Multidimensional Product Design*

André Veiga† and E. Glen Weyl‡

May 17, 2012

Abstract

Insurance contracts and other products are designed to attract individuals who are healthy or otherwise particularly valuable. An empirically relevant model of this process requires individuals heterogeneous in both preferences and values. A simple price-theoretic analysis is possible when individual heterogeneity is of high dimension relative to the firm's product design instruments. Necessary conditions for profit and welfare maximization depend on moments of the distribution of individual heterogeneity. Our main result is that the power of an instrument to sort for value is proportional to the covariance, within the set of marginal individuals, between value and marginal utility for the instrument. Existing models assume unidimensional heterogeneity or require restrictive assumptions that imply the absence of sorting. Our analysis applies in settings with non-transferable utility, consumption externalities, non-linear pricing, third-degree discrimination and imperfect competition. We discuss applications to broadcast media, the credit card industry and imperfectly competitive insurance provision.

Keywords: Platforms, screening, heterogeneity, cream-skimming, intensive v. extensive margin
JEL Classication Codes: D43, D82, I13, L15

---

*We thank Eduardo Azevedo, Bruno Jullien, Renato Gomes, Jacques Crémer, Michal Fabinger, Matthew Gentzkow, Jesse Shapiro and seminar participants at the Universities of Chicago, Iowa and Toulouse for helpful comments. We thank Jonathan Baker, Jean Tirole, Andrei Hagiu and Preston McAfee for excellent formal discussions of the paper. Anthony Zhang, Alison Lanski and Michael Olijnyk provided excellent research assistance. We acknowledge the financial support of the Net Institute, NOKIA, and the research program on intellectual property at the Institut d'Economie Industrielle. Veiga is grateful to the Becker Friedman Institute for hosting him during a visit to the University of Chicago that allowed us to develop this research.

†Toulouse School of Economics (TSE), Manufacture de Tabacs, 21 allées de Brienne, 31000 Toulouse, France; andre.veiga@tse-fr.eu.

‡Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago IL, 60637; and Toulouse School of Economics (TSE), Manufacture de Tabacs, 21 allées de Brienne, 31000 Toulouse, France weyl@uchicago.edu.

1 Introduction

Radio stations famously introduced melodramatic "soap operas" to appeal to housewives who controlled family purchase decisions. Melodrama attracted those listeners most valuable to advertisers, and therefore also most valuable to the radio stations. In this and other industries, the heterogeneity of individual *preferences* and *values* is a key feature that firms take into account when designing their products. However, existing models do not allow for multidimensional heterogeneity, or require restrictive assumptions on preferences and technologies to remain tractable, which reduces their empirical relevance. We develop a formulation of this problem that is both general and tractable. Our key modeling strategy is to allow individuals to have multidimensional heterogeneity but restrict firms to use a finite number of product design instruments. This "smooths" the model and allows us to use a multidimensional version of the Leibniz rule to obtain optimality conditions in terms of aggregate market quantities. Our main result is that the power of an instrument to sort for valuable individuals is proportional to the *covariance, within the set of marginal individuals, between the value of individuals and their marginal utility for the instrument.*

We illustrate our approach in Section 2 with the example of health insurance, a product that Rothschild and Stiglitz [1976] famously argued is designed to attract the healthiest individuals. We model a monopoly insurer's choice of a uniform premium and level of coverage, and consider individuals heterogeneous along multiple dimensions, namely their risk and risk aversion. Importantly, individuals differ in their marginal utility of coverage and in their cost of provision, and there is heterogeneity even within marginal individuals. The model produces the standard profit maximization distortions of price (upwards, as in Cournot [1838]) and coverage (catering to marginal individuals, as in Spence [1975]).

The main result of the paper is the characterization of a novel *sorting* effect. If those marginal individuals with greater marginal utility of coverage are also those with higher cost, then increasing coverage sorts for costly individuals by disproportionately catering to their preferences, which implies an additional cost for the insurer. The strenght of this sorting effect is proportional to the covariance, within marginal individuals, between marginal utility for coverage (preferences) and the cost of provision (values). The effect relies on the presence of heterogeneity within the set of marginal individuals, which is a natural consequence of multidimensional heterogeneity but is typically absent from models with unidimensional heterogeneity, which therefore ignore this effect.

In Section 3, we solve a general version of our model to illustrate the flexibility of our approach. This section is somewhat abstract and technical, and may be skipped with minimal loss of understanding. We impose weak technical assumptions to ensure differentiability. We then allow for arbitrary dimensions of individual heterogeneity, any finite number of product design instruments by the firm, for a flexible specification of individual preferences and for intensive participation decisions on the part of individuals. We also allow firms to be "platforms" by assuming that some of the characteristics that make them attractive are determined endogenously by their user base. This feature is included in a tractable manner by formulating the platform's problem as an optimization over these characteristics directly, but constrained to the equations that endogenously determine them. We derive necessary optimality conditions for the firm's product design instruments, contrast the cases of profit and welfare maximization and discuss the relevant distortions. We also show how the model can be applied to the study of multi-sided platforms and markets where non-platforms compete imperfectly.

Section 4 discusses the departures of our work from the previous literature which give rise to our results. In our model, firms determine the number of individuals purchasing their product, as in Cournot [1838], and use multiple price and non-price instruments to attract individuals, as in Spence [1975]. However, individuals in our model have asymmetric information about their value, as in Akerlof [1970] and Einav and Finkelstein [2011], which together with the multidimensionality of the firm's product design instruments makes our analysis close to that of Rothschild and Stiglitz [1976]. Nonetheles, in contrast to that paper and subsequent single-dimensional screening models like Mussa and Rosen [1978], we consider individuals with multidimensional heterogeneity. On the other hand, we depart from multidimensional screening models like Rochet and Choné [1998] by restricting firms to a finite number of product design instruments but allowing for more flexible specifications of preferences and technologies.

In Section 5, we apply our model to three industries. First, we highlight the specificities of the relationship between multiple instruments in the absence of prices by modelling a radio station faced with listeners heterogeneous in wealth. Second, we explore a credit card platform's use of non-linear pricing, as a generalization of the Rochet and Tirole [2003] model of two-sided markets. The platform's pricing influences how much each buyer uses the card, which is an intensive margin of participation. The heterogeneity of marginal buyers drives the platform's pricing to cost since high per-transaction fees repell the most valuable buyers. Our third application is a study of an imperfectly competitive insurance market where two insurers choose both premiums and levels of coverage, as an extension of Einav and Finkelstein [2011].[1] We show how greater competition mitigates the Cournot [1838] and Spence [1975] distortions, but exacerbates a selection distortion in prices similar to that in Akerlof [1970] and similarly aggravates a cream-skimming distortion in the level of coverage similar to the results of Rothschild and Stiglitz [1976].

Section 6 discusses potential avenues for future work and concludes.

2 A Simple Example

This section illustrates our approach and main result with minimal technical details and generality. We consider a monopoly health insurer and a continuum of individuals with mass normalized to 1. The insurer chooses a premium P and a level of coverage ρ (for instance, the deductible or coinsurance). Then, each individual decides whether to purchase insurance.

Individuals are heterogeneous in their risk (θ^1) and risk aversion (θ^2). The "type" vector $\boldsymbol{\theta} \equiv (\theta^1, \theta^2) \in \mathbb{R}^2$ is each individual's private information. Types are distributed in the population according to a smooth density $f(\boldsymbol{\theta})$, which is common knowledge.

If she buys insurance, individual $\boldsymbol{\theta}$ obtains utility $u \equiv s(\rho; \boldsymbol{\theta}) - P$, and imposes on the insurer the expected cost $c(\rho; \boldsymbol{\theta})$. Individuals who do not purchase insurance obtain zero utility, so the set of covered individuals and the set of marginal individuals are, respectively,

$$\Theta \equiv \{\boldsymbol{\theta} : s(\rho; \boldsymbol{\theta}) \geq P\} \qquad \text{and} \qquad \partial\Theta \equiv \{\boldsymbol{\theta} : s(\rho; \boldsymbol{\theta}) = P\}.$$

[1] In fact, Einav and Finkelstein [2011] write that "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 individuals may have multidimensional heterogeneity (like expected cost and risk preferences)." We aim to provide such a characterization.

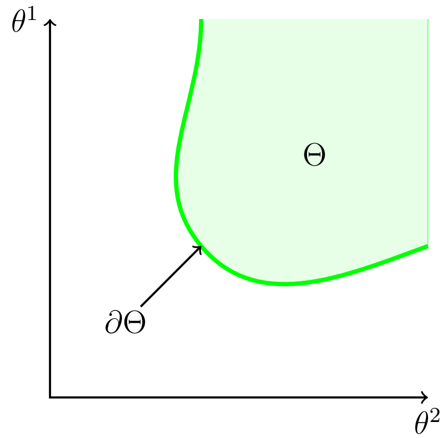

Figure 1: A set of covered and marginal individuals for $\boldsymbol{\theta} \in \mathbb{R}^2$.

The crucial features of the model are that individuals differ in their preferences for coverage and in their values to the insurer, and that there is heterogeneity even within the set of marginal individuals. Below we omit several functional arguments for notational simplicity.

We begin with the maximization of welfare, which equals the difference between value created and cost of provision, or $W \equiv \int_\Theta (s-c)\, f d\boldsymbol{\theta}$. We differentiate this integral with respect to P using a multidimensional version of the Leibniz Rule. As with its one-dimensional version, increasing the premium causes a transfer of surplus from infra-marginal individuals to the insurer (effect on the integrand) and causes marginal individuals to stop buying insurance proportionally to their marginal valuation for the premium (effect on the boundary). We obtain the first-order condition

$$\begin{aligned}
\frac{\partial W}{\partial P} &= \int_\Theta \frac{\partial}{\partial P}(s-c) f d\boldsymbol{\theta} + \int_{\partial\Theta} \frac{\partial u}{\partial P}\,(s-c)\, f d\boldsymbol{\tau} \\
0 &= 0 + \left(\int_{\partial\Theta} f d\boldsymbol{\tau}\right) \frac{\int_{\partial\Theta} -(s-c)\, f d\boldsymbol{\tau}}{\int_{\partial\Theta} f d\boldsymbol{\tau}} \\
0 &= M\mathbb{E}\left[-(s-c) \mid \partial\Theta\right] \\
0 &= P - \underbrace{\mathbb{E}\left[c \mid \partial\Theta\right]}_{\text{marginal cost}},
\end{aligned}$$

where $d\boldsymbol{\tau}$ is the vectorial area element normal to the surface $\partial\Theta$ which, intuitively, captures to what extent each marginal individual changes her participation decision in response to a change in the insurer's instruments. We discuss $d\boldsymbol{\tau}$ formally in Section 3 and Appendix A. The density of marginal individuals, $M \equiv \int_{\partial\Theta} f d\boldsymbol{\tau}$, in this setting plays a role analogous to the role played, in models with uni-dimensional heterogeneity, by the density of types evaluated at the marginal individual. Unsurprisingly, a welfare maximizing insurer covers individuals until the marginal individual's willingness to pay equals her cost. Notice that the condition is in terms of an aditional covered individual, rather than a unit increase in price. We are therefore following Spence [1975] in thinking of the firm as choosing the optimal number of covered individuals (by means of the premium).

A welfare maximizing insurer uses the level of coverage to increase the value of insurance to covered individuals and to sort for valuable marginal individuals, while the optimal number of covered individuals is held fixed by price. The Leibniz Rule captures the change in surplus and cost of infra-marginal individuals (effect on the integrand) and the extent to which marginal individuals are attracted to buying insurance, which happens proportionally to their marginal valuation for coverage (effect on the boundary). We obtain the following expression for $\frac{\partial W}{\partial \rho}$;

$$\begin{aligned}
& \int_{\Theta}\left(\frac{\partial s}{\partial \rho}-\frac{\partial c}{\partial \rho}\right) f d\boldsymbol{\theta}+\int_{\partial \Theta} \frac{\partial u}{\partial \rho}(s-c) f d\boldsymbol{\tau} \\
= \; & N \mathbb{E}\left[\frac{\partial s}{d \rho}-\frac{\partial c}{\partial \rho} \mid \Theta\right]+M \mathbb{E}\left[\frac{\partial u}{\partial \rho}(s-c) \mid \partial \Theta\right] \\
= \; & N \mathbb{E}\left[\frac{\partial s}{d \rho}-\frac{\partial c}{\partial \rho} \mid \Theta\right]+M \operatorname{Cov}\left(\frac{\partial u}{\partial \rho}, s-c \mid \partial \Theta\right)+M \mathbb{E}\left[\frac{\partial u}{\partial \rho} \mid \partial \Theta\right] \underbrace{\mathbb{E}[s-c \mid \partial \Theta]}_{0} \\
= \; & \underbrace{N \mathbb{E}\left[\frac{\partial s}{\partial \rho}-\frac{\partial c}{\partial \rho} \mid \Theta\right]}_{\text{intensive effect}}-\underbrace{M \operatorname{Cov}\left(\frac{\partial u}{\partial \rho}, c \mid \partial \Theta\right)}_{\text{sorting effect}}=0,
\end{aligned}$$

where $N \equiv \int_{\Theta} f d\boldsymbol{\theta}$ is the share of covered individuals. We have used the definition of covariance, $\mathbb{E}[XY] \equiv \operatorname{Cov}(X, Y)+\mathbb{E}[X] \mathbb{E}[Y]$, and $\mathbb{E}[s-c \mid \partial\Theta]=0$ from the optimal choice of the number of covered individuals.

