automatic, since $\hat{\mathcal{V}}_x(\cdot,\tau)/\hat{v}_{x\tau}(\cdot,\tau)$ is strictly increasing.

The second condition is equivalent to

$$\beta(x'',x') \equiv \hat{v}_{x\tau}(x'',\tau)\hat{\mathcal{V}}_{x\tau}(x',\tau) - \hat{v}_{x\tau\tau}(x'',\tau)\hat{\mathcal{V}}_x(x',\tau) \geq 0.$$

Since $\beta(x',x') \geq 0$ by (1), it is enough that where $\beta(x'',x') = 0$, $\beta_{x''}(x'',x') > 0$. But,

$$\beta_{x''}(x'',x') = \hat{v}_{xx\tau}(x'',\tau)\hat{\mathcal{V}}_{x\tau}(x',\tau) - \hat{v}_{xx\tau\tau}(x'',\tau)\hat{\mathcal{V}}_x(x',\tau) \geq 0$$

or, using that $\beta(x'',x') = 0$ and cancelling,

$$\frac{\hat{v}_{xx\tau}(x'',\tau)}{\hat{v}_{x\tau}(x'',\tau)} - \frac{\hat{v}_{xx\tau\tau}(x'',\tau)}{\hat{v}_{x\tau\tau}(x'',\tau)} \geq 0$$

which, using (2) is equivalent to

$$\hat{v}_{x\tau}\hat{v}_{xx\tau\tau} - \hat{v}_{xx\tau}\hat{v}_{x\tau\tau} \leq 0$$

which is to say that $\hat{v}_{x\tau}$ log-submodular as per (4). □

B.6 Details for the Healthcare Model

B.6.1 Assumption 2 in the Health Insurance Model

Let us return to the health insurance example. Because there are so many steps from the primitives of this model to the structure of v, it is difficult to check the conditions of Proposition 10 (they can be verified for some special cases such as when c is linear). But, it is easy to apply Proposition 11 as long as costs are not so concave as to create discontinuities in spending as l varies.

Proposition 12 *Consider the canonical model of demand for health insurance, where $\bar{\omega}$ is the maximum value of ω and where $c(a,x)$ is any analytic function satisfying our conditions and such that $c_{aa} \geq -\frac{1}{\bar{\omega}}$. Assume for simplicity that there is $\underline{\omega} > 0$ such that for all (ψ,θ) in the support of G, $\omega \geq \underline{\omega}$, and that the lower bound $\underline{\psi}$ of Ψ is strictly positive. Then, v is analytic and $\frac{v_{\omega x}(\cdot,\psi,\theta)}{v_{\psi x}(\cdot,\psi,\theta)}$ is not constant for any (ψ,θ), and so Proposition 11 applies.*

Proof To see that v is analytic note first that $a^*(l,\omega,x)$ is determined by the *FOC*

$$1 - \frac{1}{\omega}\left(a^*(l,\omega,x) - l\right) = c_a(a^*(l,\omega,x),x)$$

and so

$$a_x^*(l,\omega,x) = \frac{c_{ax}(a^*(l,\omega,x),x)}{\frac{1}{\omega} + c_{aa}(a^*(l,\omega,x),x)}.$$

But then, since $c_{aa} \geq -\frac{1}{\bar{\omega}}$, $\omega \geq \underline{\omega} > 0$, and c is analytic, a^* is analytic as well, and hence so is $z(l,\omega,x) \equiv b(a^*(l,\omega,x),l,\omega) - c(a^*(l,\omega,x),x)$. But then, since $\underline{\psi}$ is bounded away from zero $v(x,\psi,\omega) \equiv -\frac{1}{\psi}\log\int e^{-\psi z(l,\omega,x)}f$ is analytic as well.

