Multidimensional Product Design[*]

André Veiga[†] and E. Glen Weyl[‡]

Abstract

Products like radio shows and insurance contracts are designed by firms to sort for the most valuable users. An empirically relevant model of this process requires users to be heterogeneous along multiple dimensions of their preferences and their values, but existing models do not allow for multidimensional heterogeneity or require restrictive assumptions to be tractable. We show that a simple price theoretic analysis is possible when user heterogeneity is of high dimension relative to the firm's design instruments. We obtain necessary conditions for profit and welfare maximization in terms of moments of the distribution of user heterogeneity, where the power of an instrument to sort for valuable users is proportional to the *covariance, within the set of marginal users, between the value of users and their marginal utility for the instrument.* Our model allows for non-transferable utility, consumption externalities, cream-skimming distortions, adverse/advantageous selection, non-linear pricing, third-degree discrimination and imperfect competition. We discuss applications to broadcast media, the credit card industry and imperfect competition in insurance provision.

Keywords: Platforms, screening, heterogeneity, non-linear pricing, cream-skimming, intensive v. extensive margin

JEL Classication Codes: D42, D43, D82, I13, L15

[†]Toulouse School of Economics (TSE), Manufacture de Tabacs, 21 allées de Brienne, 31000 Toulouse, France; andreveiga@gmail.com.

[‡]Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago IL, 60637; weyl@uchicago.edu.

1 Introduction

Radio stations famously introduced melodramatic "soap operas" that catered to the tastes of housewives who controlled family purchase decisions. Since housewives were particularly valuable to advertisers of soap, they were also valuable to the radio stations. In this and many other industries, the heterogeneity of user *values* is a key feature that firms take into account when designing their products. A model of this screening problem must allow for users to have multiple dimensions of private information, and in particular to be heterogeneous in their preferences and their values to the firm. However, existing models do not allow for multidimensional heterogeneity, or require restrictive assumptions on preferences and technologies to remain tractable, which makes estimation challenging. We develop a formulation of this problem that is simple, general and tractable. We allow users to have heterogeneity of potentially high dimension but restrict firms to using a small number of finite dimensional product design instruments. These assumptions "smooth" the model, allowing us to use standard differential techniques without few additional assumptions. We derive profit and welfare maximizing necessary conditions that are functions of the moments of the distribution of user heterogeneity, and are thus amenable to measurement. In particular, we show that the power of an instrument to sort for valuable users is proportional to the *covariance, within the set of marginal users, between the value of users and their marginal utility for the instrument.*

We illustrate our approach and main results in Section 2 by considering a simple example of health insurance provision, a product famously designed to attract the healthiest (most valuable) patients. We model a monopoly insurer's choice of two design instruments, a uniform participation price and a uniform level of coverage. Potential patients may have multiple dimensions of heterogeneity such as health status and risk aversion and, in particular, patients differ arbitrarily in their willingness to pay for coverage (preferences) and their cost of provision (value to the firm), both increasing in the generosity of coverage. As described by Cournot [1838], the welfare maximizing price equals marginal cost, while the profit maximizing price equates marginal revenue to marginal cost. As described by Spence [1975], the welfare maximizing level of coverage depends on the benefit of coverage to the average covered patient, but the profit maximizing level depends only on the preferences of marginal patients.

The welfare and profit maximizing conditions also include a novel term, the *sorting* effect, which is one of the main contributions of our paper. It captures the extent to which, holding fixed the number of covered patients, those marginal patients to whom coverage is most attractive are also patients who are particularly costly to cover. When this relationship between preference and cost is positive, increasing coverage changes the composition of covered patients towards those most costly to cover, and thereby increases the insurer's cost. This sorting effect is intuitively proportional to the covariance, within marginal patients, between the marginal utility for coverage and the cost of provision. The presence of sorting relies on the heterogeneity of users along multiple dimensions, since homogeneity or unidimensional heterogeneity both imply that any set of marginal patents of finite density would be homogenous, causing the covariance term to vanish. Thus we obtain simple conditions for the welfare maximizing level of insurance coverage and price, describe the distortions associated with profit maximization and formalize several intuitions present in the literature, while simultaneously accomodating the multiple dimensions of private information known to be relevant in the insurance setting.

In Section 3, we solve the general version of our model, assuming only that utilities and values are sufficiently smooth and that user heterogeneity and utilities are distributed smoothly and with finite moments. The model is tractable even while allowing for any number of finite-dimensional product design instruments, non-transferable utility and intensive participation decisions on the part of users. We include heterogeneous consumption externalities between users in a straightforward way, by reformulating the firm's problem as one of constrained optimization. We derive necessary optimality conditions for each of the firm's product design instruments, contrast the cases of profit- and welfare-maximization and highlight the relevant distortions. We then extend the model to allow for for third degree discrimination between subsets of users ("sides") and to the case of imperfect competition in the absence of consumption externalities. By considering imperfect competition, we identify additional distortions caused by profit maximization while avoiding the result of market collapse common in the literature on competition under asymmetric information.

Our results stem from several departures form the existing literature, as we discuss in Section 4. In our model, firms determine the number of users purchasing their product, as in Cournot [1838], but users have assymetric information about their value, as in Akerlof [1970], and firm use non-price instruments to screen users,

3

as in Spence [1975]. This makes our analysis close in spirit to that of Rothschild and Stiglitz [1976] but, in contrast to that paper and subsequent single-dimensional screening models like Mussa and Rosen [1978], we consider users with multidimensional heterogeneity. Unlike the multidimensional screening models like [Rochet and Choné, 1998], we consider finite-dimensional instruments alongside heterogeneity of potentially high dimension, which simplifies the analysis. We also consider imperfectly competitive settings where firms use multiple instruments, unlike the model of perfect competition between price-choosing firms of [Einav and Finkelstein, 2011].

In Section 5, we apply our model to the analysis of three industries to highlight the role of ingredients like multiple "sides", non-transferable utility, non-linear pricing, adverse/advantageous selection and cream-skimming. First, we consider a radio broadcaster as an example of a multi-sided market where utility is non-transferable on one side. Our results illustrate how the the absence of transfers affects the radio station's choice of content. Second, we explore a credit card issuer's use of non-linear pricing, as a generalization of the Rochet and Tirole [2003] model of two-sided markets, and show how the heterogeneity of marginal card-holders and the elasticity of consumption of infra-marginal card-holders disciplines the use of non-linear pricing. Our third application illustrates the cream-skimming and selection distortions in an imperfectly competitive insurance market where insurers choose prices and coverage, as an extension of Einav and Finkelstein [2011].[1]

Section 6 mentions promising avenues for future works and discusses how our model can be combined with other common approaches to yield additional economic insight. These include the structural identification of unobserved economic parameters, the estimation of locally welfare improving interventions, the estimation of local approximations to optimal industrial policy, qualitative and quantitative comparative statics, and the analysis of market collapse. Section 7 concludes.

---

[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 users may have multiple dimensions of private information (like expected cost and risk preferences)." Our paper aims to provide such a characterization.

2 A Simple Example

In this section, we use a simple example to illustrate our approach and the main result of our paper, the sorting effect.[2]

We consider a monopoly health insurer and a continuum of potential patients with mass normalized to 1. First, the insurer chooses the levels of two instruments: a uniform price P and a uniform level of coverage ρ. Then, each patient makes a binary decision about whether to purchase insurance. Patients have multi-dimensional "types" $\boldsymbol{\theta} \in \mathbb{R}^{\mathcal{T}}$, a $\mathcal{T}$-vector of individual characteristics which is each patient's private information and will determine her preferences and values. In this example, $\boldsymbol{\theta}$ might include the patient's genetic profile, health status for multiple conditions, lifestyle, risk aversion and education, so patient heterogeneity is of high dimension relative to the insurer's instruments. Types are distributed in the population according the density $f(\boldsymbol{\theta})$, which is common knowledge. Patient $\boldsymbol{\theta}$ obtains utility $u = v(\rho; \boldsymbol{\theta}) - P$ from participating, and it costs the platform $c(\rho; \boldsymbol{\theta})$ to cover her. Thus, patients differ in their preferences for coverage and in their values to the insurer. We normalize the outside option of patients to be zero, so the set of covered patients is $\Theta = \{\boldsymbol{\theta} : v(\rho; \boldsymbol{\theta}) \geq P\}$ and the set of marginal patients is $\partial\Theta = \{\boldsymbol{\theta} : v(\rho; \boldsymbol{\theta}) = P\}$. We will omit functional arguments for clarity.

We begin by considering an insurer maximizing welfare, the difference between value created and cost of provision $W = \int_{\Theta} (v - c) f$. The integral W can be differentiated using Leibniz's Rule to yield[3]

$$\frac{dW}{dP} = \int_{\partial\Theta} \frac{du}{dP} (v - c) f = \left(\int_{\partial\Theta} f \right) \mathbb{E} \left[-(v - c) \mid \partial\Theta \right] = -M\mathbb{E} \left[P - c \mid \partial\Theta \right] = 0.$$

The welfare maximizing number of users is reached when price equals the cost of covering the average marginal patient, or $P = \mathbb{E}[c \mid \partial\Theta]$. Increasing price repels marginal patients proportionally to their marginal valuation du/dP, which is homogeneous and equal to -1 by quasilinearity. This changes the number of covered patients but merely redistributes surplus from infra-marginal patients to the insurer, so the Leibniz Rule effect on the interior of the integral vanishes and dW/dP boils down to the effect on the margin. The density of marginal patients $M = \int_{\partial\Theta} f$ quantifies the responsiveness

---

[2]We assume all functions are smooth and also assume the existence and uniqueness of the solution.

[3]See Flanders [1973] for details on the multidimensional version of Leibniz's Rule.

of the set of participants to changes in the instruments since it is marginal patients that change their participation decisions in response to such changes.[4] Finally, from the definition of $\partial\Theta$, we know that $v = P$ within that set.

Dividing the expression above by M expresses optimality in terms of the number of patients, rather than the price level. This is intuitive and in line with the existing literature, so throughout the paper we will think of the firm as using one instrument to select the optimal number of users.