The first term captures the *intensive* effect, or the average increase in surplus net of costs, $\mathbb{E}\left[\frac{\partial s}{\partial \rho}-\frac{\partial c}{\partial \rho} \mid \Theta\right]$, to the mass of covered individuals, N. The second term captures the *sorting effect* of coverage, or the extent to which a change in coverage affects the composition of the set of individuals buying insurance, holding fixed their number. When coverage is increased, each marginal individual is attracted proportionally to her marginal valuation, $\frac{\partial u}{\partial \rho}$. Each attracted individual then makes a different contribution to welfare since individuals differ in cost. Assuming that those who value additional coverage more highly are also less healthy, the sorting cost of additional coverage is the extent to which it disproportionally attracts these costly individuals into purchasing insurance, which is captured by $\mathrm{Cov}\left(\frac{\partial u}{\partial \rho}, c \mid \partial\Theta\right)$, the covariance among marginal individuals between preferences for coverage and cost of provision.

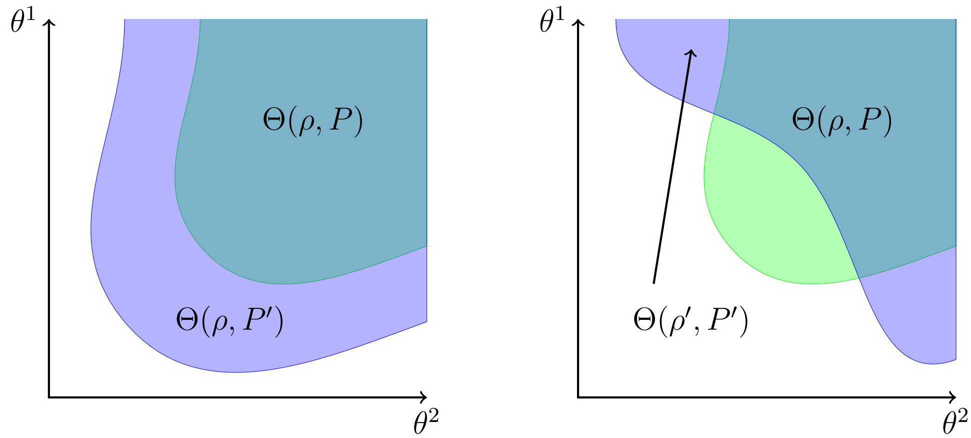

Figure 2: Changes in a set of covered individuals following a decrease in price (left), and following a decrease in coverage where price adjusts to hold fixed the number of covered individuals (right), for $P' < P, \rho' < \rho$.

Allowing for heterogeneity in both preferences and values within the set of marginal individuals is essential to characterize the sorting effect, since the covariance term vanishes otherwise. This effect is absent from most models with one dimension of heterogeneity, where there is typically a unique marginal individual, but its presence is a natural consequence of heterogeneity along multiple dimensions.[2] This is especially relevant in the insurance setting since both risk and risk aversion have been shown to be relevant by Finkelstein and McGarry [2006]. Considering this rich heterogeneity can also formalize the results of Bundorf et al. [Forthcoming], who show that the more correlated tastes for insurance are with cost of coverage, the less socially desirable it is to offer insurance at actuarially fair rates. The result above shows that, more generally, the welfare maximizing wedge between marginal utility for coverage

[2]It is possible, but uncommon, to have a heterogeneous margin with a single dimension of heterogeneity. For instance, consider the same exercise for the utility function $u = \theta\rho - \theta^2 - P$, for $\theta \in \mathbb{R}$ and arbitrary $c(\rho; \theta)$. See also Araujo and Moreira [2010] for a specification in this spirit.

and the actuarially fair marginal cost of providing it, $\mathbb{E}\left[\frac{du}{d\rho} - \frac{\partial c}{\partial \rho} \mid \Theta\right]$, is proportional to the covariance between tastes and costs, which quantifies the sorting effect.

For specific coverage dimensions, this covariance could be measured using survey data and ex-post health expenditures, as done by Hendren [2011]. For example, for a reduction in co-payment, the sorting term would equal the covariance between expected utility from a percent reduction in insurance payments and the consequent increase in expected payments by the insurer; these could be calculated from the joint distribution of risk-preferences, income and health status.[3]

Consider now the maximization of profit, $\Pi \equiv \int_\Theta (P-c)\, f d\boldsymbol{\theta}$. The first-order condition for the premium is

$$\begin{aligned}
\frac{\partial \Pi}{\partial P} &= \int_\Theta f d\boldsymbol{\theta} + \int_{\partial\Theta} \frac{\partial u}{\partial P}(P-c)\, f d\boldsymbol{\tau} \\
0 &= N + M\mathbb{E}\left[-(s-c) \mid \partial\Theta\right] \\
0 &= \underbrace{P - \frac{N}{M}}_{\text{Marginal Revenue}} - \mathbb{E}\left[c \mid \partial\Theta\right].
\end{aligned}$$

The profit maximizer covers individuals until marginal revenue is equated to marginal cost, where the wedge between premium and marginal revenue, $\frac{N}{M}$, illustrates the Cournot [1838] market power distortion.[4]

The profit maximizing level of coverage satisfies

$$\begin{aligned}
\frac{\partial \Pi}{\partial \rho} &= -\int_\Theta \frac{\partial c}{\partial \rho} f d\boldsymbol{\theta} + \int_{\partial\Theta} \frac{\partial u}{\partial \rho}(P-c)\, f d\boldsymbol{\tau} \\
0 &= -N\mathbb{E}\left[\frac{\partial c}{\partial \rho} \mid \Theta\right] + M\mathbb{E}\left[\frac{\partial u}{\partial \rho}(P-c) \mid \partial\Theta\right] \\
0 &= -N\mathbb{E}\left[\frac{\partial c}{\partial \rho} \mid \Theta\right] + M\mathrm{Cov}\left(\frac{\partial u}{\partial \rho}, P-c \mid \partial\Theta\right) + M\mathbb{E}\left[\frac{\partial u}{\partial \rho} \mid \partial\Theta\right]\underbrace{\mathbb{E}\left[P-c \mid \partial\Theta\right]}_{N/M} \\
0 &= \underbrace{N\mathbb{E}\left[\frac{\partial u}{\partial \rho} \mid \partial\Theta\right]}_{\text{Spence term}} - N\mathbb{E}\left[\frac{\partial c}{\partial \rho} \mid \Theta\right] - \underbrace{M\mathrm{Cov}\left(\frac{\partial u}{\partial \rho}, c \mid \partial\Theta\right)}_{\text{sorting effect}}.
\end{aligned}$$

[3] Weingarten [2011] performs an empirical calibration of this type in the context of market power in smartphone applications.

[4] It is common to express $-N/M$ as $Q\,(dP/dQ)$, where Q is quantity supplied by a monopoly and P is price.

We again used the definition of covariance and the optimal choice of the number of covered individuals. The benefit of additional coverage depends on the extent it allows the profit maximizer to increase the premium while holding fixed the number of covered individuals. On the other hand, the insurer must also consider the increase in the cost of covering infra-marginal individuals and the sorting cost of coverage. Since the marginal effect of coverage on the number of covered individuals is determined by the preferences of marginal individuals, a profit maximizer caters to them when choosing coverage, as described by Spence [1975]. Unlike the Cournot distortion, which always leads to under-provision, the Spence distortion may lead to excessive or insufficient coverage depending on the relative preferences of marginal and infra-marginal individuals. Finally, marginal individuals have zero utility so sorting has a first-order effect only on profit. Given a marginal set $\partial\Theta$, the incentive to sort is the same for a profit and welfare maximizing monopoly.

3 The general model

The goal of this section is to provide a formally rigorous and general illustration of the flexibility of our approach. It will also allow us to easily derive the results for the applications in Section 5. This section is more technical and abstract than the others, and can be skipped with minimal loss in understanding of the remaining sections.

We will allow the firm to use any finite number of instruments. We will also allow individuals to have a flexible specification of preferences with any finite number of dimensions of heterogeneity, to generate heterogeneous consumption externalities towards other participating individuals, to differ in their preferences over these externalities, and to choose their intensity of participation. Due to the presence of consumption externalities, we will refer to this firm as a "platform."

3.1 Setup

We consider a monopoly platform that chooses a vector of $\mathcal{R}$ *instruments*, $\boldsymbol{\rho} \equiv (\rho^1, \rho^2, ...) \in \mathbb{R}^{\mathcal{R}}$, with components indexed by $l \in \{1, 2, .., \mathcal{R}\}$. We denote $\rho^1 = \rho^\star$ as the platform's *focal instrument*. We will think of this instrument as being used by the platform to determine the number of participating individuals, much like we did for the premium in Section 2. In principle, any instrument can be thought of as focal, although the natural choice is an instrument that transfers utility from individuals to the firm, and towards which no individual is indifferent. We will assume below that marginal utilities for the focal instrument are signed and bounded away from zero.

We consider a continuum of individuals with mass normalized to 1. Each individual's type is a vector of characteristics $\boldsymbol{\theta} \in \mathbb{R}^{\mathcal{T}}$, $\mathcal{T} \in \mathbb{N}$, which represents each individual's residual private information.[5] Types are distributed in the population according to the probability density function $f(\boldsymbol{\theta}) : \mathbb{R}^{\mathcal{T}} \mapsto \mathbb{R}$, which is common knowledge. If individual $\boldsymbol{\theta}$ participates, she obtains utility $u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) : \mathbb{R}^{\mathcal{R}+\mathcal{K}+\mathcal{T}} \mapsto \mathbb{R}$ and contributes $\pi(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) : \mathbb{R}^{\mathcal{R}+\mathcal{K}+\mathcal{T}} \mapsto \mathbb{R}$ to the platform's profit, where $\boldsymbol{K}$ is a vector of platform characteristics discussed below.[6]

Each individual decides whether to join the platform, with the intensive margin of participation discussed below. We normalize outside options to zero without loss of generality.[7] The set of participating individuals and of marginal individuals are, respectively,

$$\Theta \equiv \{\boldsymbol{\theta} : u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) \geq 0\} \qquad \text{and} \qquad \partial\Theta \equiv \{\boldsymbol{\theta} : u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) = 0\}.$$

The mass of participating individuals and the density of marginal individuals are, respectively,

$$N \equiv \int_{\Theta} f(\boldsymbol{\theta})\, d\boldsymbol{\theta} \qquad \text{and} \qquad M \equiv \int_{\partial\Theta} \frac{f(\boldsymbol{\theta})}{\|\nabla_{\boldsymbol{\theta}} u\|} d\tau,$$

where $1/\|\nabla_{\boldsymbol{\theta}} u\|\, d\tau$ captures the speed of outward expansion of the boundary at each boundary point, and $d\tau$ is the surface element of $\partial\Theta$, both discussed in further detail in Appendix A. The density of marginal individuals M captures how responsive the entire set of participants is to changes in the platform's instruments. Economically, $1/\|\nabla_{\boldsymbol{\theta}} u\|$ is a measure of how reactive is the set of participants, nearby each marginal individual, to changes in the instruments. It captures the extent to which individuals similar to each marginal type also have utility close to zero so, when this term is

[5] The platform has no private information. If it can contract some dimensions of individual heterogeneity, it can segment the market according to this information and optimize its instruments within each segment.

[6] In Section 2, we had $\boldsymbol{\rho} = (P, \rho)$, $\rho^\star = P$, $u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) = s(\rho; \boldsymbol{\theta}) - P$ and $\pi(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) = P - c(\rho; \boldsymbol{\theta})$.

[7] If outside options are constant, they can be included in the type $\boldsymbol{\theta}$. Therefore, we assume that instruments and characteristics have no effect on non-participants. See Segal [1999] for a study of contracting with such externalities.

large, a small change in the instruments causes a substancial density of individuals to change their utility enough to alter their participation decision.