Let us turn to $\frac{v_{\omega x}(\cdot,\psi,\theta)}{v_{\psi x}(\cdot,\psi,\theta)}$. Recall that $v_{\psi x}(\cdot,\psi,\theta) > 0$, which is the intuitive statement that higher risk aversion makes improved insurance more valuable. We will show that $v_{\omega x}(0,\psi,\theta) = 0$, but that there exist values of x for which $v_{\omega x}(x,\psi,\theta) > 0$. Thus, $\frac{v_{\omega x}(0,\psi,\theta)}{v_{\psi x}(0,\psi,\theta)} = 0$ but for some x, $\frac{v_{\omega x}(x,\psi,\theta)}{v_{\psi x}(x,\psi,\theta)} > 0$, and thus, as claimed, $\frac{v_{\omega x}(\cdot,\psi,\theta)}{v_{\psi x}(\cdot,\psi,\theta)}$ is not constant.

To execute on this, note first that when $x = 0$, it follows from (B.6.1) that $a^*(l,\omega,0) = l$. But then, $z = -l$, and so, since $v(x,\psi,\theta) = -\frac{1}{\psi}\log\int e^{-\psi z(l,x,\omega)}f$, it follows that

$$v(0,\psi,\theta) = -\frac{1}{\psi}\log\int e^{\psi l}f$$

which is independent of ω. Thus, $v_\omega(0,\psi,\theta) = 0$. But, when $x > 0$ and using that $\omega \geq \underline{\omega} > 0$, we have $a^*(x,\psi,\theta) > l$, and thus by the envelope theorem

$$z_\omega(l,x,\omega) = b_\omega(a^*(l,x,\omega),l,\omega) = \frac{1}{2\omega^2}\left(a^*(l,x,\omega) - l\right)^2 > 0$$

and so $v_\omega(x,\psi,\theta) > 0$ as well, and thus, since $v_\omega(0,\psi,\theta) = 0$, there is $x' \in (0,x)$ such that $v_{x\omega}(x',\psi,\theta) > 0$, and we are done. □

B.7 Computational Details

Simulated Population of Consumers. We simulate a population of consumers using the parameter estimates reported in Column 3 of Table 3 and Appendix Table A.8 of Marone and Sabety (2022). We first construct a population of households in terms of simple demographic characteristics (such as age and gender), and then construct each household's type θ using the reported parameters. As in Marone and Sabety (2022), we model a household as a group of individuals, each of whom is characterized by an age, a gender, and a health risk score.

We construct a population of households to match characteristics of the U.S. population. We start the construction of each household with a "head of household." This person is female with 50 percent probability and has a uniform distribution of age between 22 and 65. We assume that 90 percent of households have a spouse present, and when present, that the spouse is of the opposite gender to the head of household. Spouses draw an age from a normal distribution with

mean equal to the age of the head of household and a standard deviation of 4, subject to bounds between 22 and 65. We further assume each household has up to 4 children, where each child exists with 15 percent probability, independently of one another and of the presence of a spouse. Conditional on existing, each child is female with 50 percent probability and draws their age from a uniform distribution between 0 and 18. Finally, we assume that all individuals draw a risk score from a log-normal distribution with mean positively related to age, such that for individual i: $\log(riskscore_i) \sim N(\frac{age_i}{20}, 1)$. We censor the right tail of the risk score distribution such that no individual can have a risk score that is more than five standard deviations above the uncensored mean. Our baseline population contains 10,000 households. Increasing the number of households does not change our results.

With this simulated population in hand, we then apply the parameter estimates to construct the type ψ and $\theta = (\omega, F)$ for each household. We make one adjustment, which is to cap the risk aversion parameter at a value of 5.[62] Summary statistics on the population distribution of demographics and resulting household types are reported in Table B.1.

Numerical Algorithm for Computing Optimal Menus. We calculate optimal premium schedules numerically given a fixed set of potential contracts $\{x^k\}_{k=0}^K$. Note that we cannot calculate optimal menus in the case of a continuum of contracts because there are no closed-form solutions for key required objects: consumer utility $v(x, \psi, \theta)$ and insurer costs $\gamma(x, \theta)$. These objects must therefore be pre-calculated for each consumer type θ and each pre-specified contract x. The largest number of contracts on which we calculate optimal menus is 65. Our theoretical convergence result implies that this approach approximates the continuum case.