The insurer then uses coverage ρ to increase the value from covered patients and sort for valuable marginal patients, while the number of covered patients is held fixed by the price. Using Leibniz's Rule, the optimal level of coverage must satisfy

$$
\begin{aligned}
\frac{dW}{d\rho} &= \int_\Theta \left(\frac{dv}{d\rho} - \frac{dc}{d\rho} \right) f + \int_{\partial\Theta} \frac{du}{d\rho} (v - c) f \\
&= \underbrace{N\mathbb{E}\left[\frac{dv}{d\rho} \mid \Theta \right] - N\mathbb{E}\left[\frac{dc}{d\rho} \mid \Theta \right]}_{\text{value from covered patients}} - \underbrace{M\text{Cov}\left(\frac{dv}{d\rho}, c \mid \partial\Theta \right)}_{\text{sorting of marginal patients}} = 0
\end{aligned}
$$

The first two terms capture the *intensive* effect, or the marginal value of coverage to covered patients net of costs, where the share of covered patients is $N = \int_\Theta f$. The second term captures the *extensive effect* or the *sorting* of marginal patients by coverage. When coverage is increased, marginal patients are attracted heterogeneously and proportionally to their marginal valuations for coverage $du/d\rho$. Each attracted patient also contributes heterogeneously to social welfare, in the amount $v - c$. This is captured by the integral over the margin $\partial\Theta$, which can be re-written as $M\mathbb{E}\left[(dv/d\rho)(v - c) \mid \partial\Theta \right]$. Then, using the definition of covariance, we can then re-write this expectation of a product as

$$
\text{Cov}\left(\frac{dv}{d\rho}, v - c \mid \partial\Theta \right) + \mathbb{E}\left[\frac{dv}{d\rho} \mid \partial\Theta \right] \cdot \underbrace{\mathbb{E}\left[v - c \mid \partial\Theta \right]}_{=0}.
$$

The last term vanishes because the FOC for the optimal number of covered patients is exactly $\mathbb{E}\left[v - c \mid \partial\Theta \right] = 0$. Moreover, since v is constant (equal to P) within the marginal set, it drops out of the covariance term, yielding the result.

---

[4]Notice M is a "surface" integral in $\mathcal{T} - 1$ dimensions. It is the multi-dimensional analogue of the density (f) evaluated at the marginal user, in models with uni-dimensional heterogeneity. M is a derivative or density, rather than a mass of patients.

The sorting effect captures the sorting cost of coverage, or the extent to which the marginal patients most attracted by coverage are also particularly costly to cover. When this is true, increasing coverage while holding the number of covered patients fixed changes the composition of covered patients towards more costly patients which implies a social cost. This sorting effect is intuitively proportional to the *covariance, among marginal patients, between preferences for coverage and cost of provision.* Notice that multidimensional heterogeneity is essential to characterize sorting, since the covariance term would vanish if marginal patients were homogeneous.

The simple optimality condition for the socially optimal level of coverage takes into account multiple dimensions of private information, shown by Finkelstein and McGarry [2006] to be relevant in this setting. The characterization depends only of aggregate market quantities and therefore seems amenable to estimation. It shows, for instance, that even a welfare maximizer reduces coverage to repel costly patients since it is not socially efficient to cover those whose willingness to pay is less than their cost, as emphasized in Einav and Finkelstein [2011]. Moreover, the sorting term formalizes the intuition of Bundorf et al. [Forthcoming], that distortions away from the first-best level of coverage are heavily determined by heterogeneity in health status but less so by heterogeneity in risk preference. The reason is that health status affects both preferences and cost, while risk preference affects only preferences. Therefore heterogeneity in risk preference has no effect on the covariance term that characterizes sorting.

Consider now the maximization of profit $\Pi = \int_\Theta (P - c) f$. Using Leibniz's Rule we obtain

$$\frac{d\Pi}{dP} = \int_\Theta f + \int_{\partial\Theta} \frac{du}{dP} (P - c) f = N - M \mathbb{E}\left[P - c \mid \partial\Theta\right] = \underbrace{P - \frac{N}{M}}_{\text{Marginal Revenue}} - \mathbb{E}\left[c \mid \partial\Theta\right] = 0.$$

The profit maximizing number of covered patients equates marginal revenue to marginal cost. The term N/M is the wedge between price (the value of participation to patients) and marginal revenue (the value of participation to the insurer), and constitutes the Cournot [1838] distortion.[5]

The profit maximizing level of coverage satisfies

---

[5]The term $-N/M = N/(dN/dP)$ is commonly expressed as $(dP/dQ) Q$, where Q is the quantity supplied by a monopoly . The price elasticity of demand is PM/N.

$$\begin{aligned}
\frac{d\Pi}{d\rho} &= \int_{\Theta} -\frac{dc}{d\rho}f + \int_{\partial\Theta} \frac{du}{d\rho}\left(P-c\right)f = \\
&= -N\mathbb{E}\left[\frac{dc}{d\rho} \mid \Theta\right] + \underbrace{N\mathbb{E}\left[\frac{dv}{d\rho} \mid \partial\Theta\right]}_{\text{Spence term}} - \underbrace{M\mathrm{Cov}\left(\frac{dv}{d\rho}, c \mid \partial\Theta\right)}_{\text{sorting of marginal patients}} = 0
\end{aligned}$$

The integral over $\partial\Theta$ captures the effect of coverage on marginal participants and can be re-written as $M\mathbb{E}\left[(dv/d\rho)\left(P-c\right) \mid \partial\Theta\right]$. This expectation of a product can then be re-written as

$$\mathrm{Cov}\left(\frac{dv}{d\rho}, P-c \mid \partial\Theta\right) + \mathbb{E}\left[\frac{dv}{d\rho} \mid \partial\Theta\right]\underbrace{\mathbb{E}\left[v-c \mid \partial\Theta\right]}_{=N/M}$$

The result follows from noticing P is constant and, from the FOC with respect to price, $\mathbb{E}\left[v-c \mid \partial\Theta\right] = N/M$.

A profit maximizer does not consider the benefit of coverage to covered patients. Coverage is valuable to a profit maximizer only to the extent that an increase in coverage leads to an upward adjustment in price to hold fixed the number of patients, which increases profit. The effect of coverage on the number of patients is determined by the preferences of marginal patients for coverage, so the adjustment of price that holds their number fixed must be proportional to it as well. Thus a profit maximizer chooses coverage catering to the taste for coverage of marginal patients ($\mathbb{E}\left[dv/d\rho \mid \partial\Theta\right]$) rather than average user, because any increase in their surplus allows the platform to raise prices to all covered patients while still maintaining the number of patients fixed. This is the distortion described by Spence [1975]. Notice that, unlike the Cournot distortion which always leads a profit maximizer to under-provide its product, the Spence distortion is not signed, so the relative preferences of marginal and infra-marginal patients determine whether coverage is over- or under-provided. Also, since marginal patients have zero utility, sorting has a first order effect only on profit so the social and private incentives to sort are the same in this case.

3 The Model

The model of Section 2 contained a number of assumptions that we now relax in order to describe the greatest generality in which our approach can be applied. Some readers might find it useful to skip the current section on a first reading. In this section, the firm uses any finite number of instruments and users have heterogeneous preferences over each instrument. Users have a more flexible specification of preferences that need not be monotonic, quasilinear or transferable.Users also generate heterogeneous consumption externalities towards other participating users, and therefore will refer to the firm as a *platform*. Each user's preferences over these externalities may also be heterogeneous, and users choose their intensity of participation in the platform. In Subsection 3.1 we extend the model to allow for third degree discrimination between "sides", and in Subsection 3.2 we consider the case of imperfect competition between non-platforms.

We consider a monopoly platform that chooses a vector of $\mathcal{R}$ *instruments*, $\boldsymbol{\rho} = \left(\rho^1, ..., \rho^l, ...\right) \in \mathbb{R}^{\mathcal{R}}$, with components indexed by $l \in \{1, 2, .., \mathcal{R}\}$. We denote $\rho^1 = \rho^\star$ as the platform's *focal instrument*, which the platform uses to determine the number of participating users. While the focal instrument is unrestricted in principle, the natural choice is an instrument that transfer utility from users to the firm, like the uniform participation price of Section 2.[6]

We consider also a continuum of potential users with mass normalized to 1. Each user has a type, a $\mathcal{T}$-vector of characteristics $\boldsymbol{\theta} \in \mathbb{R}^{\mathcal{T}}$, which is each user's private information.[7] Types are distributed in the population according to a probability density function $f(\boldsymbol{\theta})$, which is common knowledge. If user $\boldsymbol{\theta}$ participates, she obtains utility $u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta})$ and contributes $\pi(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta})$ to the platform's profit, where $\boldsymbol{K}$ is a vector of platform characteristics discussed below. Thus, such a user would contribute $w(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) = u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) + \pi(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta})$ to social welfare. In Section 2, user $\boldsymbol{\theta}$'s contribution to welfare was $v(\rho; \boldsymbol{\theta}) - c(\rho\boldsymbol{\theta})$, and to profit was $P - c(\rho; \boldsymbol{\theta})$.

Each user decides whether or not to join the platform (we will model additional intensive margins of participation below). We normalize outside options to zero so the set of participating users is $\Theta \equiv \{\boldsymbol{\theta} : u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) \geq 0\}$, and the set of marginal users

---

[6]It is also useful if user preferences for the focal instrument are homogeneous and/or have a constant sign.

[7]The firm has no private information.

is $\partial\Theta \equiv \{\boldsymbol{\theta} : u(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) = 0\}$.[8] The share of participating users is $N \equiv \int_{\Theta} f(\boldsymbol{\theta}) d\boldsymbol{\theta}$ and the density of marginal users is $M \equiv \int_{\partial\Theta} f(\boldsymbol{\theta}) d\boldsymbol{\theta}$, where $\boldsymbol{t}$ is the area element on the surface of Θ.

We allow users to generate consumption externalities towards other participating users by considering $\mathbf{K} = (K^1, ..., K^i, ...) \in \mathbb{R}^{\mathcal{K}}$, a vector of platform *characteristics* with components indexed by $i \in \{1, 2, ..., \mathcal{K}\}$. $\boldsymbol{K}$ is determined by the platform's participants. If user $\boldsymbol{\theta}$ participates, she makes a contribution $k^i(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta})$ to characteristic K^i. We then define each characteristic as the integral of the individual contributions of all participating users to a given characteristic, so $K^i \equiv \int_{\Theta} k^i(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}, \forall i$. For instance, if k^i is a user's wealth, K^i is the total wealth of participants.[9]

We model an intensive margin of user participation by allowing individual contributions to characteristics to depend on the levels of instruments and characteristics, which we illustrate in Subsection 5.2. For instance, users can make heterogeneous contributions to the total demand experienced by a platform, with individual contributions depending on the platform's choice of instruments (like price). Thus $\boldsymbol{K}$ models both consumption externalities which enter the preferences of users, and intensive participation decisions that may enter the payoff function of the firm. As is common in the literature, we assume that user expectations of $\boldsymbol{K}$ are correct in equilibrium.

We define profit and social welfare respectively as $\Pi = \int_{\Theta} \pi(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$ and $W = \int_{\Theta} w(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$.[10] In order to tackle the profit and welfare maximization problems simultaneously, and to highlight the differences between the two, we denote $h(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) \in \{\pi(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}), w(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta})\}$ and $H = \int_{\Theta} h(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}$. The platform's problem is therefore to choose $\boldsymbol{\rho}$ to maximize H.

We assume the existence of a solution, and discuss its uniqueness in Section 6. Any solution to the firm's problem must be interior because $\boldsymbol{\rho} \in \mathbb{R}^{\mathcal{R}}$. We make the following additional technical assumptions.

Assumption 1: f is twice continuously differentiable, atomless and with finite moments.

---

[8]This assumes that instruments and characteristics have no effect on non-participants. See Segal [1999] for an analysis when this is not the case.