We define the following conditional expectation operators for any $X(\boldsymbol{\theta})$,

$$\mathbb{E}\left[X(\boldsymbol{\theta}) \mid \Theta\right] \equiv \frac{\int_{\Theta} X(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}}{N} \qquad \text{and} \qquad \mathbb{E}\left[X(\boldsymbol{\theta}) \mid \partial\Theta\right] \equiv \frac{\int_{\partial\Theta} X(\boldsymbol{\theta}) \frac{f(\boldsymbol{\theta})}{\|\nabla_{\boldsymbol{\theta}} u\|} d\tau}{M}.$$

Individuals may generate consumption externalities towards other participating individuals. This is modeled by considering $\mathbf{K} \equiv (K^1, K^2, ...) \in \mathbb{R}^{\mathcal{K}}$, a vector of $\mathcal{K}$ platform *characteristics* with components indexed by $i \in \{1, 2, ..., \mathcal{K}\}$. Each component of $\boldsymbol{K}$ is determined by the composition of the platform's participants. We assume that the technology determining each characteristic is as follows: if individual $\boldsymbol{\theta}$ participates, she makes a contribution $k^i\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right)$ to characteristic K^i, and each characteristic is the integral of all individual contributions, or

$$K^i \equiv \int_{\Theta} k^i\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right) f\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta}.$$

For instance, if $k^i\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right)$ is individual wealth then K^i is the total wealth of participants.[8] We assume that individual expectations of $\boldsymbol{K}$ are correct for any choice of $\boldsymbol{\rho}$.

We model intensive margins of participation by allowing individual contributions to depend on the levels of instruments and characteristics. For instance, $k^i\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right)$ may be the demand of individual $\boldsymbol{\theta}$ for the platform's product as a function of the instruments, and other individuals or the platform itself may care about the total demand K^i.

Profit and social welfare are, respectively,

$$\Pi \equiv \int_{\Theta} \pi\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right) f\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta} \qquad \text{and} \qquad W \equiv \int_{\Theta} \left(u\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right) + \pi\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right)\right) f\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta}.$$

The problems we consider is the choice of $\boldsymbol{\rho}$ to maximize Π or W. However, the problem is made more transparent when formulated in term of choosing $(\boldsymbol{\rho}, \boldsymbol{K})$ to maximize Π or W, subject to the $\mathcal{K}$ constraints $\int_{\Theta} k^i\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}\right) f\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta} = K^i, \forall i$. The Lagrangians for the two problems are

[8] The model can accommodate preferences that depend the average or variance of contributions. For instance, if k^i is individual age, we can define an additional characteristic $K^j = N$ and define preferences to depend on K^i/N.

$$\begin{aligned}\mathcal{L}^{\Pi} &= \Pi + \sum_{i=1}^{\mathcal{K}} \left(\int_{\Theta} k^i \left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta} \right) f\left(\boldsymbol{\theta} \right) d\boldsymbol{\theta} - K^i \right) \lambda^{\Pi i} \qquad \text{and} \\ \mathcal{L}^{W} &= W + \sum_{i=1}^{\mathcal{K}} \left(\int_{\Theta} k^i \left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta} \right) f\left(\boldsymbol{\theta} \right) d\boldsymbol{\theta} - K^i \right) \lambda^{W i},\end{aligned}$$

where $\lambda^{\Pi i}$ and λ^{Wi} are the marginal values of characteristic i in each problem. It is possible to tackle both problems simultaneously by considering

$$\mathcal{L}^{V} = \int_{\Theta} v\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta} \right) f\left(\boldsymbol{\theta} \right) d\boldsymbol{\theta} - \sum_{i=1}^{\mathcal{K}} K^i \lambda^{Vi},$$

where $\lambda^{Vi} \in \left\{ \lambda^{\Pi i}, \lambda^{Wi} \right\}$, and the (private or social) value of individual $\boldsymbol{\theta}$ is

$$v\left(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta} \right) \in \left\{ \pi + \sum_{i=1}^{\mathcal{K}} k^i \lambda^{\Pi i}, u + \pi + \sum_{i=1}^{\mathcal{K}} k^i \lambda^{Wi} \right\},$$

omitting the arguments of π, u and k^i. This value consists of the contribution of individual $\boldsymbol{\theta}$ to profit, her utility (in welfare maximization), and her contributions to all characteristics weighted by the marginal values of those characteristics in the problem under consideration.[9] We assume the existence of a solution to this problem and, for notational simplicity, we will henceforth omit functional arguments.

3.2 Technical Details

We make the following technical assumptions.

1. f is twice continuously differentiable, atomless and has finite moments.

2. u, π and $k^i, \forall i$ are twice continuously differentiable.

3. u has bounded derivatives.

4. $\forall \left(\boldsymbol{\rho}, \boldsymbol{K} \right)$, the gradient $\nabla_{\boldsymbol{\theta}} u$ has Euclidean norm bounded away from zero.

5. $\forall \left(\boldsymbol{\rho}, \boldsymbol{K} \right)$, there is a finite density of marginal individuals ($M \neq 0$).

[9] Assuming that profit is a linear aggregation of individual contributions simplifies the exposition by allowing both problems to be re-cast in this way, but the model can be solved for any differentiable profit function $\Pi \left(\boldsymbol{\rho}, \boldsymbol{K} \right)$.

6. $\forall \boldsymbol{\theta}$, $\frac{du}{d\rho^\star}$ is signed and bounded away from zero.

Lemma 1. *The integrals W, Π and $K^i, \forall i$ are differentiable in $\boldsymbol{\rho}$.*

Proof. Assumptions 1-4 satisfy the conditions for differentiability in Uryas'ev [1994].[10] □

Lemma 2. *For a differentiable function $G(\rho) = \int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})\geq 0} g(\rho;\boldsymbol{\theta})f(\boldsymbol{\theta})d\boldsymbol{\theta}$, we have*

$$\frac{dG}{d\rho} \;=\; N\mathbb{E}\left[\frac{\partial}{\partial\rho}g(\rho;\boldsymbol{\theta}) \mid \boldsymbol{\theta} : u(\rho;\boldsymbol{\theta}) \geq 0\right] + M\mathbb{E}\left[g(\rho;\boldsymbol{\theta})\frac{\partial u}{\partial\rho} \mid \boldsymbol{\theta} : u(\rho;\boldsymbol{\theta}) = 0\right].$$

Proof. See Appendix A.[11] □

3.3 Results

We are now able to express our main results.

Proposition 3. *A necessary condition for the optimal choice of the focal instrument $\rho^\star$ is*

$$\underbrace{\mathbb{E}\left[v \mid \partial\Theta\right]}_{\textit{marginal value}} + \underbrace{\frac{Cov\left(\frac{\partial u}{\partial \rho^\star}, v \mid \partial\Theta\right)}{\mathbb{E}\left[\frac{\partial u}{\partial\rho^\star} \mid \partial\Theta\right]}}_{\textit{heterogeneous preference for } \rho^\star} + \underbrace{\frac{N\mathbb{E}\left[\frac{\partial v}{\partial\rho^\star} \mid \Theta\right]}{M\mathbb{E}\left[\frac{\partial u}{\partial\rho^\star} \mid \partial\Theta\right]}}_{\textit{intensive effect}} = 0. \tag{1}$$

Proof. See Appendix B. □

The optimal value of the focal instrument $\rho^\star$ is such that the total value from including an additional marginal individual attracted by $\rho^\star$ is zero. The average value of marginal individuals is $\mathbb{E}\left[v \mid \partial\Theta\right]$. Since $\mathbb{E}\left[u \mid \partial\Theta\right] = 0$, any profit maximization distortion in this term arises from the difference between $\lambda^{\Pi i}$ and $\lambda^{W i}$. When individuals are heterogeneously attracted by $\rho^\star$, the value of an additional marginal

[10] Intuitively, the differentiation of the interior is immediate from Assumption 2. Assumptions 1-4 ensure that the region of integration changes smoothly with $\boldsymbol{\rho}$. Since u has bounded derivatives, the change is not "explosive." Since $\|du/d\boldsymbol{\theta}\|$ is bounded below, the density of marginal individuals is not infinite (which would lead to a discontinuous derivative). Notice $\|du/d\boldsymbol{\theta}\|$ is a denominator in the formula for the derivative of the integrals and therefore could not be zero.

[11] For additional details, see Flanders [1973] and Uryas'ev [1994].

individual must take into consideration the extent to which those most attracted by $\rho^\star$ are also particularly valuable, which is captured by Cov $\left(\frac{\partial u}{\partial \rho^\star}, v \mid \partial\Theta\right)$.[12]

The third term generalizes the Cournot [1838] term. The value of N infra-marginal individuals changes on average by $\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right]$ when $\rho^\star$ increases to attract an additional marginal individual. To do so, the required change in $\rho^\star$ is inversely proportional to the density of marginal individuals M and to the responsiveness of marginal individuals to this instrument, $\mathbb{E}\left[\frac{du}{d\rho^\star} \mid \partial\Theta\right]$. Under welfare maximization, if $\rho^\star$ transfers utility between individuals and the platform ($\frac{du}{d\rho^\star} + \frac{d\pi}{d\rho^\star} = 0$), and if contributions to characteristics are fixed, we have $\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right] = 0$. In this case, the third term vanishes as in Section 2.[13]

Proposition 4. *A necessary condition for the optimal choice of each non-focal instrument ρ^l is*

$$\underbrace{M\,Cov\left(\frac{\partial u}{\partial \rho^l} + \frac{\partial u}{\partial \rho^\star}\frac{\partial \rho^\star}{\partial \rho^l}, v \mid \partial\Theta\right)}_{\textit{sorting effect}} + \underbrace{N\mathbb{E}\left[\frac{dv}{d\rho^l} + \frac{dv}{d\rho^\star}\frac{\partial \rho^\star}{\partial \rho^l} \mid \Theta\right]}_{\textit{intensive effect}} = 0. \tag{2}$$

where $\frac{\partial \rho^\star}{\partial \rho^l} = -\mathbb{E}\left[\frac{du}{d\rho^l} \mid \partial\Theta\right] / \mathbb{E}\left[\frac{du}{d\rho^\star} \mid \partial\Theta\right]$.

Proof. See Appendix B. □

The platform uses each non-focal instruments ρ^l to sort for valuable marginal individuals and to increase the value of infra-marginal individuals. Since we think of the number of individuals as being determined by $\rho^\star$, we should also think of any change in ρ^l as accompanied by an adjustment of $\rho^\star$ that holds fixed the number of individuals. The required adjustment is proportional to the effectiveness of ρ^l in attracting marginal individuals and inversely proportional to the effectiveness of $\rho^\star$ in repelling marginal individuals, hence $\frac{\partial \rho^\star}{\partial \rho^l} = \mathbb{E}\left[\frac{du}{d\rho^l} \mid \partial\Theta\right] / \left(-\mathbb{E}\left[\frac{du}{d\rho^\star} \mid \partial\Theta\right]\right)$.[14]

The power of ρ^l to sort for valuable marginal individuals is proportional to the density of marginal individuals, M, multiplied by the covariance between the value of

[12]In Section 2, $\mathbb{E}[v \mid \partial\Theta] = \mathbb{E}[P - c \mid \partial\Theta]$ under both profit and welfare maximization, and Cov $\left(P - c, \frac{du}{dP} \mid \partial\Theta\right) = 0$ because preferences for price are homogeneous. To our knowledge, the first instance of a covariance term being used to quantify incentives is Dixit and Sandmo [1977], who consider an optimal linear tax scheme that affects individuals heterogeneously.

[13]In Section 2, profit maximization implied $\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right] = 1$ and $\mathbb{E}\left[\frac{du}{d\rho^\star} \mid \partial\Theta\right] = -1$, so we obtained the traditional Cournot term N/M.

[14]Notice that $\frac{\partial \rho^\star}{\partial \rho^l}$ solves $0 = \frac{dN}{d\rho^l} = \frac{\partial N}{\partial \rho^l} + \frac{\partial N}{\partial \rho^\star}\frac{\partial \rho^\star}{\partial \rho^l}$.

these individuals and their preferences for the non-focal instrument, and preferences for the focal instrument to the extent that it adjusts, $\mathrm{Cov}\left(\frac{\partial u}{\partial \rho^l}+\frac{\partial u}{\partial \rho^\star}\frac{\partial \rho^\star}{\partial \rho^l}, v \mid \partial\Theta\right)$. Thus ρ^l is useful in sorting for value to the extent that valuable individuals are attracted by ρ^l more intensely than by the focal instrument.[15]

The second term generalizes the Spence [1975] distortion. A change in ρ^l induces an average change in the values of infra-marginal individuals of $\mathbb{E}\left[\frac{dv}{d\rho^l}+\frac{dv}{d\rho^\star}\frac{\partial \rho^\star}{\partial \rho^l} \mid \Theta\right]$. If there are no externalities and $\rho^\star$ transfers utility between individuals and the platform ($\frac{du}{d\rho^\star}+\frac{d\pi}{d\rho^\star}=0$), the adjustment of $\rho^\star$ causes only socially neutral redistribution. In that case, a welfare maximizer considers $\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right]=0$, thus taking into account only the direct effects of ρ^l on the average value of individuals, $\mathbb{E}\left[\frac{dv}{d\rho^l} \mid \Theta\right]$. However, a profit maximizer considers the direct impact of ρ^l on profit $\mathbb{E}\left[\frac{d\pi}{d\rho^l} \mid \Theta\right]$ as well as the extent to which this leads to higher profit through the adjustment of $\rho^\star$, $\mathbb{E}\left[\frac{d\pi}{d\rho^\star}\frac{\partial \rho^\star}{\partial \rho^l} \mid \Theta\right]$. Since the adjustment of the focal instrument, $\frac{\partial \rho^\star}{\partial \rho^l}$, is proportional to the preferences of marginal individuals for ρ^l, a profit maximizer chooses ρ^l catering to these individuals, as described by Spence [1975], to the extent that an upward adjustment in $\rho^\star$ is profitable.