Our numerical algorithm for finding optimal prices mirrors the logic of the perturbation argument underlying our necessary conditions stated in Theorem 1. The algorithm proceeds as follows. Start from a candidate price schedule $\rho(x)$, and let $p^k = \rho(x^k) - \rho(x^{k-1})$ be the incremental premium between adjacent contracts. Starting from the first increment $k = 1$, consider a small perturbation to the incremental premium p^k, holding all other incremental premiums fixed. According to Theorem 1, at any optimal menu, this perturbation should not have a first-order effect on the insurer's payoff. Use a bounded one-dimensional optimizer to find a locally optimal p^1, around which the insurer's payoff cannot be improved.[63] Proceed to the second increment $k = 2$ and repeat this process, now optimizing over p^2, holding all other incremental premiums fixed. Proceed through all remaining increments up to K. At this point, restart at $k = 1$, and repeat the entire loop again. Once payoffs are unresponsive (within some tolerance) to small perturbations at every incremental premium k, a price schedule that fulfills the necessary conditions

[62] We express monetary amounts in thousands of dollars, so dividing our coefficients of absolute risk aversion by 1,000 makes them comparable to other settings where monetary amounts are measured in dollars.

[63] We use the commercial optimization packages available through MATLAB.

for local optimality has been found.

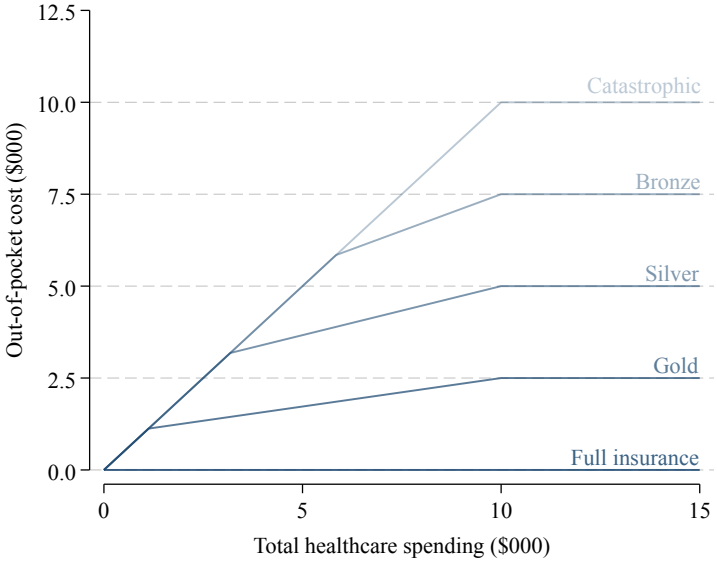
Given the complexity of this nonlinear optimization problem, the standard caveat applies that it is impossible to guarantee that a local optimum is a global optimum. In principle, this is exactly the same roadblock that prevents us from analytically deriving general sufficient conditions for optimality in the full problem. To reduce the risk of settling at a poor local optimum, we repeat this procedure from multiple starting points, each generated by varying the bounds placed on the search for the incremental premiums, and we retain the price schedule that achieves the highest payoff. In practice, the resulting solution is essentially unchanged across these starting points.

Table B.1. Population Summary Statistics

Sample demographic	Mean	Percentile				
		10	25	Median	75	90
Demographics						
Number of adults	1.9	2.0	2.0	2.0	2.0	2.0
Number of children	0.6	0.0	0.0	0.0	1.0	2.0
Average age of household adults	43.5	26.2	32.6	43.6	54.3	60.7
Dimensions of type θ						
Health state distribution parameter μ	1.5	0.3	0.8	1.6	2.2	2.7
σ	1.0	0.8	0.9	1.0	1.2	1.3
κ	0.6	0.1	0.3	0.5	0.9	1.3
Moral hazard parameter ω	1.4	0.8	1.0	1.3	1.7	1.9
Risk aversion parameter ψ	0.9	0.2	0.4	0.6	1.1	1.9
Resulting characteristics						
CE of equal odds gamble between \$0 and \$100 (\$)	48.9	47.6	48.6	49.2	49.5	49.7
Expected total spending, null contract (\$000)	9.8	2.9	4.2	7.3	12.4	20.2
full insurance (\$000)	11.2	4.0	5.6	8.7	13.9	21.6