[9]By combining multiple characteristics, it is possible to make the preferences of users depend on the average or variance of a characteristic.

[10]Assuming that profit is a linear aggregation of individual contributions is convenient for exposition and is easily relaxed by considering a general smooth profit function $\Pi(\boldsymbol{\rho}, \boldsymbol{K})$, although this complicates the exposition.

Assumption 2: Any function $g \in \{u, \pi, k^1, ..., k^j\}$ is twice continuously differentiable in all arguments and has bounded derivatives.

Assumption 3: For any $(\boldsymbol{\rho}, \boldsymbol{K})$, $\|du/d\boldsymbol{\theta}\|$ is bounded away from zero.

Assumption 4: There is some finite density of marginal users ($M \neq 0$).

These assumptions ensure that any integral $G \in \{W, \Pi, K^1, ..., K^j\}$ is well defined and differentiable in $\boldsymbol{\rho}$. For a proof of differentiability, see Uryas' ev [1994].[11]

For any x and y, we use $x_y = \partial x/\partial y$ to denote partial derivatives. We use $\widetilde{x} = \mathbb{E}[x \mid \partial\Theta]$ to denote expectations conditional on the set of marginal users, and use $\overline{x} = \mathbb{E}[x \mid \Theta]$ to denote expectations conditional on participation. We use the following additional notation in our first result.

- $\sigma\langle x, y\rangle \equiv \mathrm{Cov}(x, y \mid \boldsymbol{\theta} \in \partial\Theta)$ is the covariance, among marginal users, between x and y.

- $\widetilde{\boldsymbol{k}} = (\mathbb{E}[k^1 \mid \partial\Theta], ..., \mathbb{E}[k^i \mid \partial\Theta], ...)$ is the $\mathcal{K}$-vector of average contributions to characteristics by marginal users.

- $\overline{\boldsymbol{k}_{\rho^l}} = (\mathbb{E}[dk^1/d\rho^l \mid \Theta], ..., \mathbb{E}[dk^i/d\rho^l \mid \Theta], ...)^\intercal$ is the $\mathcal{K}$-vector of effects of any instrument ρ^l on average user contributions to characteristics (including the focal instrument $\rho^\star$).

- $\boldsymbol{\lambda^H} = (\lambda^{H1}, ..., \lambda^{H\mathcal{K}})^\intercal$ is the $\mathcal{K}$-vector of marginal values of characteristics in the maximization of H.

Proposition 1. *A necessary condition for the optimal choice of the focal instrument $\rho^\star$ is*

$$\underbrace{\widetilde{\pi} + \widetilde{\boldsymbol{k}}\boldsymbol{\lambda^H}}_{\text{marginal value}} + \underbrace{\sigma\left\langle \frac{u_{\rho^\star}}{\widetilde{u_{\rho^\star}}}, \pi + \boldsymbol{k}\boldsymbol{\lambda^H} \right\rangle}_{\text{heterogeneous incidence}} + \underbrace{\frac{1}{M\widetilde{u_{\rho^\star}}}N\left(\overline{h_{\rho^\star}} + \overline{\boldsymbol{k}_{\rho^\star}}\boldsymbol{\lambda^H}\right)}_{\text{intensive effect}} = 0. \tag{1}$$

Proof. See the appendix in Section 8. □

From Equation 1, the focal instrument $\rho^\star$ is set to that the value of an additional marginal user attracted by $\rho^\star$ is zero. The average marginal user makes direct a

---

[11]Intuitively, the differentiation of the interior is clear, and the conditions above ensure that the region of integration changes smoothly with $\boldsymbol{\rho}$: u having bounded derivative ensures the change is not "explosive", while $\|du/d\boldsymbol{\theta}\|$ being bounded below ensures that the density of marginal users is not infinite.

contribution to profit of $\widetilde{\pi}$, her classical marginal value. She makes no contribution to user surplus since marginal users have zero utility, so this term is the same in the profit and welfare-maximizing problems. The average marginal user also makes contributions to all characteristics in an amount captured by $\widetilde{\boldsymbol{k}}$, the vector of average contributions of marginal users. The marginal value to the platform of each characteristic is captured by the vector $\boldsymbol{\lambda}^H$, which we solve for below, so the total value of contributions $\widetilde{\boldsymbol{k}}$ is the dot product $\widetilde{\boldsymbol{k}}\boldsymbol{\lambda}^H$. The platform differ in the case of welfare and profit maximization because $\boldsymbol{\lambda}^H$ differs between the two cases. In the example of Section 2, the classical marginal value was $\widetilde{\pi} = P - \widetilde{c}$, and there were no consumption externalities so $\widetilde{\boldsymbol{k}}\boldsymbol{\lambda}^H = \boldsymbol{0}$.

The discussion of the previous paragraph assumes all marginal users are attracted homogeneously by the focal instrument $\rho^\star$, so the value of an additional user is simply the average value of marginal users, $\widetilde{\pi} + \widetilde{\boldsymbol{k}}\boldsymbol{\lambda}^H$. When users are heterogeneously attracted by $\rho^\star$, the optimality condition accounts for the extent to which those users most attracted by $\rho^\star$ are also particularly valuable, which is captured by the covariance between (normalized) preferences for the focal instrument and value to the platform, $\sigma\left\langle u_{\rho^\star}/\widetilde{u_{\rho^\star}}, \pi + \boldsymbol{k}\boldsymbol{\lambda}^H\right\rangle$. This term vanishes in Section 2 because user preferences over prices are homogenous.

The term $(1/M\widetilde{u_{\rho^\star}})\, N\left(\overline{h_{\rho^\star}} + \overline{\boldsymbol{k}_{\rho^\star}}\boldsymbol{\lambda}^H\right)$ captures the cost of using instrument $\rho^\star$ and generalizes the Cournot [1838] term of Section 2. It reflects the change in value of infra-marginal users when $\rho^\star$ changes by the amount needed to attract a marginal user. The required change in $\rho^\star$ is inversely proportional to the effectiveness of $\rho^\star$ in attracting marginal users $(1/M\widetilde{u_{\rho^\star}})$ and induces in all participating users (N) an average change in value of $\overline{h_{\rho^\star}} + \overline{\boldsymbol{k}_{\rho^\star}}\boldsymbol{\lambda}^H$. This is composed of the change in classical value $(\overline{h_{\rho^\star}})$ and the vector of changes in contributions to characteristics $(\overline{\boldsymbol{k}_{\rho^\star}})$ weighted by the marginal values of characteristics $(\boldsymbol{\lambda}^H)$. If there are no consumption externalities, as in Section 2, or if contributions are invariant to the focal instrument (which seems common), then $\overline{\boldsymbol{k}_{\rho^\star}}\boldsymbol{\lambda}^H = \boldsymbol{0}$. The more common distortions of profit maximization are illustrated by the term $\overline{h_{\rho^\star}}$. Under the reasonable assumption that $\rho^\star$ transfers utility from users to the firm (perhaps imperfectly), then $\overline{\pi_{\rho^\star}} > \overline{\pi_{\rho^\star}} + \overline{u_{\rho^\star}}$ so a profit maximizer will set its focal instrument above the socially optimal level. For the monopolist of Section 2, the cost of lowering price to attract an additional user is the lower revenue from covered patients. In that case, $-\overline{u_{\rho^\star}} = \overline{\pi_{\rho^\star}} = 1$, so the cost of using price (the Cournot distortion) becomes $(1/M(-1))\cdot N\cdot 1 = N/M$.

On the other hand, the welfare maximizer of Section 2 takes into account that utility is transferable so changes in price cause only welfare neutral redistribution among infra-marginal users. In that case, $\overline{w_{\rho^\star}} = 0$ and this term vanishes. Thus a welfare maximizer equates price to marginal cost under transferable utility, because there is no social cost of increasing price beyond its impact on participation.

Proposition 2. *A necessary condition for the optimal choice of each non-focal instrument ρ^l is*

$$\underbrace{M\sigma\left\langle u_{\rho^l} - \frac{\widetilde{u_{\rho^l}}}{\widetilde{u_{\rho^\star}}}u_{\rho^\star}, \pi + \boldsymbol{k}\boldsymbol{\lambda^H}\right\rangle}_{\text{sorting effect}} + \underbrace{N\left(\overline{h_{\rho^l}} + \overline{\boldsymbol{k}_{\rho^l}}\boldsymbol{\lambda^H} - \frac{\widetilde{u_{\rho^l}}}{\widetilde{u_{\rho^\star}}}\left(\overline{h_{\rho^\star}} + \overline{\boldsymbol{k}_{\rho^\star}}\boldsymbol{\lambda^H}\right)\right)}_{\text{intensive effect}} = 0. \quad (2)$$

Proof. See the appendix in Section 8. □

The platform uses its non-focal instruments ρ^l to sort the set of participating users towards those with greater value and to increase the value of infra-marginal users. Since the number of users is determined and held fixed by $\rho^\star$, any change in ρ^l should be thought of as being accompanied by a compensating adjustment of $\rho^\star$ that holds fixed the number of participants. The effect of ρ^l on the number of users is determined by the effectiveness of ρ^l in attracting marginal users, so the compensating change in $\rho^\star$ must also be proportional to $\widetilde{u_{\rho^l}}$ and is inversely proportional to the effectiveness of $\rho^\star$ is repelling marginal users $(-\widetilde{u_{\rho^\star}})$. The adjustment is therefore proportional to $-\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}$.

The power of ρ^l to sort for valuable marginal users is quantified by the density of marginal users (M) multiplied by $\sigma\left\langle u_{\rho^l} - \left(\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}\right)u_{\rho^\star}, \pi + \boldsymbol{k}\boldsymbol{\lambda^H}\right\rangle$, the covariance between the total value of marginal users $(\pi + \boldsymbol{k}\boldsymbol{\lambda^H})$ and the preferences for the non-focal instrument (u_{ρ^l}) taking also into account the marginal preferences for the focal instrument $(u_{\rho^\star})$ to the extent that the focal instrument adjusts $(-\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}})$. When users are homogeneous in their preferences for the focal instrument, $-\left(\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}\right)u_{\rho^\star}$ drops out of the covariance, as in Section 2.

A change in ρ^l also induces an average change in the classical values of infra-marginal users of $\overline{h_{\rho^l}} - \left(\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}\right)\overline{h_{\rho^\star}}$, which generalizes the the Spence [1975] distortion. Let us begin by assuming $\rho^\star$ perfectly transfers utility between users and the platform. A welfare maximizer $(h = w)$ ignores the adjustment of the focal instrument since it causes only socially neutral redistribution $(\overline{w_{\rho^\star}} = 0)$, so the term

13

$-\left(\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}\right)\overline{w_{\rho^\star}}$ vanishes. Thus, a welfare maximizer considers only the direct effects of ρ^l on the welfare of the average user $(\overline{w_{\rho^l}})$. In Section 2, the welfare maximizer considers the total effect of coverage on surplus and costs, $\overline{w_{\rho^l}} = \overline{v_{\rho^l}} - \overline{c_{\rho^l}}$.

