American Economic Review 100 (September 2010): 1642–1672
http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.4.1642

A Price Theory of Multi-Sided Platforms

By E. Glen Weyl*

I develop a general theory of monopoly pricing of networks. Platforms use insulating tariffs to avoid coordination failure, implementing any desired allocation. Profit maximization distorts in the spirit of A. Michael Spence (1975) *by internalizing only network externalities to marginal users. Thus the empirical and prescriptive content of the popular Jean-Charles Rochet and Jean Tirole* (2006) *model of two-sided markets turns on the nature of user heterogeneity. I propose a more plausible, yet equally tractable, model of heterogeneity in which users differ in their income or scale. My approach provides a general measure of market power and helps predict the effects of price regulation and mergers.* (*JEL* D42, D85, L14)

The pricing problems of payment and advertising platforms have much in common. Both seek to attract two distinct groups of users: AmEx needs cardholders and merchants, while the *New York Times* recruits readers and advertisers. Because the value each group takes from using these services depends on the size of the *other side of the market*, the platform's pricing and marketing strategies to each group are closely linked. Therefore policy directed at alleviating distortions caused by market power in these industries must take account of how interventions on one side affect welfare and platform behavior on the other.

Yet despite credit cards and newspapers both being canonical *two-sided markets*, the economics of these industries seem intuitively quite different. Consumers most likely to carry AmEx are those who most value the opportunity to use the card. These loyal cardholders therefore value the participation of merchants more than those indifferent between AmEx and another payment form do. Given its limited ability to price discriminate, AmEx fails to fully internalize the preferences of loyal users, putting too little effort into attracting merchants and charging them a higher price than would be socially optimal. However, when the costs of attracting cardholders rise and therefore cardholder incentives fall, AmEx will tend to serve only users who value merchant participation more strongly, leading them to attract more merchants with lower fees. This logic is the basis of the burgeoning literature on two-sided markets pioneered by Rochet and Tirole (2003).

* Harvard Society of Fellows and Toulouse School of Economics: 78 Mount Auburn Street, Cambridge, MA 02138 (e-mail: weyl@fas.harvard.edu). This paper was split off as one part of a previous working paper "Monopolies in Two-Sided Markets: Comparative Statics and Identification." I am grateful to el Ministerio de Hacienda de Chile and the Centro de Investigación Económica at the Instituto Tecnológico Autonomo de México (CIE-ITAM), which hosted me on a visit while I conducted parts of this research. The Harvard Milton Fund supported the last stage of this project and financed the excellent research assistance provided by Will Weingarten, Alex White, and especially Stephanie Lo. Ahmed Jaber and Dan Sacks also acted as research assistants with the permission of their employer Amy Finkelstein, to whom I am grateful for her generosity. I also appreciate the helpful comments and advice on this research supplied by Mark Armstrong, David Evans, Jeremy Fox, Andrea Hawkley, Alisha Holland, Scott Kominers, Luis Rayo, Bill Rogerson, Dick Schmalensee, Jesse Shapiro, and especially Maxim Engers, and seminar participants at the Banco de México, the Barcelona Graduate School of Economics, CIE-ITAM, Harvard University, IESE Business School, Northwestern University, Télécom ParisTech Conference on the Economics of Information and Communications Technology, the Toulouse School of Economics (TSE), University of Chicago, and University of California at Berkeley. I owe a special debt to Pai-Ling Yin for detailed discussions of the paper. All confusion and errors are my own.

Matters are quite different for the *New York Times*. Its loyal customers are high income readers who dislike advertising but are willing to pay more for the paper's content than marginal readers who are less sensitive to advertising.[1] Thus the *Times* fails to internalize loyal readers' distaste for advertising, leading to potentially excessive advertising as a result of *below* optimal pricing to advertisers despite market power. Increases in the costs of distribution that reduce the number of subscribers will tend to *reduce* advertisements as the paper internalizes the costs to its wealthier readers. While intuitive in many markets, this opposite case has been assumed away by previous theoretical and empirical analysis of two-sided markets.

This paper shows that both of these are special cases of a simplified reformulation of the Rochet and Tirole, henceforth RT, (2006) model of monopoly in two-sided markets. The crucial difference between cases is the source of user heterogeneity. While credit card users primarily differ in the *interaction* (or usage) value they take from merchants accepting cards, newspaper readers differ most importantly in their *membership* value from reading the paper's content. This distinction is crucial because participation on one side of the market effectively determines the quality of the platform on the other side. Therefore, like any monopolist who must choose a single quality as well as quantity, the platform internalizes network effects to marginal rather than average participating users (Spence 1975).