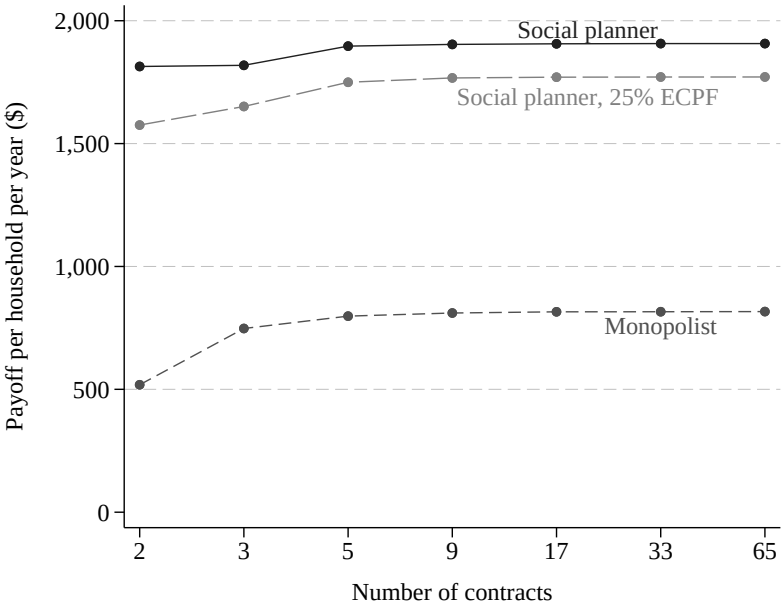
Notes: The table shows descriptive statistics for our simulated population of 10,000 households. Note that the moral hazard parameter and coefficient of absolute risk aversion are reported relative to thousands of dollars.

Figure B.2. Potential Contracts



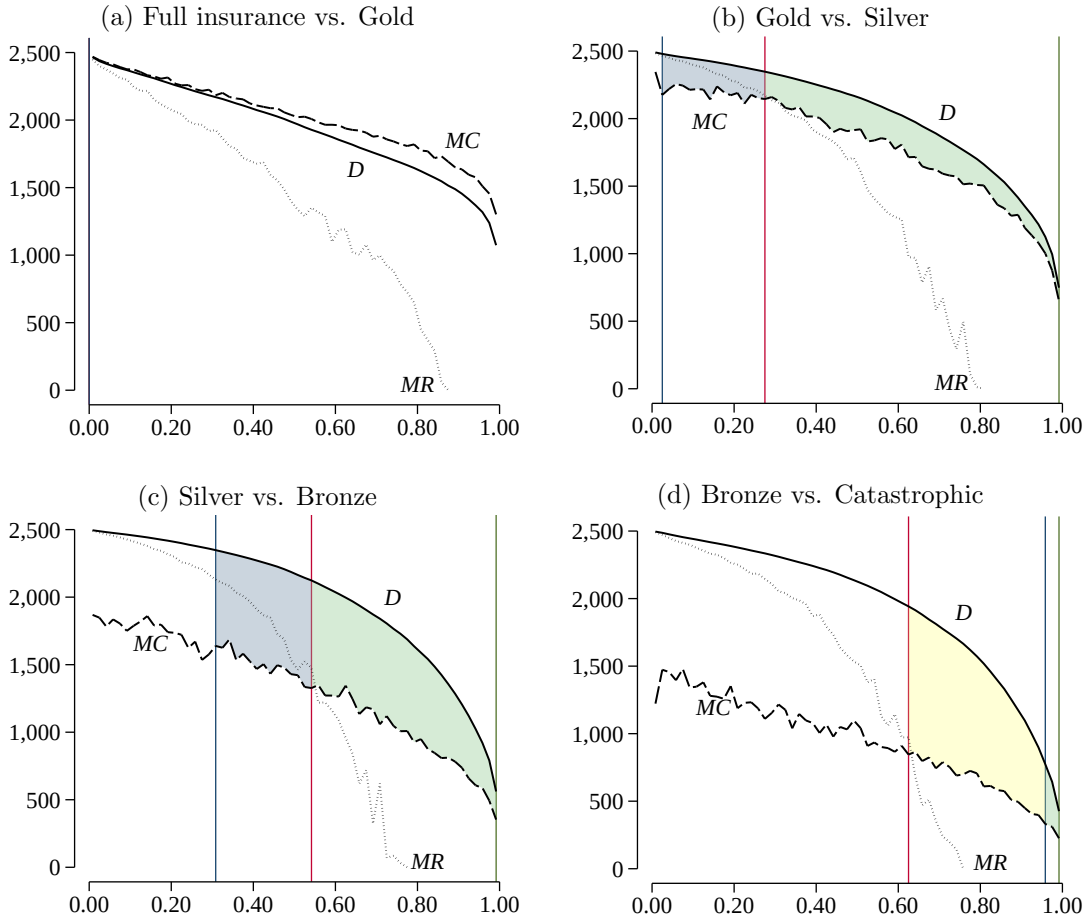
Notes: The figure shows our focal set of potential contracts. The base level of coverage x_0 provided by the government is the Catastrophic contract.