However, a profit maximizer $(h = \pi)$ considers the effect of ρ^l on the contributions to profit of infra-marginal users $(\overline{\pi_{\rho^l}})$, but not on user surplus. It also considers the effect of ρ^l in causing an adjustment in $\rho^\star$, to the extent that this adjustment increases the platform's profit from infra-marginal users, $-\left(\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}\right)\overline{\pi_{\rho^\star}}$. Since the adjustment of $\rho^\star$ depends on the preferences of marginal users for ρ^l, a profit maximizing platform chooses ρ^l catering to the tastes of marginal users, which constitutes the Spence [1975] distortion. Notice the importance of the signs of $\widetilde{u_{\rho^\star}}$ and $\overline{\pi_{\rho^\star}}$: a profit maximizer will only positively consider the preferences of marginal users for ρ^l when the implied adjustment of $\rho^\star$ increases profits, which need not happen for a non-price focal instrument, as we illustrate in Subsection 5.1. In Section 2, this becomes $-\left(\widetilde{v_{\rho^l}}/\left(-1\right)\right)1 = \widetilde{v_{\rho^l}}$.

The term $N\left(\overline{\boldsymbol{k_{\rho^l}}} - \left(\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}\right)\overline{\boldsymbol{k_{\rho^\star}}}\right)\boldsymbol{\lambda^H}$ captures the effect of a ($\rho^\star$-compensated) change in ρ^l on the platform values of infra-marginal users. The logic of this term is similar to what is described above for classical values, although a change in ρ^l affects contributions to multiple characteristics, hence the vectors $\overline{\boldsymbol{k_{\rho^l}}}$ and $\overline{\boldsymbol{k_{\rho^\star}}}$, which multiply the vector of marginal values of characteristics, $\boldsymbol{\lambda^H}$.

Our third result requires one additional piece of notation:

- $\overline{\boldsymbol{k_{K^j}}} = \left(\overline{k_{K^j}^1}, ..., \overline{k_{K^j}^i}, ...\right)^\mathsf{T} = \left(\mathbb{E}\left[dk^1/dK^j \mid \Theta\right], ..., \mathbb{E}\left[dk^i/dK^j \mid \Theta\right], ...\right)^\mathsf{T}$ is the $\mathcal{K}$-vector of average effects of characteristic K^j on the average contributions to all characteristics by infra-marginal users.

Proposition 3. *The vector of marginal values of characteristics* $\boldsymbol{\lambda^H}$ *solves the system of* $\mathcal{K}$ *equations of the form*

$$\underbrace{M\sigma\left\langle u_{K^j} - \frac{\widetilde{u_{K^j}}}{\widetilde{u_{\rho^\star}}}u_{\rho^\star}, \pi + \boldsymbol{k}\boldsymbol{\lambda^H}\right\rangle}_{\textit{sorting effect}} + \underbrace{N\left(\overline{h_{K^j}} + \overline{\boldsymbol{k_{K^j}}}\boldsymbol{\lambda^H} - \frac{\widetilde{u_{K^j}}}{\widetilde{u_{\rho^\star}}}\left(\overline{h_{\rho^\star}} + \overline{\boldsymbol{k_{\rho^\star}}}\boldsymbol{\lambda^H}\right)\right)}_{\textit{intensive effect}} = \lambda^{Hj}.$$

(3)

for $j \in \{1, ..., \mathcal{K}\}$.

Proof. See the appendix in Section 8. $\qquad\square$

The marginal value of characteristic j (λ^{Hj}) is defined using the recursive Equation 3 (since λ^{Hj} is on the RHS and is also a part of $\boldsymbol{\lambda}^H$). This stems from the fact that a change in the levels of characteristics has a self-reinforcing (or self-defeating) feedback effect on characteristics themselves. Moreover, a change in one characteristic has an effect on all other characteristics, and each of these effects then produces a secondary effects on all characteristics proportional to the first, and so forth. The marginal value of characteristics is therefore the solution to the system of equations represented above.

The logic of Equation 3 is, perhaps surprisingly, analogous to that of Equation 2. The main difference is that instruments can be directly decided by the platforms, whereas characteristics are only indirectly influenced. This is clear from the Right Hand Sides (RHS) of Equations 2 and 3, which are are shadow values of instrument ρ^l and characteristic K^j (0 and λ^{Hj}, respectively).

Beyond this, there is little difference. As in Equation 2, a ($\rho^\star$-compensated) change in characteristics K^j induces, an average change in classical value of $\overline{h_{K^j}} - (\widetilde{u_{K^j}}/\widetilde{u_{\rho^\star}})\,\overline{h_{\rho^\star}}$, an average change in platform value of $\left(\overline{\boldsymbol{k_{K^j}}} - (\widetilde{u_{K^j}}/\widetilde{u_{\rho^\star}})\,\overline{\boldsymbol{k_{\rho^\star}}}\right)\boldsymbol{\lambda^H}$ and sorts for valuable marginal users proportionally to $\sigma\left\langle u_{K^j} - (\widetilde{u_{K^j}}/\widetilde{u_{\rho^\star}})\,u_{\rho^\star}, \pi + \boldsymbol{k}\boldsymbol{\lambda^H}\right\rangle$. The conditions shows why consumption externalities produce Spence distortions even when the platform uses only a price instruments, as in Weyl [2010]. With multiple instruments, the platform chooses each non-focal instrument taking into account its feedback effect on the focal instrument, which depends on the tastes of marginal users for the non-focal instrument. When there are consumption externalities, the platform chooses the focal instrument taking into account its effect on characteristics and the subsequent feedback effect on the focal instrument itself, which depends on the tastes of marginal users for characteristics.

3.1 Multiple Sides

With additional notation, the model can explicitly accommodate cases where platforms third-degree discriminate between "sides". We require only that, for each side, there is a focal instrument affecting only users on that side which can be used to determined the optimal number of users.

We consider $\mathcal{S}$ sides indexed by $s \in \{1, 2, ..., \mathcal{S}\}$. Users on side s have types $\boldsymbol{\theta}^s \in \mathbb{R}^{\mathcal{T}^s}$, with $\boldsymbol{\theta}^s \sim f^s(\boldsymbol{\theta}^s)$. Let $\boldsymbol{\rho}^s$ be the vector containing all non-focal instruments

and the focal instrument on side s, and let $\boldsymbol{\rho}$ be the vector containing all instruments. If user $\boldsymbol{\theta}^s$ participates, she obtains $u^s\left(\boldsymbol{\rho}^s, \boldsymbol{K}, \boldsymbol{\theta}^s\right)$ and increases the platform's profit by $\pi^s\left(\boldsymbol{\rho}^s, \boldsymbol{K}, \boldsymbol{\theta}^s\right)$, where $\boldsymbol{K}$ is the vector of all characteristics. Outside options are zero, the set of participating users on side s is $\Theta^s = \{\boldsymbol{\theta}^s : u^s \geq 0\}$ and the set of marginal users on that side is $\partial\Theta^s = \{\boldsymbol{\theta}^s : u^s = 0\}$. Let $\mathbf{K} = (K^1, ..., K^i, ...) \in \mathbb{R}^{\mathcal{K}}$ be the vector of all characteristics, generated by any side. Without loss of generality, we assume each characteristic is generated by a single side.[12] We refer to the side generating characteristic i as side i, and denote by $k^i\left(\boldsymbol{\rho}^i, \boldsymbol{K}, \boldsymbol{\theta}^i\right)$ the contribution to k^i of user $\boldsymbol{\theta}^i$ on side i. We can then define $K^i = \int_{\Theta^i} k^i f^i$. We define N^s, M^s, w^s, h^s in a similar way. Then problem can then be stated as the choice of $(\boldsymbol{\rho}, \boldsymbol{K})$ that maximizes $H = \sum_{s=1}^{\mathcal{S}} \int_{\Theta^s} h^s(\boldsymbol{\rho}, \boldsymbol{K}; \boldsymbol{\theta}^s)$ subject to the constraints $K^i = \int_{\Theta^i} k^i f^i, \forall i$. The solution can then be characterized as follows.

Proposition 4. *For a multi-sided platform, a necessary condition for the optimal choice of the focal instrument $\rho^{s\star}$ on side s is*

$$\underbrace{\widetilde{\pi^s} + \widetilde{\boldsymbol{k}^s}\boldsymbol{\lambda^{Hs}}}_{\text{marginal value}} + \sigma^s \underbrace{\left\{\frac{u_{\rho^{s\star}}^s}{u_{\rho^{s\star}}^s}, \pi^s + \boldsymbol{k}^s\boldsymbol{\lambda}^s\right\}}_{\text{heterogeneous incidence}} + \underbrace{\frac{1}{M^s\widetilde{u_{\rho^{s\star}}^s}}N^s\left(\overline{h_{\rho^{s\star}}^s} + \overline{\boldsymbol{k}_{\boldsymbol{\rho}^{s\star}}^s}\boldsymbol{\lambda^{Hs}}\right)}_{\text{intensive effect}} = 0 \quad (4)$$

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

$$\sum_{s=1}^{\mathcal{S}} \underbrace{M^s\sigma^s\left\{u_{\rho^l}^s - \frac{\widetilde{u_{\rho^l}^s}}{\widetilde{u_{\rho^1}^s}}u_{\rho^{s\star}}^s, \pi^s + \boldsymbol{k}^s\boldsymbol{\lambda}^s\right\}}_{\text{sorting}} + \underbrace{N^s\left(\overline{h_{\rho^l}^s} + \overline{\boldsymbol{k}_{\boldsymbol{\rho}^l}^s}\boldsymbol{\lambda^{Hs}} - \frac{\widetilde{u_{\rho^l}^s}}{\widetilde{u_{\rho^\star}^s}}\left(\overline{h_{\rho^{s\star}}^s} + \overline{\boldsymbol{k}_{\boldsymbol{\rho}^{s\star}}^s}\boldsymbol{\lambda^{Hs}}\right)\right)}_{\text{intensive effect}} = 0$$

$$(5)$$

The marginal value of each characteristic j (λ^{Hj}) can be obtained by solving the following system of $\mathcal{K}$ equations:

$$\sum_{s=1}^{\mathcal{S}} \underbrace{M^s\sigma^s\left\{u_{K^j}^s - \frac{\widetilde{u_{K^j}^s}}{\widetilde{u_{\rho^1}^s}}u_{\rho^{s\star}}^s, \pi^s + \boldsymbol{k}^s\boldsymbol{\lambda}^s\right\}}_{\text{sorting}} + \underbrace{N^s\left(\overline{h_{K^j}^s} + \overline{\boldsymbol{k}_{\boldsymbol{K}^j}^s}\boldsymbol{\lambda^{Hs}} - \frac{\widetilde{u_{K^j}^s}}{\widetilde{u_{\rho^1}^s}}\left(\overline{h_{\rho^{s\star}}^s} + \overline{\boldsymbol{k}_{\boldsymbol{\rho}^{s\star}}^s}\boldsymbol{\lambda^{Hs}}\right)\right)}_{\text{intensive effect}} = \lambda^{Hj}$$

$$(6)$$

---

[12] Any preferences depending on the sum of contributions from multiple sides can be re-defined appropriately as depending on the sum of the separate contributions, for instance.