The final step is to determine the marginal values of characteristics.

Proposition 5. *The marginal values of characteristics* $\left\{\lambda^{Vj}\right\}_{j=1}^{j=\mathcal{K}}$ *solve the system of* $\mathcal{K}$ *equations of the form*

$$\underbrace{M\,Cov\left(\frac{\partial u}{\partial K^j}+\frac{\partial u}{\partial \rho^\star}\frac{\partial \rho^\star}{\partial K^j}, v \mid \partial\Theta\right)}_{\textit{sorting effect}}+\underbrace{N\mathbb{E}\left[\frac{dv}{dK^j}+\frac{dv}{d\rho^\star}\frac{\partial \rho^\star}{\partial K^j} \mid \Theta\right]}_{\textit{intensive effect}}=\lambda^{Vj}, \tag{3}$$

where $\frac{\partial \rho^\star}{\partial K^j}=-\mathbb{E}\left[\frac{du}{dK^j} \mid \partial\Theta\right]/\mathbb{E}\left[\frac{du}{d\rho^\star} \mid \partial\Theta\right]$.

Proof. See Appendix B. □

This system of equations defines the marginal values of characteristics recursively since $\left\{\lambda^{Vj}\right\}_{j=1}^{j=\mathcal{K}}$ is present on the left-hand side of each equation, as part of the definition of v. The reason is that a change in the level of one characteristic changes the participating set and thereby has an effect on the levels of all characteristics. In turn,

[15] When preferences for $\rho^\star$ are homogeneous, $\frac{\partial u}{\partial \rho^\star}\frac{\partial \rho^\star}{\partial \rho^l}$ drops out of the covariance term as in Section 2.

each of these effects produces secondary effects on all characteristics proportional to the first, and so forth recursively. The logic of Equation (3) is otherwise analogous to that of Equation (2), illustrating that characteristics are essentially instruments that cannot be directly determined. Hence the right-hand sides of the two equations, which are the shadow price of instruments (zero) and characteristics ($\left\{\lambda^{Vj}\right\}_{j=1}^{j=\mathcal{K}}$).

3.4 Multiple Sides

Platforms like radio stations can discriminate between "sides" of individuals (listeners and advertisers) who affect each other's payoffs via consumption externalities (advertisers care about the wealth of participating listeners). The model can be applied to multi-sided markets by repeating the structure and procedures described above for each side, while taking into account that the value of a characteristic generated by one side depends on the preferences of individuals on another side for that characteristic.

For simplicity, we consider a monopoly platform and two sides, $s \in \{1, 2\}$. Individuals on side s have types $\boldsymbol{\theta}^s \in \mathbb{R}^{\mathcal{T}^s}$ distributed according to $f^s(\boldsymbol{\theta}^s)$, obtain utility $u^s\left(\boldsymbol{\rho}^s, K^{-s}; \boldsymbol{\theta}^s\right)$, and contribute $\pi^s\left(\boldsymbol{\rho}^s, K^{-s}; \boldsymbol{\theta}^s\right)$ to the platform's profit. The instruments for side s are $\boldsymbol{\rho}^s = (\rho^{s\star}, \rho^{s2}, ...) \in \mathbb{R}^{\mathcal{R}^s}$ where $\rho^{s\star}$ is the focal instrument for that side. The sets of participating and marginal individuals on side s are, respectively, $\Theta^s = \{\boldsymbol{\theta}^s : u^s \geq 0\}$ and $\partial\Theta^s = \{\boldsymbol{\theta}^s : u^s = 0\}$. The share of participating individuals and the density of the margin on that side are, respectively, N^s and M^s.

We assume that $\boldsymbol{\rho}^s$ affects only side s, and that each side generates only one characteristic, which enters the preferences of individuals on the opposite side.[16] Let $k^s\left(\boldsymbol{\rho}^s, K^{-s}; \boldsymbol{\theta}^s\right)$ be the contribution by individual $\boldsymbol{\theta}^s$ to K^s, the characteristic generated by side s. We then define $K^s = \int_{\Theta^s} k^s f^s d\boldsymbol{\theta}^s$. The (private or social) value of individual $\boldsymbol{\theta}^s$ is

$$v^s \in \left\{\pi^s + k^s\lambda^{\Pi s}, u^s + \pi^s + k^s\lambda^{Ws}\right\},$$

where $\lambda^{\Pi s}$ and λ^{Ws} are the marginal values of characteristic K^s.

[16] It is possible to include additional sides, allow each side to generate multiple characteristics, and allow each non-focal instrument to enter the utility of individuals on all sides. The only requirement is a focal instrument on each side affects only that side. In such a setting, the optimality conditions sum over the sorting and intensive effects for each side while using the focal instruments to hold fixed the number of individuals on each side. Assuming each characteristic is generated by a single side is without loss of generality since preferences can be changed to depend on the sum of the two characteristics instead of a jointly determined characteristic. Details are available from the authors upon request.

The optimality conditions for $\boldsymbol{\rho}^s$ are obtained from Equations 1 and 2 by considering individual value v^s, utility u^s, and sets Θ^s and $\partial\Theta^s$. To determine the marginal values of characteristics, one need only to keep in mind that characteristic $-s$ affects the utilities and values of individuals on side s. We include this last condition below.

Proposition 6. *The marginal values of characteristics $\lambda^{V1}, \lambda^{V2}$ solve the system of 2 equations of the form*

$$M^s \, Cov\left(\frac{\partial u^s}{\partial K^{-s}} + \frac{\partial u^s}{\partial \rho^{s\star}} \frac{d\rho^{s\star}}{dK^{-s}}, v^s \mid \partial\Theta^s\right) + N^s \mathbb{E}\left[\frac{dv^s}{dK^{-s}} + \frac{dv^s}{d\rho^{s\star}} \frac{d\rho^{s\star}}{dK^{-s}} \mid \Theta^s\right] = \lambda^{V(-s)}, \quad (4)$$

where $\frac{d\rho^{s\star}}{dK^{-s}} = -\mathbb{E}\left[\frac{\partial u^s}{\partial K^{-s}} \mid \partial\Theta^s\right] / \mathbb{E}\left[\frac{\partial u^s}{\partial K^{-s}} \mid \partial\Theta^s\right]$.

3.5 Imperfect Competition

The model can also be applied to imperfectly competitive markets, by distinguishing between individuals indifferent between buying from one firm and not buying (the *exiting* margin), and those indifferent between the two firms (the *switching* margin).

We consider a mass 1 of individuals, with types $\boldsymbol{\theta} \in \mathbb{R}^{\mathcal{T}}$, distributed $\boldsymbol{\theta} \sim f$, and a duopoly indexed by $d \in \{1, 2\}$. Firm d chooses instruments $\boldsymbol{\rho}^d \in \mathbb{R}^{\mathcal{R}}$, where $\rho^{d\star}$ is the focal instrument. Each individual purchases from the firm that gives her the highest utility. If individual $\boldsymbol{\theta}$ purchases from d, she obtains $u^d\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right)$ and contributes $\pi^d\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right)$ to the profit of that firm.

Individuals purchasing from d are the set $\Theta^d = \left\{\boldsymbol{\theta} : u^d \geq \max\left\{u^{-d}, 0\right\}\right\}$ with mass N^d. The set of marginal participants is the union of the exiting and switching margins,

$$\partial\Theta^{Xd} \cup \partial\Theta^S = \left\{\boldsymbol{\theta} : u^d = 0\right\} \cup \left\{\boldsymbol{\theta} : u^d = u^{-d} \geq 0\right\},$$

with densities M^{Xd} and M^S.

We assume that the joint distribution of (π^1, π^2, u^1, u^2) induced by f is symmetric in (u^1, u^2) at any symmetric $\boldsymbol{\rho}^d = \boldsymbol{\rho}^{-d} = \boldsymbol{\rho}$. We further assume that firms have symmetric technologies, so $\pi^d(\boldsymbol{\rho}; \boldsymbol{\theta}) = \pi^{-d}(\boldsymbol{\rho}; \boldsymbol{\theta}) = \pi(\boldsymbol{\rho}; \boldsymbol{\theta})$, and also consider an equilibrium where firms choose their instruments symmetrically. Each firm faces the same residual demand curve and therefore, for symmetric $\boldsymbol{\rho}$, both firms solve the same maximization problem.

Profit maximization conditions for firm d can be obtained directly from Equations 1 and 2 by considering individual values π^d and marginal set $\partial\Theta^{Xd} \cup \partial\Theta^S$, since a duopolist acts as an oligopolist faced with the market's residual demand curve.

A welfare maximizer considers the utility of all individuals and joint industry profits. Our symmetry assumptions imply that it is irrelevant to industry profits which firm an individual is served by, so the industry is viewed as a single firm and the relevant margin is only the exiting margin $\partial\Theta^{Xd}$. Welfare maximization conditions can then be obtained for firm d from Equations 1 and 2 by considering individual value $\pi + u^d$ and restricting attention to the exiting margin $\partial\Theta^{Xd}$.

To see this, notice that a welfare maximizing firm d solves

$$\begin{aligned} & \max_{\boldsymbol{\rho}^d} \int_{\mathbb{R}^{\mathcal{T}}} \left(\left(\pi^d + u^d\right) \mathbf{1}_{\left\{\boldsymbol{\theta} \in \Theta^d\right\}} + \left(\pi^{-d} + u^{-d}\right) \mathbf{1}_{\left\{\boldsymbol{\theta} \in \Theta^{-d}\right\}} \right) f d\boldsymbol{\theta} \\ = \; & \max_{\boldsymbol{\rho}^d} \int_{\Theta^d} \left(\pi^d + u^d - \left(\pi^{-d} + u^{-d}\right) \mathbf{1}_{\{\boldsymbol{\theta} \in \partial\Theta^S\}}\right) f d\boldsymbol{\theta}, \end{aligned}$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function. By Equations 1 and 2 and the symmetry assumptions above, the solution to this problem is the same as that obtained by considering individual value $\pi + u^d$ and restricting attention to the exiting margin $\partial\Theta^{Xd}$, as described above. For instance, we would obtain

$$\left(M^{Xd} + M^S\right) \mathbb{E}\left[\pi^d + u^d - \left(\pi^{-d} + u^d\right) \mathbf{1}_{\{\boldsymbol{\theta} \in \partial\Theta^S\}} \mid \partial\Theta^{Xd} \cup \partial\Theta^S\right] = M^{Xd} \mathbb{E}\left[\pi \mid \partial\Theta^{Xd}\right].$$

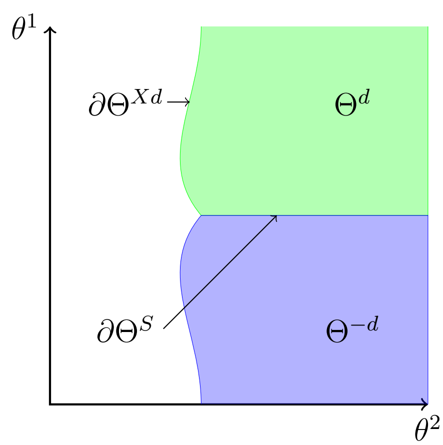

Figure 3: The exiting and switching margins for $\boldsymbol{\theta} \in \mathbb{R}^2$ in a symmetric equilibrium.

3.6 Second-Order Conditions and Multiplicity

The analysis above focuses on necessary first-order conditions for optimality. We are not aware of non-trivial conditions on primitives that ensure the satisfaction of second-order conditions in general. We have worked out some special cases, which are available upon request, but a detailed discussion is beyond the scope of this paper.