Figure B.3. Convergence



Notes: The figure shows optimal insurer payoffs as a function of the number of contracts used in the potential contract space. Insurer payoffs are reported on a per-consumer per-year basis, and are measured relative to allocating all consumers to the Catastrophic contract.

Figure B.4. Welfare Comparison: Monopoly versus Competition



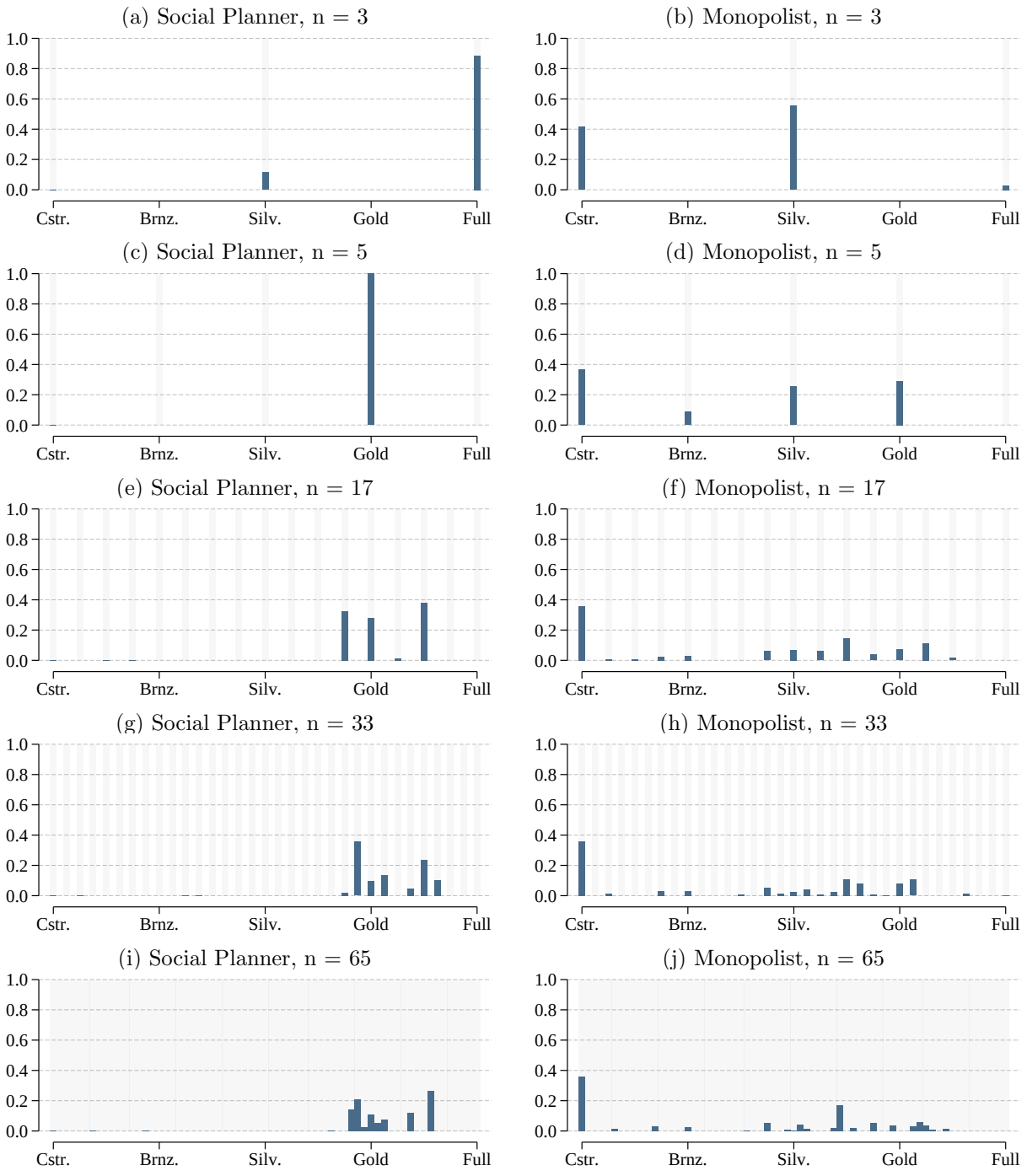
Notes: The figure illustrates the differential welfare consequences of a monopoly insurer versus perfect competition, increment by increment. Each panel represents the "market for incremental coverage" between a pair of adjacent contracts. The vertical axes are measured in dollars, and the horizontal axes report the percentage of consumers choosing a given incremental level of coverage, ordered by their marginal willingness to pay. The solid line (WTP) represents consumers' willingness to pay, the dotted line (MC) the marginal cost of providing the increment, and the dashed line (MR) a monopolist's marginal revenue. On each panel, the red vertical line marks the monopoly allocation (where MR meets MC), the blue vertical line marks the competitive equilibrium allocation of Azevedo and Gottlieb (2017) (computed separately), and the green vertical line marks the efficient allocation (where WTP meets MC). The green-shaded area is the deadweight loss common to both market structures; the yellow-shaded area is the additional deadweight loss under monopoly; and the blue-shaded area is the deadweight loss under competition. The MC and MR curves are constructed as connected binned scatter plots using 100 points.