Since $\rho^{s\star}$ affects only side s, the FOC with respect to this focal instrument takes exactly the form of Equation 1. The FOC with respect to any non-focal instrument ρ^l, while holding fixed the number of users on each side s by means of the appropriate focal instrument $\rho^{s\star}$, takes a form similar to that of Equation 2, where one must simply add the effect of ρ^l on the classical and platform values generated by all sides. The marginal value of characteristics can be obtained from a system of equations, each of which takes a form similar to the scalar version of Equation 3. Thus all the intuitions from the one-sided case extend to the case of multiple sides when one accounts for the effect on all sides of each non-focal instrument and each characteristic.[13]

3.2 Imperfect Competition

We now extend the model to a simple case of imperfect competition between non-platforms. The crucial insight here is that an oligopolist faces two "margins:" the *exiting margin* of users indifferent between the oligopolist's product and no product at all, and the *switching margin* of users indifferent between the oligopolist's product and that of a competition. Under reasonable assumptions, changes along the switching margin are socially neutral and therefore ignored by a welfare maximizer, but not by a profit maximizer.

For simplicity, we consider only a duopoly with firms indexed by $d \in \{A, B\}$, where $-d$ refers to the competitor of d. Let d chooses a vector of instruments $\boldsymbol{\rho}^d$, and let $\rho^{d\star}$ be its focal instrument. There is a mass 1 of users with types $\boldsymbol{\theta} \in \mathbb{R}^{\mathcal{T}}$, distributed $\boldsymbol{\theta} \sim f$. If user $\boldsymbol{\theta}$ purchases from d, she obtains $u\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right)$ and makes a contribution of $\pi\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right)$ to the profit of d. Then $w\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right) = u\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right) + \pi\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right)$ is $\boldsymbol{\theta}$'s contribution to social welfare. Users purchase from a single firm (if at all) and choose the one that gives them highest utility. The set of users purchasing from d is $\Theta^d = \left\{\boldsymbol{\theta} : u\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right) \geq \text{Max}\left\{u\left(\boldsymbol{\rho}^{-d}; \boldsymbol{\theta}\right), 0\right\}\right\}$. The exiting margin of d is $\partial\Theta^{Xd} = \left\{\boldsymbol{\theta} : u\left(\boldsymbol{\rho}^{-d}; \boldsymbol{\theta}\right) \geq u\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right) = 0\right\}$ and the switching margin (of both firms) is $\partial\Theta^S = \left\{\boldsymbol{\theta} : u\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right) = u\left(\boldsymbol{\rho}^{-d}; \boldsymbol{\theta}\right) \geq 0\right\}$.[14] Let N^d be the share of users who purchase from d, and let the density of its margins be $M^d = M^{Xd} + M^S = \int_{\partial\Theta^{Xd}} f + \int_{\partial\Theta^S} f$. The profit of d is $\Pi^d = \int_{\Theta^d} \pi\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right) f$ and welfare is $W = \sum_{d=1}^{2} \int_{\Theta^d} w\left(\boldsymbol{\rho}^d; \boldsymbol{\theta}\right) f$.

Thus set of users joining each firm depends on the choices of both firms, a firm's instruments affect only the surplus of users purchasing from that firm, and firms have

---

[13]Expressing Equation 6 in its vector-valued form would require the use of tensors of order 3.

[14]The sets $\partial\Theta^{Xd}$ and $\partial\Theta^S$ are disjoint apart from a set of measure zero in $\mathcal{T} - 2$ dimensions.

the same production function. For simplicity, we make the additional assumption of a symmetrical equilibrium where both firms choose the same level of their instruments $(\boldsymbol{\rho}^d = \boldsymbol{\rho}^{-d} = \boldsymbol{\rho})$. Then, movement of users along the switching margin has no effect on user surplus or on industry profits. This yields a particularly simple result: a welfare maximizer will ignore all effects occurring along the switching margin as socially neutral, whereas a profit maximizer will consider both margins.

In the following result, we omit the superscript d for clarity. To consider jointly the cases of welfare and profit maximization, we index the margins by $m \in \{X, X+S\}$ and we denote $h \in \{\pi, w\}$ as above. Then, $\widetilde{x^m} = \mathbb{E}[x \mid \partial\Theta^m]$ and $\sigma^m \langle x, y \rangle = \mathrm{Cov}(x, y \mid \partial\Theta^m)$.

Proposition 5. *For a non-platform duopolist in a symmetrical equilibrium, a necessary condition for the optimal level of the focal instrument $\rho^\star$ is*

$$\underbrace{\widetilde{\pi^m}}_{marginal\ value} + \underbrace{\sigma^m \left\langle \frac{u_{\rho^\star}}{\widetilde{u_{\rho^\star}}}, \pi \right\rangle}_{heterogenous\ incidence} + \underbrace{\frac{N}{M^m \widetilde{u_{\rho^\star}^m}} \overline{h_{\rho^\star}}}_{intensive\ effect} = 0 \tag{7}$$

and a necessary condition for the optimal level of each non-focal instrument ρ^l is

$$\underbrace{M^m \sigma^m \left\langle u_{\rho^l} - \frac{\widetilde{u_{\rho^l}}}{\widetilde{u_{\rho^\star}}} u_{\rho^\star}, \pi \right\rangle}_{sorting\ effect} + \underbrace{N \left(\overline{h_{\rho^l}} - \frac{\widetilde{u_{\rho^l}^m}}{\widetilde{u_{\rho^\star}^m}} \overline{h_{\rho^\star}} \right)}_{intensive\ effect} = 0 \tag{8}$$

where a profit maximizer considers $m = S + X$ and $h = \pi$, whereas a welfare maximizer considers $s = X$ and $h = w$.

There are two crucial differences between the setting with imperfect competition and the previous settings. First, in Equation 7, a profit maximizer $(m = X + S)$ considers the marginal revenue of attracting users everywhere along his margin $(\widetilde{\pi^{X+S}})$ whereas a welfare maximizer $(m = X)$ considers only revenue originating from the exiting margin $(\widetilde{\pi^X})$. By making the focal instrument excessively responsive to the value of users infra-marginal to the industry as a whole, competition introduces a *selection* distortion in the spirit of Akerlof [1970].

In Equation 8, the sorting term $\sigma^m \left\langle u_{\rho^l} - \left(\widetilde{u_{\rho^l}}/\widetilde{u_{\rho^\star}}\right) u_{\rho^\star}, \pi \right\rangle$ also differs between the two cases. A profit maximizer uses each non-focal instrument to sort for valuable marginal users on its entire margin $(m = X + S)$ so it does not internalize the extent

to which the profit it gains is being poached from its competitor. Competition then leads profit maximizers to use non-focal instruments to *cream-skim* valuable users from competitors, as in Rothschild and Stiglitz [1976]. We delay a more detailed explanation of these results to the more concrete setting of competition in insurance provision of Subsection 5.3.

3.3 Second Order Conditions and Multiplicity

The analysis above abstracted from second order derivatives. Determining plausible conditions (on primitives or otherwise) that would ensure that the results above describe a global maximum are beyond the scope of this paper. Without assumptions on second derivatives, our model predicts how an equilibrium reacts to a small exogenous shock and, under the conditions described by Milgrom and Shannon [1994], a global shift in one term of a first-order derivative implies a global shift in the optimal level of an action for a monopolist. Combining our model with assumptions on the signs of second derivatives, it is possible to obtain qualitative comparative statics, as done in Hale et al. [1999]. It would be useful to find, for this setting, conditions similar to those found for a single-product monopoly by Weyl and Fabinger [2011], who show that the measurement of second order derivatives allows for quantitative measurements of (or bounds on) comparatives statics. If distortions are sufficiently small, measurements of the gradient and curvature at equilibrium may even be informative about the level (not just the direction) of socially optimal policy, as discussed in Jaffe and Weyl [2012].[15] These results are left for future research.

The presence of consumption externalities tends to produce multiple equilibria, since the decision of each user depends on her expectations about the decisions of other users, as pointed out by Rohlfs [1974].[16] However, our analysis assumes the a platform can choose its preferred equilibrium. Competition between platforms is even more complex since the characteristics of each firm depend on the decisions taken by all firms and on user coordination between platforms. In a multi-sided monopoly setting with quasilinear preferences and homogeneous values, Weyl [2010] obtains uniqueness by having the platform's prices be contingent on the number of users on every side, thereby allowing the platform to make the number of users on each

---

[15]Quantitative results are sensible only if the stability conditions of [Samuelson, 1941] are satisfied.
[16]See also Katz and Shapiro [1985] and Caillaud and Jullien [2003]

side invariant to ("insulated from") changes in user expectations.[17] In an imperfectly competitive (but otherwise similar) setting, White and Weyl [2011] propose a unique *insulated equilibrium* in which firms adjust prices to keep the number of users on each side constant. These techniques do not apply immediately to our setting where users differ in their preferences for instruments and in their values. There are intuitive conditions under which it is possible to extend these results to our setting, although it would lead us too far astray from the the main flow of our work. Necessary and sufficient conditions can be found in Sandberg [1981].[18] Intuitively, a platform must have as many effective instruments as there are characteristics, instruments must have sufficiently independent effects on characteristics 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 illuminate connections between economics effects stemming from a diverse set of models. We pay special attention to six areas of the literature: classic product design, classic contract theory, multidimensional screening, empirical work in industrial organization, recent empirically motivated research on markets with asymmetric information, and recent price theory papers. We use the following subsections to discuss how our work relates to each of these literatures and how the assumptions typically made in each field determine which effects manifest themselves. At the end of this section, Table 1 summarizes these connections.

One assumption common to these fields is the transferability of utility. Firms often have no access to instruments which perfectly transfer utility from users, as argue by by Anderson and Coate [2005] in the context of broadcast media and White [2008] in the context of online search engines. Our model relaxes the transferability assumption and characterizes the role of the wedge between user utility and firm profit, as we have discussed above and illustrate in Subsection 5.1.

---

[17]This can be thought of as a reduced-form model of dynamic pricing, as argued by Cabral [2011].

[18]Let $\hat{K}$ be user expectations of characteristics and K be their realizations. Let $K = \kappa\left(\rho, \hat{K}\right)$ and let instruments ρ be contingent on expectations. Assume a desired feasible equilibrium where $K = K^\star$ and $\rho = \rho^\star$. This can be implemented uniquely when $\kappa\left(\rho, \hat{K}\right) = K^\star$ has a unique global implicit function solution $\rho = \rho^\star\left(\hat{K}\right)$. Sandberg [1981] shows when this is the case.

4.1 Classic product design

The classical treatment of product design is Spence [1975]'s analysis of a quality-choosing monopoly, which allows for users with arbitrary preference heterogeneity. The main result of this paper, the Spence distortion, was described in Sections 2 and 3. Weyl [2010] and White and Weyl [2011] show that the same principle holds for price-setting multi-sided platforms, where the number of users on each side plays the role of quality.