We emphasize that first-order conditions on their own are useful for at least four reasons. First, they quantify the marginal incentive to sort and show under which assumptions the sorting effect is necessarily absent. Second, if the value of some terms in the first-order conditions may be estimated then, given an observed equilibrium, other terms may be identified from the first-order conditions in the spirit of Rosse [1970], an approach common in the new empirical industrial organization literature. Third, first-order conditions identify what quantities must be estimated to determine non-parametrically the social costs and benefits of a small change in instruments departing from an observed equilibrium, as in Chetty [2009]. Fourth, in the case of a monopolist or social planner as in Milgrom and Shannon [1994], or in the case of a symmetric stable equilibrium between competing oligopolists as in Echenique [2002], a global rise in the first-order condition for a given instrument leaving others fixed will lead to a rise in that relevant instrument. Additional applications requiring further analysis of second-order conditions are suggested as future research in Section 6.

Multiplicity of equilibria in the decisions of consumers is common in the presence of consumption externalities. This arises because, given the instruments chosen by the platform, the decision of each individual depends on her expectations about the decisions of other individuals, as pointed out by Rohlfs [1974].[17] In a multi-sided monopoly setting with quasilinear preferences and homogeneous values, Weyl [2010] obtains uniqueness by having the platform's prices depend on the number of individuals on every side, thereby allowing the platform to make the number of individuals on each side invariant to changes in expectations.[18] These techniques do not apply immediately to our setting since individuals differ in their preferences and values. However, there are conditions under which it is possible to extend these results to our environment, although a detailed discussion would be excessively lenghty. The

[17]See also Katz and Shapiro [1985] and Caillaud and Jullien [2003] regarding multiplicity in individual decisions.

[18]This is interpreted as a reduced-form model of dynamic pricing. This is modeled explicitly by Cabral [2011], where the equilibrium is unique.

technical details can be found in Sandberg [1981].[19] The intuition is that instruments must have sufficient power over each characteristic independently in order to make the levels of characteristics invariant to changes in expectations. A platform must have as many effective instruments as there are characteristics, instruments must have a sufficiently independent effect on each characteristic, and the impact of instruments on characteristics must be sufficiently strong to overcome the feedback effects within characteristics.

4 Literature Review

One goal of our paper is to describe connections between a diverse set of existing models. We pay special attention to six areas of the literature: classic product design, classic contract theory, multidimensional screening, empirical work in industrial organization, empirical analyses of markets with asymmetric information, and recent price theoretical papers. Below we discuss how the assumptions typically made in each field determine which of the effects we describe are present. Table 1 summarizes these connections with several examples.

The classical treatment of product design is Spence [1975]'s model of a quality-choosing monopoly, which allows arbitrary preference heterogeneity. Weyl [2010] and White and Weyl [2011] show that the distortion described by Spence is present also in settings with price-choosing multi-sided platforms, where the number of individuals on each side plays the role of quality to the other sides. The sorting effect is absent in these models because individual values are homogeneous and individuals have only an extensive participation margin.[20] We illustrate the importance of these features by modeling a radio station platform faced with individuals heterogenous in their tastes and income in Subsection 5.1, and a credit card platform faced with individuals heterogeneous in their usage elasticities in Subsection 5.2.

Classic contract theory, surveyed by Bolton and Dewatripont [2004], has tended to

[19] Let $\hat{\boldsymbol{K}}$ be expectations of characteristics and $\boldsymbol{K}$ be their realizations. Let $\boldsymbol{K} = \boldsymbol{\kappa}\left(\boldsymbol{\rho}, \hat{\boldsymbol{K}}\right)$ for some function $\boldsymbol{\kappa} : \mathbb{R}^{\mathcal{R}+\mathcal{K}} \mapsto \mathbb{R}^{\mathcal{K}}$ and let instruments $\boldsymbol{\rho}$ be contingent on expectations. That is, instruments are the function $\boldsymbol{\rho}(\hat{\boldsymbol{K}})$. Assume a desired feasible equilibrium where $\boldsymbol{K} = \boldsymbol{K}^{\star}$. This can be implemented uniquely when $\boldsymbol{\kappa}\left(\boldsymbol{\rho}(\hat{\boldsymbol{K}}), \hat{\boldsymbol{K}}\right) = \boldsymbol{K}^{\star}$ has a unique global implicit function solution $\boldsymbol{\rho}\left(\hat{\boldsymbol{K}}\right)$. Necessary and sufficient conditions are given in Sandberg [1981].

[20] Individuals purchasing different amounts typically have different values to the firm.

focus on settings with a single dimension of heterogeneity. This restriction typically implies that there is a single marginal individual (in monopoly models like Stiglitz [1977]) or that a positive mass of individuals is marginal (in perfectly competitive models like Rothschild and Stiglitz [1976]). In the first case, there is no sorting because there is no heterogeneity among marginal individuals. In the second, the incentive to cream-skim from competitors is infinitely large, preventing the existence of equilibria, as show by Riley [1979]. Also, when a single parameter captures multiple aspects of incentives, one must typically choose between modeling the effect of nonlinear pricing on intensive participation decisions, as in Mussa and Rosen [1978], and modeling the effect of prices on selection along an extensive margin, as in Akerlof [1970]. By considering imperfect competition, we obtain finite cream-skimming incentives which we can analyze jointly with the effect of non-linear pricing.[21] The application of Subsection 5.2 to the credit card industry features an analysis of both an extensive and intensive margins, and our modeling of imperfect competition in insurance in Subsection 5.3 yields finite cream-skimming distortions.

The literature on multidimensional screening, surveyed in Rochet and Stole [2003], addresses many of the concerns above by considering individuals heterogeneous along multiple dimensions and firms that use infinite-dimensional instruments (nonlinear price functions). Individuals have endogenously heterogeneous values to the firm since they purchase different goods. Analytic treatments of these models tend to require strong assumptions on preferences and on the distribution of heterogeneity, as in McAfee and McMillan [1988], Armstrong [1996], and Rochet and Choné [1998], so the literature has tended to focus on conditions under which full separation of types is possible. Our model departs from this literature by restricting the firm to an arbitrary, but finite, set of instruments which allows us to relax standard restrictions on the distribution of preferences and technologies.[22] We abstract from the literature's focus on type separation since the relative richness of heterogeneity in our model makes bunching inevitable and, in fact, the sorting effect stems from the bunching of marginal individuals. To illustrate these distinctions, Subsection 5.2 contrasts Rochet

[21]When the preferences and values are determined by a single parameter, the classic Spence [1973]-Mirrlees [1971] single crossing condition is often assumed to determine a (global) relationship between the two that allows the firm to screen for individual value based on their preferences. The (local) analogue of this condition in our model is the sign of the covariance term. See Araujo and Moreira [2010] for an example of a model where the single crossing condition does not hold.

[22]Wilson [1993] shows that, in many settings, tariffs with only a few parts are close to optimal.

and Stole [2002], who assume a particular distribution of heterogeneity and arbitrary non-linear tariffs, to a model with a two-part tariff and arbitrary heterogeneity.

Empirical work in industrial organization has increasingly incorporated the concerns discussed above. For instance, Berry [1994] and Berry et al. [1995] estimate preference heterogeneity in discrete choice demand models and Hendel [1999] considers an intensive margin. Mazzeo [2002] and Gentzkow et al. [2011] analyze firm choices of non-price product design instruments, while Nosko [2010] allows for second-degree discrimination and Leslie [2004] for third-degree discrimination. Chiappori and Salanié [2000] and Cardon and Hendel [2001] discuss selection, and Starc [2010] considers welfare losses from adverse selection and market power. A recent literature, surveyed by Rysman [2009], considers platform markets where heterogeneous externalities are found by Rysman [2004], Ryan and Tucker [Forthcoming], Cantillon and Yin [2008] and Lee [2010]. This literature, while allowing for rich heterogeneity, focuses on the computational analysis of parametric demand systems. The contribution that we aim for in this context is twofold. First, we aim to formalize which empirical moments are most important to identify in a flexible manner in order to quantify the economic effects of interest.[23] Second, the optimality conditions we derive can be used to estimate unobserved quantities, as in Rosse [1970].

A number of recent papers estimate the effects of asymmetric information in several settings, and develop models similar to ours in the richness of individual heterogeneity relative to firm instruments. Einav and Finkelstein [2011] characterize selection in perfectly competitive insurance provision where they allow only price to be chosen by insurers.[24] Einav et al. [Forthcoming] consider the choice of continuous non-price product characteristics by a monopolist, but use a reduced-form approach that does not allow for welfare analysis and does not relate the results to the primitives of individual heterogeneity.[25] Our contribution here is to provide a framework that is easily adaptable to the wide variety of settings tackled by this literature. In Subsection 5.3 we extend Einav and Finkelstein [2011] by modeling imperfect competition

[23]For example, Gentzkow and Shapiro [2010]'s analysis of newspaper choices of political slant includes a test of whether newspaper readers have heterogeneous values to the newspaper. Applying our model to this setting shows that their procedure would correspond to testing whether the covariance between the preferences of readers for slant and the values of readers to advertisers is different from zero.

[24]See Einav et al. [2010a] for an empirical application of this logic.

[25]Einav et al. [2010b] surveys a number of papers in the insurance setting which, like ours, consider intensive effects ("moral hazard") and third-degree discrimination ("pricing on observables").

between insurers in both premia and coverage levels.

Some recent theoretical work emphasizes the importance of moments of individual heterogeneity similar to those we discuss. Weyl and Tirole [2011] consider the use of market power to screen innovations heterogeneous in their consumer surplus and Azevedo and Leshno [2012] study the matching of heterogeneous users to firms in large markets.

Table 1 contains examples of the various relationships of our paper to the literature.

	Classical Product Design	Classical Contract Theory	Multi-dimensional Screening	Empirical Industrial Organization	Applied Work on Asymmetric Information
Cournot Distortion	Cournot	Stiglitz [77]	Armstrong [96]	Berry Levinsohn Pakes [95]	
Spence Distortion	Spence [75]			Mazzeo [02]	
Selection Distortion		Akerlof [70]	Rochet Stole [02]	Starc [11]	Einav Finkelstein [11]
Sorting			Armstrong [96]	Gentzkow Shapiro [10]	Einav et al. [Forthcoming]
Cream--Skimming Distortion		Rothschild Stiglitz [76]	Rochet Stole [02]	Gentzkow et al. [11]	
Intensive Margin		Mussa Rosen [78]	Rochet Chone [98]	Hendel [99]	Einav et al. [Forthcoming]
Consumption Externalities	Weyl [10]			Rysman [04]	
Third-Degree Discrimination	Pigou [32]	Segal [99]		Leslie [04]	

Table 1: Effects we describe and the literature.

5 Applications

The following section applies our model to three industries. The goal is to illustrate how it can be easily applied in practice, which of its feature are relevant in particular settings, and what insights can be obtained from applying the model to specific industries.

5.1 Non-Transferability in Broadcast Media

Broadcasters like television channels and radio stations offer free programming to viewers and listeners in order to raise revenue from advertisers. Anderson and Coate

[2005] modeled these settings as multi-sided platforms restricted in their transfers to individuals on one side of the market. We build off their model to illustrate, when individuals have heterogeneous preferences for all instruments, how the sorting power of an instrument depends on the covariance among marginal individuals between value and the relative preferences for several instruments.

We consider a radio station broadcaster of soap operas, as mentioned in Section (1), and two sides of the market, advertisers (A) and listeners (L). Advertisers have uni-dimensional types $\theta^A \sim f^A$ and obtain utility $u^A = \theta^A W - P^A$, where P^A is the uniform price charged to advertisers and W is the aggregate wealth of listeners. The preferences of advertisers are in the style of Anderson and Coate [2005]: listeners are vertically differentiated and there is a unique marginal advertiser, so there is no sorting of advertisers.

Listeners have multidimensional types $\boldsymbol{\theta}^L \sim f^L$, which may include gender, age, geography and other demographic traits. They obtain utility $u^L(\rho, m; \boldsymbol{\theta}^L)$, where ρ is the overall quality of the programming and m is the level of its melodrama.

Outside options are zero. For each side $s \in \{A, L\}$, the participating set is $\Theta^s = \{\boldsymbol{\theta}^s : u^s \geq 0\}$ and the set of marginal participants is $\partial\Theta^s = \{\boldsymbol{\theta}^s : u^s = 0\}$. The share of participants on side s is N^s and the density of the margin is M^s. Listener $\boldsymbol{\theta}^L$ has wealth $w(\boldsymbol{\theta}^L)$ so the aggregate wealth of listeners is $W = \int_{\Theta^L} w f^L d\boldsymbol{\theta}^L$. That is, listeners generates externalities but advertisers do not. Notice that the unique marginal advertiser has type $\theta^A = \frac{P^A}{W}$.

The platform incurs a cost of c^A per advertiser and a cost of $c^L(\rho, m)$ per listener. We will consider only the case of a profit maximizer, for whom the values of individual advertisers and listeners are

$$\pi^A = P^A - c^A \qquad \text{and} \qquad \pi^L = -c^L + w\lambda^W,$$

where λ^W is the marginal value of listener wealth to the profit maximizer. We will think of ρ and P^A as being used to determine the number of participants on each side.