Table B.2. Contract-by-Contract Markups

	Average cost *$000s*				Premium *$000s*				Premium / Average cost			
Scenario	Brnz.	Slvr.	Gold	Full	Brnz.	Slvr.	Gold	Full	Brnz.	Slvr.	Gold	Full
Social planner, $w = (1, 1, 1)$			4.08		0.29	0.51	0.83	3.33			0.20	
Social planner, $w = (0.8, 1, 1)$	0.17	1.02	4.76		1.44	2.80	4.64	7.17	8.75	2.76	0.97	
Monopolist, $w = (0, 1, 0)$	0.52	2.18	5.71		1.93	4.05	6.38	8.91	3.69	1.86	1.12	
Competitive equilibrium	0.76	3.11	5.58		0.76	3.11	5.58	8.09	1.00	1.00	1.00	

Notes: The table shows average costs, premiums, and the ratio of premium to average cost at the optimal menu of our three focal insurers as well as at the competitive equilibrium. Costs are incremental relative to the Catastrophic contract, and note that the premium and insurer cost for consumers in the Catastrophic contract is always zero.

Figure B.5. Optimal Allocations as Density of Contract Space Increases



Notes: The figure shows the percentage of consumers allocated to each contract under the optimal menus chosen by a social planner and a monopolist as the density of the contract space increases. The gray bars identify the set of *potential* contracts available to the menu designer, while the blue bars show the actual allocations. The left-hand side panels show the allocations chosen by the social planner, while the right-hand side panels show the allocations chosen by the monopolist. The rows correspond to 3, 5, 17, 33, and 65 potential contracts, respectively.