This literature typically assumes homogeneity of user values (to the firm or to other users), and assumes that users make discrete rather than intensive decisions about their purchases.[19] By relaxing these assumptions, our model shows why the sorting effect is absent in models where users preferences or user values are homogenous, and allows us to account for the sorting effect in settings richer than those previously considered. We described in Section 2 the relationship between the covariance term we obtain and some existing literature, and we further illustrate the importance of heterogeneous user values in Subsection 5.1, which models a radio broadcaster of melodramatic soap operas faced with users with heterogenous preferences and heterogeneous exogenously fixed values. Allowing for an intensive margin of participation lets the firm affect the values of users, as we illustrate in Subsection 5.2 where a credit card issuer's choice of a two-part tariff determines the heterogeneity of user usages of the card.

4.2 Classical contract theory

The classic contract theoretical literature, surveyed by Bolton and Dewatripont [2004], focuses on the incentive schedules chosen by firms faced with users with private information. The focus of this literature has been the case where individuals have a single dimension of private information. This implies several restrictions, which we avoid by allowing for multidimensional heterogeneity.

A single dimension of private information implies that either there is a single marginal user (usually in monopoly models like Stiglitz [1977]) or that a positive mass of users is marginal (usually in perfectly competitive models like Rothschild and Stiglitz [1976]). In the first case the sorting effect is zero because there is no

---

[19]The two issues are linked because users purchasing different amounts would generally have different values to the firm.

heterogeneity among marginal users. In the second case, the incentive to cream-skim valuable users from competitors is infinitely large, causing equilibrium to typically not exist, as discussed by Riley [1979]. Also, when a single parameter captures multiple aspects of incentives, it is unfeasible to distinguish important effects in a unified model. For instance, one must typically choose between modeling the effect of nonlinear pricing on the intensive participation decisions of infra-marginal users, 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 analyze jointly with the intensive participation effects of non-linear pricing and the.[20] We use the modeling of non-linear pricing in the credit card industry of Subsection 5.2, and of competition in insurance of Subsection 5.3, to illustrate how our model reproduces and generalizes several results typical of this literature.

4.3   Multidimensional screening

The literature on multidimensional screening, surveyed in [Armstrong and Rochet, 1999], addresses many of the concerns above by enriching user heterogeneity to multiple dimensions and allowing firms to use multiple infinite-dimensional instruments (nonlinear price schedules) to screen users. This literature allows for endogenous heterogeneity in user values since the firm's price schedules induces users to purchase different amounts. Classical models are [McAfee and McMillan, 1988], [Armstrong, 1996] and Rochet and Choné [1998]. However, an analytic treatment of these models is generally challenging. Results tend to focus on conditions under which full separation of types is possible and desirable, they tend to rely on parametric assumptions on preferences for tractability, and closed-form solutions are not always possible. The models therefore tend to lack a general characterization, comparative statics and intuitive optimality conditions.

Our model allows the firm to use any finite number of instruments, but not instruments of infinite dimension. This restriction allows us to consider design instruments

---

[20]Note that, when the incentives of users (their preferences) and the incentives of the firm (user 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's instruments to attract valuable users based on user preferences. In our model, this condition has a natural analog in the sign of the covariance term we describe, which quantifies the (local) power of instruments to attract valuable users based on their preferences.

other than nonlinear pricing, such as different aspects of product quality, but also nonlinear price tariffs with a finite number of parts.[21] Considering finite instruments also allows for the use of simple differential techniques to characterize optimality conditions, while simultaneously allowing for more flexible specifications of preferences and for straightforward extensions to the settings like imperfect competition and third degree discrimination. Additionally, we abstract from the literature's focus on separation versus bunching because enriching the heterogeneity of users relative to the firm's instruments makes bunching inevitable and, in fact, it is the bunching of marginal users that generates the covariance terms that characterize our results.

To emphasize the distinctions between our approach and this literature, Subsection 5.2 contrasts the results of Rochet and Stole [2002], who assume a particular distribution of heterogeneity but arbitrary differentiable non-linear tariffs, to a model of a credit card issuer where we consider a two-part tariff but allow for arbitrary user heterogeneity.

4.4 Empirical industrial organization

Empirical work in industrial organization has increasingly incorporated the concerns and effects discussed above. 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, Nosko [2010] allows for second-degree discrimination, Leslie [2004] incorporates third-degree discrimination, and Chiappori and Salanié [2000] and Cardon and Hendel [2001] discuss selection effects. Starc [2010] estimates the welfare loss from adverse selection and imperfect competition between price-choosing insurers. A number of recent papers, surveyed by Rysman [2009], consider platform markets and strong evidence of heterogeneity of externalities in these markets has been found by Rysman [2004],Ryan and Tucker [Forthcoming], Cantillon and Yin [2008] and Lee [2010], although these papers estimate demand without considering optimal firm responses.

This literature tends to consider structural and computational analyses based on parameterized demand systems which yield limited insight into economic mechanisms. The contribution that we aim for in this context is the simple formalization of the

---

[21]We believe that this is a good approximation for most applications. See Wilson [1993] on this topic.

empirical moments most important to identify persuasively in order to quantify the economic effects of interest, and which aspects of structural models are most important to estimate flexibly. For example, Gentzkow and Shapiro [2010]'s analysis of newspaper choices of political slant includes, in an appendix, an intuitive 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.

4.5 Applied price theory of asymmetric information

A number of recent papers estimate the effects of asymmetric information in specific settings, and develop models similar to ours in the richness of user heterogeneity relative to firm instruments. Einav et al. [2010a] and Einav and Finkelstein [2011] provide a characterization of selection perfectly competitive insurance provision, but allow only price to be chosen by insurers. Einav et al. [Forthcoming] consider the choice of continuous non-price product characteristics by a monopolist, but use a reduced-form approach based on elasticities that does not allow for welfare analysis, does not identify the distortions caused by profit maximization and does not relate the results to the primitives of user heterogeneity. Einav et al. [2010b] surveys a number of papers in the insurance setting which, like ours, consider intensive effects and third-degree discrimination (referred to as "moral hazard" and "pricing on observables," respectively).

Our contribution to this literature is to provide a model easily adaptable to the variety of settings addressed by this literature. The sorting effect characterizes the interactions between instruments in a manner that is simpler and more amenable to empirical measurement than what has been described in the literature. We aim to illustrate this contribution in Subsection 5.3, where we consider imperfect competition in prices and insurance coverage levels, thereby extending Einav and Finkelstein [2011].

4.6 Recent price theory and market design

A few recent theoretical papers have discussed the importance of moments of user heterogeneity similar to those we emphasize. Weyl and Tirole [2011] consider the

use of market power to screen valuable innovations and obtaining a characterization in terms of a covariance between innovation characteristics that is closely related to our result. Azevedo and Leshno [2011] shows that a game between firms in a matching market can be characterized in terms of moments of the distribution of user heterogeneity similar to the ones we emphasize. However, many of the aspects we consider, such as third-degree discrimination and consumption externalities, do not play a role in the more specific environments of these papers. They can therefore be seen as specific micro-foundations for the more general mechanisms we describe.

Table 1 gives examples of the various effects in the literature.

	Classical Product Design	Classical Contract Theory	Multi--dimensional Screening	Empirical IO	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 study and examples in the literature.

5   Applications

This section illustrates how our model can be applied to the analysis of several industries to yield substantive insights. We abstract from technical details like differentiability and uniqueness of the equilibrium and follow the notation of Section 3 unless

otherwise specified.[22]

5.1 Broadcast Media

Broadcast Media platform are usually multi-sided and restricted to zero prices on one side, where utility is therefore non-transferable.[23] To fix ideas, we consider the case of radio stations mentioned in Section 1.

A radio station charges price P^A to advertisers (A) and attracts listeners (L) by choosing the level of melodrama (m) in its programming. Listeners have types $\boldsymbol{\theta}^L \sim f^L$, which may include gender, age and other demographic traits. For simplicity, advertisers have uni-dimensional types $\theta^A \sim f^A$, so there is no sorting of advertisers. Participating listeners obtain $u^L(m, D; \boldsymbol{\theta}^L)$ where D is the distraction generated by advertisers and advertiser utility is $u^A = \theta^A W - P^A$, where W is the wealth of listeners. The preferences of advertisers are in the style of Rochet and Tirole [2003], implying that listeners are vertically differentiated for advertisers and that there is a unique marginal advertiser $(\widetilde{\theta^A} = P^A/W)$.

Outside options are zero. For $i \in \{A, L\}$, we define Θ^i, N^i and M^i as above. Listener $\boldsymbol{\theta}^L$ controls wealth in the amount $w(\boldsymbol{\theta}^L)$, and advertiser θ^A generates distraction in the amount of $d(\theta^A)$, so platform values are heterogeneous but fixed. The wealth of listeners is $W = \int_{\Theta^L} w f^L$ and the distraction of advertisers is $D = \int_{\Theta^A} d f^A$. Thus the (focal) instruments are $\left(P^A, m\right)$ and characteristics are (W, D). The platform incurs a cost of $c^L m$ per listener and c^A per advertiser, so classical marginal values are homogeneous $(\pi^L = -c^L m$ and $\pi^A = P^A - c^A)$. We will focus on the maximization or profit, $\Pi = N^A \pi^A + N^L \pi^L$.

We use Equation 6 to find the value of characteristics to the platform. The value of listener wealth is $\lambda^W = N^A P^A/W$ since there is no sorting of advertisers and no intensive platform effects. The value of distraction the platform is

$$\underbrace{M^L \sigma^L \left\langle u_D^L - \frac{\widetilde{u_D^L}}{\widetilde{u_m^L}} u_m^L, w\lambda^W \right\rangle}_{\text{sorting}} + \underbrace{N^L \frac{\widetilde{u_D^L}}{\widetilde{u_m^L}} c^m}_{\text{classical intensive effect}} = \lambda^D.$$

---

[22]In particular, we use $x_y = \partial x/\partial y$, $\widetilde{x} = \mathbb{E}\left[x \mid \boldsymbol{\theta} \in \partial\Theta\right]$ and $\overline{x} = \mathbb{E}\left[x \mid \boldsymbol{\theta}\Theta\right]$ for any x, y.

[23]This is the case in Anderson and Coate [2005]. White [2008] shows that the same principle applies to the case of online search engines.

Since π^L is homogenous it drops out of the covariance, and $N^L c^L \left(\widetilde{u_D^L} / \widetilde{u_m^L} \right)$ is the cost of using melodrama to hold fixed the number of listeners following an increase in distraction. In this case, the covariance term includes both preferences for distraction and for melodrama (rather than the simpler form of Section 2) because listeners have heterogeneous preferences for the focal instrument.