Equation (1) prescribes that the price to advertisers equates marginal revenue to marginal cost, $P^A - \frac{N^A}{M^A} = c^A$. By Equation (4), $\lambda^W = N^A \frac{\partial \pi^A}{\partial P^A}\left[\frac{du^A}{dW} \mid \partial\Theta^A\right] = N^A \frac{P^A}{W}$ is the value of listener wealth.

From equation (1), the profit maximizing level of quality satisfies

$$\underbrace{-c^L + \mathbb{E}\left[w\lambda^W \mid \partial\Theta^L\right]}_{\text{value of marginal listener}} + \underbrace{\frac{\operatorname{Cov}\left(\frac{\partial u^L}{\partial \rho}, w\lambda^W \mid \partial\Theta^L\right)}{\mathbb{E}\left[\frac{\partial u^L}{\partial \rho} \mid \partial\Theta^L\right]}}_{\text{heterogeneous taste for quality}} - \underbrace{\frac{N^L \frac{\partial c^L}{\partial \rho}}{M^L \mathbb{E}\left[\frac{\partial u^L}{\partial \rho} \mid \partial\Theta^L\right]}}_{\text{market power of quality}} = 0.$$

The covariance term in this expression captures the extent to which valuable users are attracted by quality. The third term is the analogue of the Cournot market power term. It captures the cost of the quality necessary to attract an additional listener, much like price-setting profit maximizer lowers its revenue from infra-marginal buyers to attract an additional marginal buyer.

Equation (2) implies that the optimal level of melodrama satisfies

$$M^L \underbrace{\operatorname{Cov}\left(\frac{\partial u^L}{\partial m} + \frac{\partial u^L}{\partial \rho}\frac{\partial \rho}{\partial m}, w\lambda^W \mid \partial\Theta^L\right)}_{\text{purchasing power of women}} - \underbrace{N^L\left(\frac{\partial c^L}{\partial m} + \frac{\partial c^L}{\partial \rho}\frac{\partial \rho}{\partial m}\right)}_{\text{cost of melodrama}} = 0,$$

where $\frac{\partial \rho}{\partial m} = -\mathbb{E}\left[\frac{\partial u^L}{\partial m} \mid \partial\Theta^L\right] / \mathbb{E}\left[\frac{\partial u^L}{\partial \rho} \mid \partial\Theta^L\right]$ is the change in quality that holds fixed the number of listeners as melodrama. Melodrama is a useful instrument to attract valuable listeners if valuable marginal listeners respond to melodrama more than they respond to overall quality. Thus $\operatorname{Cov}\left(\frac{\partial u^L}{\partial m} + \frac{\partial u^L}{\partial \rho}\frac{\partial \rho}{\partial m}, w\lambda^W \mid \partial\Theta^L\right)$ captures the sorting power of a non-focal instrument when individuals have heterogeneous preferences for the focal instrument. This feature is absent when utility is quasilinear in prices, as in Section 2, but becomes relevant in industries where it is non-transferable. The sorting term quantifies the desirability of melodrama by the radio station and, the greater is this covariance in the population of listeners, the greater is the expected level of melodrama in the station's programming. Notice that the Spence [1976] distortion is evident in the consideration of the tastes of marginal listeners in $\frac{\partial \rho}{\partial m}$ and the tastes of marginal advertisers in λ^W.

5.2 Non-Linear Pricing by a Credit Card Platform

The credit card industry is a canonical example of a multi-sided market, famously studied by Rochet and Tirole [2003] upon whose model we base this application. Where broadcasters have no access to a price instrument on the listener side, credit card platforms often use rich pricing schemes, charging a two part tariff to buyers in the form of an annual fee and awards of "points" in proportion to the use of the card. This makes them a natural setting in which to analyze how multidimensional heterogeneity interacts with the non-linear pricing schemes that have been the focus of classical contract theory. We show that the platform's use of nonlinear pricing is disciplined by the elasticity of demand of infra-marginal buyers, as in Mussa and Rosen [1978], but also by the variance of the demand of marginal buyers, which captures the sorting effect.

We consider a platform and individuals on two sides of the market, retailers (R) and buyers (B). We follow Rochet and Tirole [2003] in assuming that individuals interact at random. Retailers have uni-dimensional types $\theta^R \sim f^R$ so there is no sorting of retailers, as was the case for advertisers in the previous application. Participating retailers obtain utility $u^R = \left(\theta^R - \phi^R\right) Q$, where ϕ^R is the fee per transaction charged to retailers and Q is the total demand for transactions by buyers.

Buyers have multidimensional types $\boldsymbol{\theta}^B \sim f^B$ which may account for their wealth, gender, impulsiveness, and other characteristics. The platform charges buyers a two-part tariff in the form of a yearly price P^B and a per transaction fee ϕ^B. Participating buyers obtain $u^B = N^R S(\phi^B; \boldsymbol{\theta}^B) - P^B$, where N^R is the share of retailers participating in the platform, and $S\left(\phi^B; \boldsymbol{\theta}^B\right) = \max_q v\left(q; \boldsymbol{\theta}^B\right) - \phi^B q$ is the transactional surplus of buyer $\boldsymbol{\theta}^B$, which is maximized at her demand for transactions $q(\phi^B; \boldsymbol{\theta})$. A crucial feature of this model is that the specification of buyer preferences allows us to use the envelope theorem to find $\frac{du^B}{d\phi^B} = -N^R q$, which will enable us to obtain a closed form version of the sorting effect.

Outside options are zero. For sides $s \in \{R, B\}$, the set of participants is $\Theta^s = \{\boldsymbol{\theta}^s : u^s \geq 0\}$, the set of marginal participants is $\partial\Theta^s = \{\boldsymbol{\theta}^s : u^s = 0\}$, the share of participants is N^s and the density of the margin is M^s. The relevant characteristics are the share of participating retailers $N^R = \int_{\Theta^R} f^R d\boldsymbol{\theta}^R$, and total demand of buyers $Q = \int_{\Theta^B} q f^B d\boldsymbol{\theta}^B$.

The platform incurs a fixed cost c for each of the QN^R executed transactions. We will think of ϕ^R and P^B as being used to determine the number of participants on each side and will focus on ϕ^B, the slope of the two-part tariff.

For convenience, we assign the transaction cost c in equal shares to retailers and buyers, whose values to society are $v^R = \left(\theta^R - \frac{c}{2}\right) Q + \lambda^{WN}$ and $v^B = N^R S + (\phi^B - \frac{c}{2}) q N^R + q\lambda^{WQ}$. Here, λ^{WN} and λ^{WQ} are the marginal values of N^R and Q.

From Equation (4), $\lambda^{WQ} = N^R\left(\mathbb{E}\left[\theta^R \mid \Theta^R\right] - \frac{c}{2}\right)$. There is no sorting of retailers because they have a single dimension of heterogeneity. From Equation (2) we obtain $\phi^B = c - \mathbb{E}\left[\theta^R \mid \Theta^R\right]$, as prescribed by Pigou [1912]. The covariance term disappears because, at the welfare maximizing optimum, all buyers have the same value: their marginal utility of transactions is ϕ^B and the marginal value of additional demand is equal to its cost.

Conversely, the values of individual retailers and buyers to a profit maximizer are

$$\pi^R = \left(\phi^R - \frac{c}{2}\right)Q + \lambda^{\Pi N} \qquad \text{and} \qquad \pi^B = \left(\phi^B - \frac{c}{2}\right)qN^R + P^B + q\lambda^{\Pi Q},$$

where $\lambda^{\Pi N}$ and $\lambda^{\Pi Q}$ are the marginal values of N^R and Q. From Equation (4), $\lambda^{\Pi Q} = N^R\left(\phi^R - \frac{c}{2}\right)$. From Equation (2), the optimal transaction fee to buyers satisfies, after some manipulation,

$$N^B\left(\mathbb{E}\left[q \mid \Theta^B\right] - \mathbb{E}\left[q \mid \partial\Theta^B\right]\right) + \left(\phi^B + \phi^R - c\right)\left(N^B\mathbb{E}\left[\frac{dq}{d\phi^B} \mid \Theta^B\right] - N^R M^B \mathrm{Var}\left(q \mid \partial\Theta^B\right)\right) = 0.$$

Increasing ϕ^B raises the revenue from the $N^B\mathbb{E}\left[q \mid \Theta^B\right]$ infra-marginal transactions but, to hold the number of buyers fixed, P^B changes proportionally to $\mathbb{E}\left[\frac{du^B}{d\phi^B} \mid \partial\Theta^B\right] = -N^R\mathbb{E}\left[q \mid \partial\Theta^B\right]$, by the envelope theorem. Moreover, increasing ϕ^B changes the demand for transactions in two ways. First, the N^B infra-marginal buyers reduce their demand on average by $\mathbb{E}\left[\frac{dq}{d\phi^B} \mid \Theta^B\right]$. Second, among marginal buyers, the sorting effect is proportional to $\mathrm{Cov}\left(\frac{\partial u^B}{\partial\phi^B}, q \mid \partial\Theta^B\right) = -N^R\mathrm{Var}(q \mid \partial\Theta^B)$, again by the envelope theorem. The sensitivity of each buyer to ϕ^B is proportional to her demand, as is her value. This is typical of price descrimination models, where individuals have different values precisely because they differ in preferences. The change in demand has a value to the platform of $\phi^B + \phi^R - c$ per unit. The negative sign in front of the sorting effect term shows that increasing ϕ^B repels valuable marginal buyers, which explains why ϕ^B is typically a subsidy in the form of airline miles or "points," rather than a fee.

We highlight several connections to the existing literature by expressing this condition as

$$\frac{\phi^B + \phi^R - c}{\phi^B} = \left(1 - \frac{\mathbb{E}\left[q \mid \partial\Theta^B\right]}{\mathbb{E}\left[q \mid \Theta^B\right]}\right)\left(\epsilon_X \frac{N^R\mathrm{Var}\left(q \mid \partial\Theta^B\right)}{\mathbb{E}\left[q \mid \partial\Theta^B\right]} + \epsilon_I\right)^{-1},$$

where $\epsilon_I = -\frac{\phi^B}{\mathbb{E}[q|\Theta^C]}\mathbb{E}\left[\frac{dq}{d\phi^B} \mid \Theta^B\right]$ is the average intensive or infra-marginal elasticity of individual demand and $\epsilon_X = M^B \mathbb{E}\left[q \mid \partial\Theta^B\right] \frac{\phi^B}{Q}$ is the elasticity of the demand due to "exiting" buyers.

If individuals are homogenous ex-ante, then $\mathbb{E}\left[q \mid \partial\Theta^C\right] = \mathbb{E}\left[q \mid \Theta^C\right]$ and we obtain $\phi^B + \phi^R = c$. The platform acts as a welfare maximizer towards buyers and imposes a Spence [1975] distortion to retailers, as in Bedre-Defolie and Calvano [2010].[26] If there are no externalities and marginal individuals have zero consumption, as in Mussa and Rosen [1978], then $\text{Var}(q \mid \partial\Theta) = \mathbb{E}\left[q \mid \partial\Theta^B\right] = \theta^R = 0$ and we recover the Wilson [1993] inverse elasticity formula $\left(\phi^B - c\right)/\phi^B = 1/\epsilon_I$. Rochet and Stole [2002] emphasize the role of the exiting elasticity ϵ_X in disciplining the platform's use of nonlinear pricing by showing that, as competition increases, transaction fees are driven to cost. The analysis above emphasizes that this disciplining effect of marginal buyers is weighted by the heterogeneity of their consumption, $\text{Var}\left(q \mid \partial\Theta^B\right)$. Rich heterogeneity reduces the platform's incentive to charge a mark-up on marginal transactions, driving transaction fees down to cost.

5.3 Imperfect Competition in Insurance

We extend the model of Section 2 to allow for imperfect competition, and extend the analysis of insurance provision of Einav and Finkelstein [2011] to allow for market power and for non-price instruments (coverage levels). This application illustrates how competition, rather than market power, generates the selection and cream-skimming distortions typically associated with asymmetric information.

We consider a duopoly of insurers, indexed by $d \in \{1, 2\}$, and assume a symmetric equilibrium where each insurer chooses the same levels of a premium P and coverage ρ. Individuals have types $\boldsymbol{\theta} \sim f$ which may account for their risk aversion, risk, search costs, outside options and preferences over insurers. Individual $\boldsymbol{\theta}$ obtains $u^d = s^d(\rho; \boldsymbol{\theta}) - P$ by purchasing from insurer d. Outside options are zero and individuals purchase from a single insurer if at all. Individual $\boldsymbol{\theta}$'s cost of provision is $c(\rho, \boldsymbol{\theta})$ irrespective of her insurer. Individual contributions to profit are $\pi = P - c$.