By Equation 4, the optimal price to advertisers is $P^A - c^A + \widetilde{d}\lambda^D = N^A/M^A$. Price equates marginal revenue to marginal cost net of the value of the distraction generated by an additional advertiser, which corresponds to the pricing principle originally expressed by Pigou [1912]. The optimal level of melodrama satisfies

$$\underbrace{-c^L m + \widetilde{w}\lambda^W}_{\text{marginal value}} + \underbrace{\sigma^L \left\langle u_m^L, w\lambda^W \right\rangle}_{\text{housewives' spending power}} - \underbrace{\frac{1}{M^L \widetilde{u_m^L}} N^L c^m}_{\text{intensive effect}} = 0$$

The marginal effect on wealth of an increase in melodrama depends on the extent to which listeners who prefer melodrama also have a control over family purchase decisions. This effect, captured by $\sigma^L \left\langle u_m^L, w \right\rangle$ explains why melodrama is a useful instrument for the radio station to attract valuable listeners. The analogue of the Cournot market power term is $-N^L c^m / M^L \widetilde{u_m^L}$, which captures the cost of using melodrama to attract an additional listener and is intuitively inversely proportional to the effectiveness of the instrument, $M^L \widetilde{u_m^L}$.

5.2 Credit Cards With Non-Linear Pricing

Credit cards issuers must attract accepted merchants and often charge an annual fee and offer a per-transaction cash-back subsidy. This makes them a canonical example of the non-linear pricing that has been a focus of the contract theory literatures and simultaneously an example of a (two-sided) platform as modeled by Rochet and Tirole [2003].

We consider a credit card issuer faced with users (C) and retailers (R). The issuer charges consumers an annual P^C and a per-transaction fee p.[24] It charges retailers a price P^R. Consumers have types $\boldsymbol{\theta}^C \sim f^C$ which may account for their wealth, gender, impulsiveness and other demographic characteristics. Retailers have uni-dimensional types $\theta^R \sim f^R$, so there is no sorting of retailers. User $\boldsymbol{\theta}^C$ chooses an amount of

---

[24]We impose no restrictions on the sign of p, although it is normally negative in equilibrium.

transactions q to maximize her transactional surplus $S\left(p; \boldsymbol{\theta}^C\right) = \max_q v\left(q; \boldsymbol{\theta}^C\right) - pq$. Thus q is each user's intensive margin of participation. When N^R retailers participate, user $\boldsymbol{\theta}$ obtains $u^C = N^R S(p; \boldsymbol{\theta}^C) - P^C$. By the envelope theorem, $du^C/dp = -N^R q$. Participating retailers obtain utility $u^R = \theta^R Q - P^R$, where Q is the total demand for transactions by consumers. We follow Rochet and Tirole [2003] in assuming that consumers and retailers interact with each other at random.

Outside options are zero, so $\Theta^R = \left\{\theta^R \geq P^R/Q\right\}$ and $\Theta^C = \left\{\boldsymbol{\theta}^C : S \geq P^C/N^R\right\}$. For for $i \in \{C, R\}$, N^i and M^i are defined in the usual way. Total demand is $Q = \int_{\Theta^C} q(p; \boldsymbol{\theta}^C) f^C$. A total of QN^R transactions occur, and the platform incurs a cost c per transaction. We can therefore define $\pi^R = P^R$ and $\pi^C = P^C + (p - c)qN^R$. Profit is $\Pi = \int_{\Theta^R} \pi^R + \int_{\Theta^C} \pi^C$. The platform's instruments are $\left(P^R, P^C, p\right)$, the focal instruments are P^R and P^C, and the characteristics are $\left(N^R, Q\right)$.

Consider first welfare maximization. From Equation 6, the value of Q to a welfare maximizer is $\lambda^{WQ} = N^R \overline{\theta^R} + N^R (p - c)$, the value to the average retailer and to the platform. The value of an additional merchant is $\lambda^{WN} = N^C \overline{S} + N^C (p - c)\overline{q}$, its value to users and to the platform. There are no sorting effects because there is a unique marginal retailer and all marginal consumers have marginal utility of N^R equal to $\widetilde{S} = P^C/N^R$.

From Equation 4, the welfare maximizing price to consumers is equal to the marginal value of a consumer to retailers ($P^C = -\widetilde{q}\lambda^{WQ}$) because there is no cost of an additional consumer. Similarly, the welfare maximizing price to retailers is $P^R = -\lambda^{WN}$. From Equation 5, the welfare maximizing per-transaction price p is such that the benefit of additional demand equals the cost ($\lambda^{WQ} = 0$). This implies $p = c - \overline{\theta^R}$, the price of an additional transaction is equated to its cost net of its value to retailers. This further implies that $P^C = 0$, as is intuitive since there is no direct cost of an additional consumer.

Consider now profit maximization. By Equation 6, we obtain $\lambda^{\Pi Q} = N^R \left(\widetilde{\theta^R} + p - c\right)$ and $\lambda^{\Pi N} = N^C \left(\widetilde{S} + (p - c)\overline{q}\right)$. The Spence distortion is clearly illustrated by the issuer considering the preferences of marginal retailers $(\widetilde{\theta^R})$, and of marginal consumers $(\widetilde{S})$. From Equation 4, the profit maximizing price to users $P^C - N^C/M^C = -\widetilde{q}\lambda^Q$ and to retailers is $P^R - N^R/M^R = -\lambda^N$. Both equate marginal revenue is equated to marginal cost (zero) net of the value of the relevant externality.

From Equation 5, the optimal transaction fee p is such that

$$-\underbrace{M^C N^R \mathrm{Var}(q \mid \partial\Theta) \lambda^{\Pi Q}}_{\text{sorting}} + \underbrace{N^C N^R (\overline{q} - \widetilde{q}) + N^C \overline{\frac{dq}{dp}} \lambda^{\Pi Q}}_{\text{intensive effect}} = 0$$

Increasing p increases the revenue from existing transactions by $N^C N^R \overline{q} = Q N^R$, but requires decreasing P^C to hold fixed the number users leading to a loss of $N^C \widetilde{u_p^C} = -N^C N^R \widetilde{q}$. Increasing p also reduces the number of transactions executed by infra-marginal users by $N^C \overline{dq/dp}$. Finally, the power of p to sort for valuable marginal users is quantified by $\sigma^C \langle u_p^C, q \rangle = -N^R \mathrm{Var}(q \mid \partial\Theta)$. The sensitivity of each consumer to p is proportional to her number of transactions, as is her value to retailers and to the platforms.

The following manipulation highlights how the model relates to the existing literature:

$$\frac{p - c + \widetilde{\theta^R}}{p} = \left(1 - \frac{\widetilde{q}}{\overline{q}}\right) \left(\epsilon_M \frac{N^R \mathrm{Var}(q \mid \partial\Theta)}{\widetilde{q}} + \overline{\epsilon_I}\right)^{-1}$$

where $\overline{\epsilon_I} = -\overline{dq/dq}\,(\overline{q}/p)$ is the average elasticity of transactions of infra-marginal participants, and $\epsilon_M = M^C \widetilde{q} p / Q$ is the elasticity of transactions due to marginal consumers. If users are homogenous ex-ante, as in Bedre-Defolie and Calvano [2010], then $\overline{q} = \widetilde{q}$, the RHS vanishes and we recover the welfare maximizing condition for p, as that paper concludes. If there are no externalities from users to retailers and all marginal users had zero consumption, then $\mathrm{Var}(q \mid \partial\Theta) = \widetilde{q} = \widetilde{\theta^R} = 0$ so we recover the Wilson [1993] inverse elasticity formula $(p - c)/p = 1/\overline{\epsilon_I}$. If participation is random as in Rochet and Stole [2002], marginal users have heterogeneous levels of consumption, so $\mathrm{Var}(q \mid \partial\Theta) \neq 0$ disciplines the use of non-linear pricing. This is not the case in Mussa and Rosen [1978], for instance, where the marginal consumer has zero consumption.

5.3   Imperfectly Competitive Insurance Provision

The following application aims to extend the approach of Einav and Finkelstein [2011] and of the model in Section 2 to allow for imperfectly competitive insurers who choose both prices and levels of coverage.

We consider two insurers, indexed by $i \in \{A, B\}$, each choosing a price-coverage pair (P^i, ρ^i). Patients have types $\boldsymbol{\theta} \sim f$ which may account for their risk preference, health status and other traits. Patients obtain $u = v(\rho^i; \boldsymbol{\theta}) - P^i$ from purchasing from insurer i. Outside options are zero and that patients purchase from a single insurer. Patient $\boldsymbol{\theta}$'s cost of provision by insurer i is $c(\rho^i, \boldsymbol{\theta})$, and insurer i's profit is $\Pi = \int_{\Theta^i} [P - c(\rho, \boldsymbol{\theta})] f$. As in Subsection 3.2, we consider the exiting margin of i as those patients indifferent between i and no insurance, and the switching margin as those patients indifferent between the two insurers. We assume a symmetrical equilibrium where insurers choose the same levels of both instruments. This implies changes along the switching margin are socially neutral and therefore ignored by a welfare maximizer but not by a profit maximizer. Since insurers are symmetrical, we eliminate the superscript i for notational clarity.

By Equations 7 and 8, a welfare maximizer considers only the exiting margin and sets $P = \widetilde{c^X}$, price is equal to the cost of coverage the average user on the exiting margin. The welfare maximizing level of coverage is $N(\overline{v_\rho} - \overline{c_\rho}) - M^X \sigma^X \langle v_\rho, c \rangle = 0$, as in Section 2. Conversely, a profit maximizing duopolist considers both the switching and exiting margins and ignores the surplus of patients. it chooses P and ρ such that

$$P - \underbrace{\frac{N}{M^{X+S}}}_{\text{Cournot distortion}} = \underbrace{\widetilde{c}^{X+S}}_{\text{Akerloff-Einav-Finkelstein term}}$$

$$N \underbrace{\widetilde{v}_\rho^{S+X}}_{\text{Spence term}} - N\overline{c_\rho} - \underbrace{M^{X+S}\sigma^{X+S} \langle u_\rho, c \rangle}_{\text{Rothschild-Stiglitz term}} = 0$$

An oligopolist's profit maximizing conditions include the Cournot [1838] and Spence [1975] distortions as expected. Additionally, when setting the level of its focal instrument, a profit maximizing oligopolist considers the marginal revenue and cost of users on the switching margin, as described in Subsection 3.2. The optimal number of patients is determined based on the cost of patients on the switching, rather than the exiting margin, and patients on the switching margin are infra-marginal to the industry and therefore are likely to be more similar to average patients than to patients on the exiting margin. The responsiveness of prices to the cost of infra-marginal patients is a distortion from the welfare maximizing prices that arises from competition under

assymetric information about heterogeneous values, and is therefore in the same spirit as those described by Akerlof [1970] and Einav et al. [2010a].

When setting the level of its non-focal instrument, a profit maximizer caters to the tastes of users on the switching margin to the extent that those users are valuable. Thus a profit maximizer uses its non-focal instrument to poach valuable patients form its competitor and does not take into account the externality this imposes. This cream-skimming distortion is similar in spirit to that described by Rothschild and Stiglitz [1976] and Akerlof [1976].

Echenique [2002] shows that in a stable symmetric equilibrium, symmetric competition is equivalent to an individual optimization problem. This implies that increased competition (in the sense of differentiated Bertrand) is equivalent to an increase in the density of the switching margin. In that case, greater competition increases M^{X+S} thereby mitigating the Cournot distortion but also increasing the difference between $\widetilde{\pi^X}$ and $\widetilde{\pi^{X+S}}$ which exacerbates the Akerlof-Einav-Finkelstein distortion.

Regarding non-focal instruments, if users on the switching margin are more similar to average users than to users on the exiting margin, competition will mitigate the Spence distortion. However, the incentive to cream-skim increases with the density of the switching margin, so competition exacerbates the Rothschild and Stiglitz [1976]. In fact, the optimality condition shows that the cream-skimming distortion always leads to market collapse when competition is sufficiently intense, while this is not necessarily the case for the selection distortion.

6 Additional Extensions

In this section, we discuss how our results can be combined with standard techniques to yields insight on economically important problems.

Rosse [1970] showed when the marginal cost of competing firms may be inferred from observations of a relevant demand elasticity. The approach has become central to modern empirical industrial organization research such as Berry [1994] and Berry et al. [1995]. In richer settings, our results generate the moment conditions necessary to recover marginal costs of participants (using the focal instrument) or the marginal cost of non-focal instruments. Because such applications rely on the first-order optimality conditions we derive, they follow directly from our analysis.

Chetty [2009] summarizes public economics applications of a local approximation approach popularized by Harberger [1964] to measure marginal distortions from social welfare at a profit maximizing equilibrium. In these applications, measuring this gradient is simplified by envelope conditions that reduce the estimation problem to a small number of market aggregate quantities, as we have done above. The gradient of welfare provides an estimate of the marginal distortions at that equilibrium and the locally optimal direction for policy interventions. In our multidimensional environment, the gradient of welfare indicates the direction of steepest ascent in the instrument space. This approach can also provide a non-parametric estimate of the welfare gains associated with a small change of the instruments.

Previous versions of this paper included several additional applications. These include analyzes of newspaper choices of political slant as in Gentzkow and Shapiro [2010], labor markets as in Akerlof [1976], and industrial policy as in Restuccia and Rogerson [2008]. Some of the applications above could obviously also be enriched by adding more parts to tariffs, allowing menus of health coverage, etc. One application of theoretical interest would be to model imperfect competition when users can join multiple platforms the characteristics generated by an individual depend on the subset of platforms she participates on, as in Ambrus and Reisinger [2006]. Such an application would require confronting the issues in platform competition that we discuss in Section 3.2.

7   Conclusion

We proposed a new approach to modeling firms' choices of many product characteristics when users differ in many dimensions of preferences and values to the firm or other users. By enriching heterogeneity and considering a finite numbers of product design instrument, we obtained a smooth, price theoretic analysis that contrasts with previous characterizations of similar problems and is amenable both to theoretical analysis and empirical estimation. We obtain intuitive necessary conditions for profit and welfare maximization in terms terms of aggregate market quantities and and the moments of the distribution of user preferences and values. Namely, we characterize the power of an instrument to sort for valuable marginal users, which is quantified by the density of marginal users and the covariance, within that set , between their preferences for the instrument and their value to the platform.

We allow for a number of considerations typical of distinct literatures, such as contract theory and Industrial Organization. These include non-transferable utility, cream-skimming distortions, adverse/advantageous selection, moral hazard, nonlinear pricing, third-degree price discrimination and imperfect competition. Our model is also able to accommodate consumption externalities between users in a mathematically straightforward way by expressing the firm's problems in terms of constrained optimization. Our model can be applied to several industries, including multi-sided markets and settings with imperfect competition.

The modeling of asymmetric imperfectly competitive equilibria and of competition between platforms is beyond the scope of this paper and a promising avenue for future research. Regarding the later, a change in one platform's instruments affects the platform's own characteristics as well and those of its competitors, so the characteristics of competitors can be included as constraints in each firm's optimization problem. A similar approach could be used to model a multi-product monopolist. If consumers purchase at most one product, this would require only consideration of the relevant switching margins between products within a firm.

We have not considered externalities to unserved individuals as in Segal [1999]. We have allowed users to make discrete choices only regarding their participation, although including other discrete choices would blur the sharp distinction between the intensive and extensive effects. Finally, in our extension to competition, markets clear through user choices. Markets might instead clear based on firm choices over users, as in Gale and Shapley [1962].[25] 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. The extension of these lines of research to the richly heterogeneous settings we have been discussing is a promising path for future work.

---

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

References

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

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

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.

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

Mark Armstrong and Jean-Charles Rochet. Multi-dimensional screening: A user's guide. *European Economic Review*, 43(4-6):959–979, 1999.

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

Eduardo M. Azevedo and Jacob D. Leshno. A supply and demand framework for two-sided matching markets. http://www.people.fas.harvard.edu/~jleshno/papers/Simple_Large_Matching.pdf, 2011.

Özlem Bedre-Defolie and Emilio Calvano. Pricing payment cards. https://sites.google.com/site/ozlembedre/, 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 Bru 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://www.ecares.org/ecare/personal/cantillon/web/bund.pdf, 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.

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.

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.

Douglas Hale, George Lady, John Maybee, and Jam Quirk. *Nonparametric Comparative Statics and Stability*. Princeton University Press, 1999.

Arnold C. Harberger. The measurement of waste. *American Economic Review*, 54 (3):58–76, 1964.

I. 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://www.soniajaffe.com/articles/Jaffe_Weyl_Merger_Analysis.pdf, 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://pages.stern.nyu.edu/∼rslee/papers/VIExclusivity.pdf, 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.

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

Diego Restuccia and Richard Rogerson. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4):707–720, 2008.

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

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

Paul A. Samuelson. The stability of equilibrium: Comparative statics and dynamics. *Econometrica*, 9(2):97–120, 1941.

Irwin Sandberg. Global implicit function theorems. *Circuits and Systems, IEEE Transactions on*, 28(2):145–149, 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. Starc. Insurer pricing and consumer welfare: Evidence from medigap. *Manuscript. Harvard University*, 2010.

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

S. 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.

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://www.people.fas.harvard.edu/∼fabinger/Restatement.pdf, 2011.

E. Glen Weyl and Jean Tirole. Market power screens willingness-to-pay. http://idei.fr/doc/by/tirole/market_power_circulating.pdf, 2011.

Alexander White. Search Engines: Left Side Quality Versus Right Side Profits. http://alex-white.net/Home/Research_files/White_LSRS.pdf, 2008.

Alexander White and E. Glen Weyl. Insulated platform competition. http://alex-white.net/Home/Research_files/WWIPC.pdf, 2011.

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

8   Appendix

Proof of Proposition 1

Proof. The original problem requires choosing $\boldsymbol{\rho}$ to maximize H. However, the feedback effects between characteristics are made more transparent by considering the (equivalent) problem of choosing $(\boldsymbol{\rho}, \boldsymbol{K})$ to maximize H subject to the $\mathcal{K}$ constraints $\int_{\Theta} k^i f = K^i, \forall i$. The Lagrangian for this problem is $\mathcal{L}^H = \int_{\Theta} hf + \sum_{i=1}^{\mathcal{K}} \left(\int_{\Theta} k^i f - K^i \right) \lambda^{Hi}$ where λ^{Hi} is the Lagrange multiplier on constraint i. The FOC for any instrument ρ^l (including the focal instrument $\rho^\star$) is

$$
\begin{aligned}
\frac{d\mathcal{L}^H}{d\rho^l} &= \int_{\Theta} \frac{dh}{d\rho^l} f + \int_{\partial\Theta} \frac{du}{d\rho^l} hf + \sum_{i=1}^{\mathcal{K}} \left(\int_{\Theta} \frac{dk^i}{d\rho^l} f + \int_{\partial\Theta} \frac{du}{d\rho^l} k^i f \right) \lambda^{Hi} = \\
&= N\overline{h_{\rho^l}} + M\sigma \left\langle u_{\rho^l}, \pi \right\rangle + \sum_{i=1}^{\mathcal{K}} \left(N\overline{k_{\rho^l}^i} + M\sigma \left\langle u_{\rho^l}, k^i \right\rangle \right) \lambda^{Hi} + \widetilde{u_{\rho^l}} \left(M\widetilde{\pi} + \sum_{i=1}^{\mathcal{K}} M\widetilde{k^i} \lambda^{Hi} \right) = 0
\end{aligned}
$$

Here, we use the definition of covariance to transform all expectations of products. We also use the fact that, conditional on the marginal set, $h = \pi$.

The FOC for any characteristic K^j is

$$
\begin{aligned}
\frac{d\mathcal{L}^H}{dK^j} &= \int_{\Theta} \frac{dh}{dK^j} f + \int_{\partial\Theta} \frac{du}{dK^j} hf + \sum_{i=1}^{\mathcal{K}} \left(\int_{\Theta} \frac{dk^i}{dK^j} f + \int_{\partial\Theta} \frac{du}{dK^j} k^i f \right) \lambda^{Hi} - \lambda^{Hj} = \\
&= N\overline{h_{K^j}} + M\sigma \left\langle u_{K^j}, \pi \right\rangle + \sum_{i=1}^{\mathcal{K}} \left(M\sigma \left\langle u_{K^j}, k^i \right\rangle + N\overline{k_{K^j}^i} \right) \lambda^{Hi} - \lambda^{Hj} + \widetilde{u_{K^j}} \left(M\widetilde{\pi} + \sum_{i=1}^{\mathcal{K}} M\widetilde{k^i} \lambda^{Hi} \right)
\end{aligned}
$$

Taking the FOC for the focal instrument $\rho^\star$ and re-arranging immediately yields Equation 1. Using the FOC for $\rho^\star$ to eliminate $\left(M\widetilde{\pi} + \sum_{i=1}^{\mathcal{K}} M\widetilde{k^i} \lambda^{Hi} \right)$ in the FOC for each other instrument ρ^l yields Equation 2. Using the FOC for $\rho^\star$ to eliminate $\left(M\widetilde{\pi} + \sum_{i=1}^{\mathcal{K}} M\widetilde{k^i} \lambda^{Hi} \right)$ in the FOC for each K^j yields the expression

$$
\lambda^{Hj} = M\sigma \left\langle u_{K^j} - \frac{\widetilde{u_{K^j}}}{\widetilde{u_{\rho^\star}}} u_{\rho^\star}, \pi + \boldsymbol{k}\boldsymbol{\lambda^H} \right\rangle + N \left(\overline{h_{K^j}} - \frac{\widetilde{u_{K^j}}}{\widetilde{u_{\rho^\star}}} \overline{h_{\rho^\star}} \right) + N \left(\overline{\boldsymbol{k_{K^j}}} - \frac{\widetilde{u_{K^j}}}{\widetilde{u_{\rho^\star}}} \overline{\boldsymbol{k_{\rho^\star}}} \right) \boldsymbol{\lambda^H}
$$

This system of $\mathcal{K}$ equations can be expressed as vector-valued equation, using the matrices defined above, which yields Equation 3. □