We assume that the distribution of (π, u^1, u^2) induced by f for any symmetric level of (P, ρ), is symmetric, so both insurers face the same residual demand curve

[26] The ratio $\mathbb{E}\left[q \mid \partial\Theta^B\right]/\mathbb{E}\left[q \mid \Theta^B\right]$ is a standard result for optimal two-part tariffs, discussed in Carlton and Perloff [1994] and Varian [1992]. Oi [1971] shows that homogeneity leads to the efficiency of two part tariffs in a one sided setting. See also Schmalensee [1981].

and therefore solve the same maximization problems.

Since the set of individuals purchasing from each insurer has the same size and composition, with some abuse of notation we will denote the set of users purchasing from each insurer by $\Theta = \left\{\boldsymbol{\theta} : u^d \geq \max\left\{u^{-d}, 0\right\}\right\}$ and the mass of these individuals by N. The exiting margin of each insurer contains those individuals indifferent between purchasing from that insurer and not purchasing at all: $\partial\Theta^X = \left\{\boldsymbol{\theta} : u^d = 0\right\}$ with density M^X. The switching margin of users indifferent between the two insurers is $\partial\Theta^S = \{\boldsymbol{\theta} : u^1 = u^2 \geq 0\}$.

A welfare maximizer views the industry as a single firm and, since total profit is invariant to which insurer covers each user, the switching margin is ignored. Given the symmetry assumptions above, the welfare maximizing choices for either insurer can be obtained from Equations (1) and (2) by considering values $s - c$ and considering only the exiting margin $\partial\Theta^X$. We obtain

$$P = \mathbb{E}\left[c \mid \partial\Theta^X\right] \qquad \text{and} \qquad N\mathbb{E}\left[\frac{\partial s}{\partial\rho} - \frac{\partial c}{\partial\rho} \mid \Theta\right] = M^X \mathrm{Cov}\left(\frac{\partial s}{\partial\rho}, c \mid \partial\Theta^X\right),$$

a result similar to that of Section 2.

A profit maximizer views the value of individual $\boldsymbol{\theta}$ as $P - c(\rho, \boldsymbol{\theta})$ and considers both the switching and exiting margins, $\partial\Theta^X \cup \partial\Theta^S$. From Equations (1) and (2), it chooses P and ρ such that

$$P - \underbrace{\frac{N}{M^X + M^S}}_{\text{Cournot}} = \underbrace{\mathbb{E}\left[c \mid \partial\Theta^X \cup \partial\Theta^S\right]}_{\text{Akerlof-Einav-Finkelstein}}$$

$$\underbrace{N\,\mathbb{E}\left[\frac{\partial s}{\partial\rho} \mid \partial\Theta^X \cup \partial\Theta^S\right]}_{\text{Spence}} - N\mathbb{E}\left[\frac{\partial c}{\partial\rho} \mid \Theta\right] = \underbrace{\left(M^X + M^S\right)\mathrm{Cov}\left(\frac{\partial s}{\partial\rho}, c \mid \partial\Theta^X \cup \partial\Theta^S\right)}_{\text{Rothschild-Stiglitz}}$$

A duopolist's profit maximizing conditions include the Cournot [1838] distortion of prices upwards and the Spence [1975] distortion of coverage catering to marginal individuals. However, competition introduces two additional distortions.

When setting its focal instrument (choosing the number of individuals to serve), a profit maximizing duopolist considers the cost of individuals on both the exiting and switching margins, whereas a welfare maximizer considers only the exiting margin. We hypothesize that individuals on the switching margin, since they are infra-marginal to the industry, are more similar to infra-marginal individuals than to individuals on the exiting margin. Therefore a profit maximizing duopolist distorts price away from marginal cost in the direction of average cost, in the spirit of Akerlof [1970] and Einav et al. [2010a].

When setting its non-focal instrument, a profit maximizer caters to the tastes of valuable individuals on both the exiting and switching margins. The incentive to poach valuable individuals from its competitor, without internalizing the externality this imposes, is a distortion similar to that described in Rothschild and Stiglitz [1976] and Akerlof [1976].

One can consider the effects of an increase in (differentiated Bertrand) competition as equivalent to an increase in the density of the switching margin, M^S. The optimality conditions above show that competition mitigates the Cournot distortion since the M^S is the numerator of that distortion term. It is also reasonable to think that greater competition mitigates the Spence distortion, since more weight is placed on the preferences of individuals on the switching margin, who can be conjectured to more closely resemble infra-marginal than marginal individuals.

However, an increase in competition increases the extent to which price responds to the cost of individuals on the switching margin. Since this increases the difference between $\mathbb{E}\left[c \mid \partial\Theta^X\right]$ and $\mathbb{E}\left[c \mid \partial\Theta^X \cup \partial\Theta^S\right]$, it exacerbates the Akerlof [1970]-Einav et al. [2010a]. Under perfect competition ($M^S \to \infty$) price are equated to average cost as in Einav et al. [2010a]. Similarly, the marginal incentive to cream-skim increases with the density of the switching margin, since M^S multiplies the covariance term that quantifies its benefit. Thus competition also exacerbates the Rothschild and Stiglitz [1976] distortion. In fact, the expression above makes clear that perfect competition ($M^S \to \infty$) always leads to market collapse, as it does in Rothschild and Stiglitz [1976], since the incentive to cream-skim is infinite in that case. This is in contrast to the Akerlof [1970]-Einav et al. [2010a] selection distortion, which need not lead to such a collapse.

This analysis raises important questions about traditional approaches to antitrust policy in industries where asymetric information is important. While competition tends to reduce the distortions on which the industrial organization literature has traditionally focused (the Cournot and Spence distortions), it can exacerbate the Akerlof-Einav-Finkelstein and the Rothschild-Stiglitz distortions. Therefore one potential benefit of a merger might be to reduce harms associated with these distortions.

6 Conclusion

We model a firm's choices of product characteristics when individuals differ in their preferences and in their values to the firm or to other individuals. By considering a finite numbers of product design instruments the problem is amenable to a simple price-theoretical analysis. We obtain necessary conditions for profit and welfare maximization in terms of aggregate market quantities and moments of the distribution of individual preferences and values. Namely, we characterize the power of an instrument to sort for valuable marginal individuals which is proportional to the covariance, within that set, between their preferences for the instrument and their value to the firm. In the general version of our model we allow for non-transferable utility, consumption externalities, cream-skimming distortions, adverse/advantageous selection, intensive margins of participation, non-linear pricing, third-degree discrimination, and imperfect competition. We apply our model to the study of a broadcast media platform, a credit card platform, and imperfectly competing insurers.

We leave for future work the modeling of imperfectly competitive markets where non-platform firms have asymmetric technologies, choose asymmetric product designs (for instance, to focus a market niche), or design multiple products. Another interesting extension is the consideration of externalities to unserved individuals, as in Segal [1999].

The issue of competition between platforms implies that the characteristics of each firm depends on the decisions taken by other firms and introduces multiplicity of equilibria among firms as well as individuals. White and Weyl [2011] propose a unique *insulated equilibrium* in which platforms adjust prices to keep the number of individuals on each side constant, but an extension of this result to our setting is left for future research. Another application of theoretical interest would be to imperfectly competitive settings where individuals can join multiple platforms (multi-home) and the contributions to characteristics by each individual depend on the subset of platforms she participates on, as in Ambrus and Reisinger [2006] and Athey et al. [2010].

We have allowed individuals to make discrete choices only regarding their participation, although including other discrete choices would blur the sharp distinction between the intensive and extensive effects. Additionally, the competitive markets we consider clear through individual choices, although one might also consider markets that clear based on firm choices over individuals, as in Gale and Shapley [1962].[27]

We also assume the existence of a (positive sales) market equilibrium, although asymmetric information can cause markets to shut down. Recent work like Hendren [2011] explores conditions on primitives that rationalize the non-existence of markets. An extension of these conditions to the settings described above might provide insight into the role played by specific dimensions of heterogeneity in market failure.

We focused exclusively on first-order conditions, but investigation of second-order conditions, as in Weyl and Fabinger [2011], can help quantify comparative statics and other global properties. It can also help approximate the level, rather than just the direction, of socially optimal policy starting from a sub-optimal private equilibrium, as in Jaffe and Weyl [2012]. A deeper analysis of second-order conditions is therefore another important direction for future research.

[27] Azevedo and Leshno [2012] and Azevedo [2011] obtain characterizations of such a setting and obtain results related to ours.

References

George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3):488–500, 1970.

George A. Akerlof. The economics of caste and of the rat race and other woeful tales. *Quarterly Journal of Economics*, 90(4):599–617, 1976.

Attila Ambrus and Markus Reisinger. Exlusive vs. overlapping viewers in media markets. http://epub.ub.uni-muenchen.de/1178/1/MWWB_Ambrus_Reisinger.pdf, 2006.

Simon P. Anderson and Stephen Coate. Market provision of broadcasting: A welfare analysis. *Review of Economic Studies*, 72(4):947–972, 2005.

Aloisio Araujo and Humberto Moreira. Adverse selection problems without the Spence-Mirrlees condition. *Journal of Economic Theory*, 145(3):1113 – 1141, 2010.

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

Susan Athey, Emilio Calvano, and Joshua Gans. The impact of the internet on advertising markets for news media. http://kuznets.harvard.edu/ athey/papers/Impact-Internet-Adv-Mkts-4-27-12.pdf, 2010.

Eduardo M. Azevedo. Imperfect competition in two-sided matching markets. *Games and Economic Behavior*, 2011.

Eduardo M. Azevedo and Jacob D. Leshno. A supply and demand framework for two-sided matching markets. http://www.people.fas.harvard.edu/ azevedo/papers/azevedo_jmp.pdf, 2012.

Özlem Bedre-Defolie and Emilio Calvano. Pricing payment cards. http://ssrn.com/paper=1522026, 2010.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

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

Patrick Bolton and Mathias Dewatripont. *Contract Theory*. The MIT Press, 2004.

Kate Bundorf, Jonathan Levin, and Neale Mahoney. Pricing and welfare in health plan choice. *American Economic Review*, Forthcoming.

Luis Cabral. Dynamic price competition with network effects. *Review of Economic Studies*, 78(1):83–111, 2011.

Bernard Caillaud and Bruno Jullien. Chicken and egg: Competition among intermediation service providers. *RAND Journal of Economics*, 34(2):309–328, 2003.

Estelle Cantillon and Pai-Ling Yin. Competition between exchanges: Lessons from the battle of the bund. http://ssrn.com/paper=1307513, 2008.

James H. Cardon and Igal Hendel. Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *RAND Journal of Economics*, 32(3):408–427, 2001.

Dennis W. Carlton and Jeffrey M. Perloff. *Modern industrial organization*. Harper Collins, New York, NY, 1994.

Raj Chetty. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(1):451–488, 2009.

Pierre-André Chiappori and Bernard Salanié. Testing for asymmetric information in insurance markets. *Journal of Political Economy*, 108(1):56–78, 2000.

Antoine A. Cournot. *Recherches sur les principes mathematiques de la theorie des richess*. Paris, 1838.

Avinash K. Dixit and Agnar Sandmo. Some simplified formulae for optimal income taxation. *The Scandinavian Journal of Economics*, pages 417–423, 1977.

Federico Echenique. Comparative statics by adaptive dynamics and the correspondence principle. *Econometrica*, 70(2):833–844, 2002.

Liran Einav and Amy Finkelstein. Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic Perspectives*, 25(1):115–138, 2011.

Liran Einav, Am Finkelstein, and Mark R. Cullen. Estimating welfare in insurance markets using variation in prices. *Quarterly Journal of Economics*, 125(3):877–921, 2010a.

Liran Einav, Amy Finkelstein, and Jonathan Levin. Beyond testing: Empirical models of insurance markets. *Annual Review of Economics*, 2(1):311–336, 2010b.

Liran Einav, Mark Jenkins, and Jonathan Levin. Contract pricing in consumer credit markets. *Econometrica*, Forthcoming.

Amy Finkelstein and Kathleen McGarry. Multiple dimensions of private information: Evidence from the long-term care insurance market. *American Economic Review*, 96(4):938–958, 2006.

Harley Flanders. Differentiation under the integral sign. *American Mathematical Monthly*, 80(6):615–627, 1973.

David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

Matthew Gentzkow and Jesse M. Shapiro. What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1):35–71, 2010.

Matthew Gentzkow, Jesse M. Shapiro, and Michael Sinkinson. Competition and ideological diversity: Historical evidence from us newspapers. http://faculty.chicagobooth.edu/jesse.shapiro/research/competition.pdf, 2011.

Igal Hendel. Estimating multiple-discrete choice models: An application to computerization returns. *The Review of Economic Studies*, 66(2):423, 1999.

Nathaniel Hendren. Private information and insurance rejections. http://econ-www.mit.edu/files/7488, 2011.

Sonia Jaffe and E. Glen Weyl. The first-order approach to merger analysis. http://ssrn.com/paper=1765024, 2012.

Michael L. Katz and Carl Shapiro. Network externalities, competition, and compatibility. *American Economic Review*, 75(3):424–440, 1985.

Robin S. Lee. Vertical integration and exclusivity in platform and two-sided markets. http://ssrn.com/paper=1022682, 2010.

Phillip Leslie. Price discrimination in broadway theater. *RAND Journal of Economics*, 35(3):520–541, 2004.

Michael J. Mazzeo. Product choice and oligopoly market structure. *RAND Journal of Economics*, 33(2):221–242, 2002.

R. Preston McAfee and John McMillan. Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, 46(2):335–354, 1988.

Paul Milgrom and Chris Shannon. Monotone comparative statics. *Econometrica*, 62 (1):157–180, 1994.

James A. Mirrlees. An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38(2):175–208, 1971.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317, 1978.

Chris Nosko. Competition and Quality Choice in the CPU Market. http://faculty.chicagobooth.edu/workshops/marketing/archive/winter11/nosko_jmp.pdf, 2010.

Walter Y. Oi. A disneyland dilemma: Two-part tariffs for a mickey mouse monopoly. *The Quarterly Journal of Economics*, 85(1):77–96, 1971.

Arthur C. Pigou. *Wealth and Welfare*. Macmillan and Co., Limited, 1912.

John G. Riley. Informational equilibrium. *Econometrica*, 47(2):331–359, 1979.

Jean C. Rochet and Jean Tirole. Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029, June 2003.

Jean-Charles Rochet and Philippe Choné. Ironing, sweeping, and multidimensional screening. *Econometrica*, 66(4):783–826, 1998.

Jean-Charles Rochet and Lars A. Stole. Nonlinear pricing with random participation. *Review of Economic Studies*, 69(1):277–311, 2002.

Jean-Charles Rochet and Lars A. Stole. The economics of multidimensional screening. In Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, editors, *Advances in Economics and Econometrics: Theory and Applications*, volume 1 of *Eighth World Congress*. Cambridge University Press, 2003.

Jeffrey Rohlfs. A theory of interdependent demand for a communications service. *Bell Journal of Economics and Management Science*, 5(1):16–37, 1974.

James N. Rosse. Estimating cost function parameters without using cost data: Illustrated methodology. *Econometrica*, 38(2):256–275, 1970.

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

Stephen Ryan and Catherine Tucker. Heterogeneity and the dynamics of technology adoption. *Quantitative Marketing and Economics*, Forthcoming.

Marc Rysman. Competition between networks: A study of the market for yellow pages. *Review of Economic Studies*, 71(2):483–512, 2004.

Marc Rysman. The economics of two-sided markets. *Journal of Economic Perspectives*, 23(3):125–143, 2009.

Irwin Sandberg. Global implicit function theorems. *Circuits and Systems, IEEE Transactions on*, 28(2):145–149, 1981.

Richard Schmalensee. Monopolistic two-part pricing arrangements. *The Bell Journal of Economics*, pages 445–466, 1981.

Ilya Segal. Contracting with externalities. *Quarterly Journal of Economics*, 114(2): 337–388, 1999.

A. Michael Spence. Job market signaling. *Quarterly Journal of Economics*, 87(3): 355, 1973.

A. Michael Spence. Monopoly, quality, and regulation. *Bell Journal of Economics*, 6 (2):417–429, 1975.

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

Amanda Starc. Insurer pricing and consumer welfare: Evidence from medigap. 2010.

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

Stanislav Uryas'ev. Derivatives of probability functions and integrals over sets given by inequalities. *Journal of Computational and Applied Mathematics*, 56(1-2):197–223, 1994.

Hal R. Varian. *Microeconomic analysis*, volume 2. Norton New York, 1992.

William F. Weingarten. Empirical calibration of optimal innovation policy in smartphone application markets. http://ssrn.com/paper=1855823, 2011.

E. Glen Weyl. A price theory of multi-sided platforms. *American Economic Review*, 100(4):1642–1672, 2010.

E. Glen Weyl and Michal Fabinger. A restatement of the theory of monopoly. http://ssrn.com/paper=1851326, 2011.

E. Glen Weyl and Jean Tirole. Market power screens willingness-to-pay. http://ssrn.com/paper=1659449, 2011.

Alexander White and E. Glen Weyl. Insulated platform competition. http://ssrn.com/abstract=1694317, 2011.

Robert B. Wilson. *Nonlinear Pricing.* Oxford University Press, 1993.

A Appendix A

Proof of Lemma 2:

Differentiating G yields

$$\frac{dG}{d\rho} = \int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})\geq 0} \frac{\partial}{\partial \rho} g(\rho;\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} + \int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})=0} g(\rho;\boldsymbol{\theta}) f(\boldsymbol{\theta}) \left(\nabla_\rho \boldsymbol{\theta} \cdot \boldsymbol{n}\right) d\tau.$$

The gradient $\nabla_\rho \boldsymbol{\theta}$ is the velocity with which the boundary changes with ρ, $\boldsymbol{n}$ is the unit vector normal to the surface defined by $\{\boldsymbol{\theta} : u(\rho;\boldsymbol{\theta}) = 0\}$, "$\cdot$" is the vector dot product, and τ is the (scalar) area element of that surface. Then the outward velocity of expansion of the boundary at each point is given by

$$\nabla_\rho \boldsymbol{\theta} \cdot \boldsymbol{n} = \nabla_\rho \boldsymbol{\theta} \cdot \frac{\nabla_{\boldsymbol{\theta}} u}{\|\nabla_{\boldsymbol{\theta}} u\|} = \frac{\partial u}{\partial \rho} \frac{\nabla_u \boldsymbol{\theta} \cdot \nabla_{\boldsymbol{\theta}} u}{\|\nabla_{\boldsymbol{\theta}} u\|} = \frac{\partial u}{\partial \rho} \frac{1}{\left\| \frac{du}{d\boldsymbol{\theta}} \right\|}.$$

We define the expectation operators on the interior and boundary of Θ as

$$N = \int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})\geq 0} f(\boldsymbol{\theta}) d\boldsymbol{\theta} \text{ and } M = \int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})=0} \frac{1}{\|\nabla_{\boldsymbol{\theta}} u\|} f(\boldsymbol{\theta}) d\sigma$$

and obtain

$$\begin{aligned}
\frac{dG}{d\rho} &= \int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})\geq 0} f(\boldsymbol{\theta}) d\boldsymbol{\theta} \frac{\int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})\geq 0} \frac{\partial}{\partial \rho} g(\rho;\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})\geq 0} f(\boldsymbol{\theta}) d\boldsymbol{\theta}} \\
&\quad + \int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})=0} \frac{1}{\|\nabla_{\boldsymbol{\theta}} u\|} f(\boldsymbol{\theta}) d\sigma \frac{\int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})=0} g(\rho;\boldsymbol{\theta}) f(\boldsymbol{\theta}) \frac{\partial u}{\partial \rho} \frac{1}{\left\| \frac{du}{d\boldsymbol{\theta}} \right\|} d\sigma}{\int_{\boldsymbol{\theta}:u(\rho;\boldsymbol{\theta})=0} \frac{1}{\|\nabla_{\boldsymbol{\theta}} u\|} f(\boldsymbol{\theta}) d\sigma} \\
&= N\mathbb{E}\left[\frac{\partial}{\partial \rho} g(\rho;\boldsymbol{\theta}) \mid \boldsymbol{\theta} : u(\rho;\boldsymbol{\theta}) \geq 0\right] + M\mathbb{E}\left[g(\rho;\boldsymbol{\theta}) \frac{\partial u}{\partial \rho} \mid \boldsymbol{\theta} : u(\rho;\boldsymbol{\theta}) = 0\right].
\end{aligned}$$

B Appendix B

Proof of Proposition 1: We differentiate the Lagrangian

$$\mathcal{L}^V = \int_\Theta v f - \sum_{i=1}^{\mathcal{K}} K^i \lambda^{Vi}$$

with respect to the focal instrument $\rho^\star$ to obtain

$$\begin{aligned}
\frac{d\mathcal{L}^H}{d\rho^\star} &= N\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right] + M\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} v \mid \partial\Theta\right] \\
0 &= N\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right] + M\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]\mathbb{E}\left[v \mid \partial\Theta\right] + M\mathrm{Cov}\left(\frac{\partial u}{\partial \rho^\star}, v \mid \partial\Theta\right) \\
0 &= \mathbb{E}\left[v \mid \partial\Theta\right] + \frac{N\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right]}{M\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]} + \frac{\mathrm{Cov}\left(\frac{\partial u}{\partial \rho^\star}, v \mid \partial\Theta\right)}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]}.
\end{aligned}$$

Proof of Proposition 2: First, we establish that

$$0 = \frac{dN}{d\rho^l} = \frac{\partial N}{\partial \rho^l} + \frac{\partial N}{\partial \rho^\star}\frac{\partial \rho^\star}{\partial \rho^l} = M\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right] + M\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]\frac{\partial \rho^\star}{\partial \rho^l},$$

thereby justifying out definition of $\frac{\partial \rho^\star}{\partial \rho^l} = -\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right] / \mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]$. The definition of $\frac{\partial \rho^\star}{\partial K^j}$ is similar.

We differentiate the Lagrangian above with respect to ρ^l to obtain

$$\frac{d\mathcal{L}^H}{d\rho^l} = \mathbb{E}\left[v \mid \partial\Theta\right] + \frac{N\mathbb{E}\left[\frac{dv}{d\rho^l} \mid \Theta\right]}{M\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right]} + \frac{\mathrm{Cov}\left(\frac{\partial u}{\partial \rho^l} v \mid \partial\Theta\right)}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right]} = 0,$$

and use the first-order condition with respect to $\rho^\star$ above to eliminate $\mathbb{E}\left[v \mid \partial\Theta\right]$ in this equation. We obtain

$$\frac{N\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right]}{M\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]} + \frac{\mathrm{Cov}\left(\frac{\partial u}{\partial \rho^\star}, v \mid \partial\Theta\right)}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]} = \frac{N\mathbb{E}\left[\frac{dv}{d\rho^l} \mid \Theta\right]}{M\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right]} + \frac{\mathrm{Cov}\left(\frac{\partial u}{\partial \rho^l}, v \mid \partial\Theta\right)}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right]}$$

$$\begin{aligned}
0 = {} & \frac{N}{M}\mathbb{E}\left[\frac{dv}{d\rho^l} \mid \Theta\right] - \frac{N}{M}\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right]\frac{\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right]}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]} \\
& + \mathrm{Cov}\left(\frac{\partial u}{\partial \rho^l}, v \mid \partial\Theta\right) - \mathrm{Cov}\left(\frac{\partial u}{\partial \rho^\star}, v \mid \partial\Theta\right)\frac{\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right]}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]}.
\end{aligned}$$

We obtain the result by using the definition of $\frac{\partial \rho^\star}{\partial \rho^l} = -\mathbb{E}\left[\frac{\partial u}{\partial \rho^l} \mid \partial\Theta\right] / \mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]$.

Proof of Proposition 3: We differentiate the Lagrangian with respect to K^j to obtain

$$\begin{aligned}
\frac{d\mathcal{L}^H}{dK^j} &= N\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right] + M\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} v \mid \partial\Theta\right] - \lambda^{Vj} \\
0 &= \mathbb{E}\left[v \mid \partial\Theta\right] + \frac{N\mathbb{E}\left[\frac{dv}{dK^j} \mid \Theta\right] - \lambda^{Vj}}{M\mathbb{E}\left[\frac{\partial u}{\partial K^j} \mid \partial\Theta\right]} + \frac{\mathrm{Cov}\left(\frac{\partial u}{\partial K^j} v \mid \partial\Theta\right)}{\mathbb{E}\left[\frac{\partial u}{\partial K^j} \mid \partial\Theta\right]}.
\end{aligned}$$

Using the optimality condition with respect to $\rho^\star$above, we obtain

$$\begin{aligned}
0 &= \frac{N}{M}\mathbb{E}\left[\frac{dv}{dK^j} \mid \Theta\right] - \frac{N}{M}\mathbb{E}\left[\frac{dv}{d\rho^\star} \mid \Theta\right] \frac{\mathbb{E}\left[\frac{\partial u}{\partial K^j} \mid \partial\Theta\right]}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]} \\
&\quad + \mathrm{Cov}\left(\frac{\partial u}{\partial K^j}, v \mid \partial\Theta\right) - \mathrm{Cov}\left(\frac{\partial u}{\partial \rho^\star}, v \mid \partial\Theta\right) \frac{\mathbb{E}\left[\frac{\partial u}{\partial K^j} \mid \partial\Theta\right]}{\mathbb{E}\left[\frac{\partial u}{\partial \rho^\star} \mid \partial\Theta\right]}.
\end{aligned}$$

We then obtain the result by using the definition of $\frac{\partial \rho^\star}{\partial K^j}$.