The diversity of possibilities in two-sided markets does not eliminate the theory's predictive power. Because the distortions caused by market power (Section II) are linked to observable comparative statics (Section IV) through user heterogeneity, both intuition about the sources of this heterogeneity and empirical data can help calibrate the model in applications. Further restrictions may then be placed on the models (Section V) with a clear sense of how special assumptions increase predictive and prescriptive power. Together these results inform policy analysis in two-sided markets (Section VI), providing a general measure of market power and helping to predict the effects of regulation and mergers.

My analysis rests on a view of the platform's problem as choosing participation rates on the two sides rather than the prices supporting this allocation. This approach (Section I) is justified by an appropriate platform pricing strategy, the *insulating tariff*, that avoids potential coordination failures, thereby allowing the platform to achieve any desired allocation as a unique equilibrium. It applies, along with much of my analysis, more generally (Section III) than the RT (2006) model, allowing a simple approach to analyzing network industries with arbitrary heterogeneous utility, network effects, and any number of sides. I thereby answer perhaps the oldest open question in the theory of network industries (S. J. Liebowitz and Stephen E. Margolis 1994): does a monopolistic platform internalize and therefore neutralize network effects? The answer is yes, but imperfectly given the Spence distortion.

Of course this article is only a first pass at a general analysis of network pricing. Section VII therefore concludes by discussing directions for future research. Longer and less instructive proofs are collected into an Appendix available at http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.4.1642.

[1] Note that the intuitive stories I tell about these industries are not intended to be specific prescriptions about policy in these industries, but rather concrete instances of general theoretical possibilities. It is the mapping between the story's assumptions and the conclusions about policy, rather than the validity of the assumptions in a particular industry, that I am interested in here. Thus I do not have any empirical evidence substantiating my stories, evidence that would be highly desirable before reaching definite policy conclusions. For example, Ulrich Kaiser and Minjae Song (2009) argues that users do not actually dislike advertising, as my story assumes. Furthermore, in some cases at least, content may be viewed as an interaction rather than membership benefit if it is tightly tailored to accompany an ad, though I doubt this is the case for newspapers.

I. Framework

The definition of two-sided markets is controversial (RT 2006; Marc Rysman 2009). For me, the phrase denotes a style of industrial organization modeling[2] developed by, among others, Bernard Caillaud and Bruno Jullien (2001; Caillaud and Jullien (2003; Rochet and Tirole (2003); Simon P. Anderson and Stephen Coate (2005); Mark Armstrong (2006); and Rochet and Tirole (2006). These models tend to emphasize three features.[3]

i) *Multi-product firm*: A *platform* provides distinct services to two *sides of the market*, which can be explicitly charged different prices.

ii) *Cross network effects*: Users' benefits from participation depend on the extent of user participation on the other side of the market, which varies with market conditions.

iii) *Bilateral market power*: Platforms are price setters (monopolistic or oligopolistic) on both sides of the market and typically set uniform prices.

The failure of any of these conditions makes simpler and better-understood models more appropriate. If a platform does not explicitly charge different prices to different groups of users, it is best viewed as a standard, one-sided network. When participation does not vary on both sides a vertical monopolies model fits better. An absence of market power allows us to model the firm as a distributor. However, many industries[4] relevant to industrial policy exhibit all of the above characteristics. RT (2006) introduce a "canonical model," in their words, of monopoly capturing these three features in a simple manner that still generalizes the two most influential models in the literature, those of RT (2003) and Armstrong (2006).

A. *The RT* (*2006*) *model*

Before describing it more formally, I highlight a few key assumptions of the model, on top of the familiar notions of monopoly and constant marginal cost.

i) User valuations are taken as exogenous to any direct interactions between users on the two sides. Thus the RT (2006) model takes a "macro" approach, in the terminology of Nicholas Economides (1996). While "micro" models that directly incorporate such interactions (Julian Wright 2004; Andrei Hagiu 2006; Graeme Guthri and Julian Wright 2007) have similar positive forms, their normative implications can be quite different.

ii) Network effects are all *across*, not *within* the two sides. This rules out, for example, negative within-side effects from competition among software creators or positive collaboration effects among operating system users.

[2] These can either be viewed as models aimed at capturing important features of some true class of "two-sided markets" or they can be viewed as a style of modeling that captures some elements of "two-sidedness" that are more or less important in different industries. I lean towards the second view.

[3] I am grateful to Bruno Jullien and Patrick Rey for helping me refine these criteria.

[4] For example, credit cards, newspapers, operating systems, Internet service providers and others discussed by RT (2003) and Armstrong.

iii) Users on each side interact with either all or a random subset of users on the other side, price discrimination[5] within a particular side based on the number of such interactions is impossible, and user values are affine in the number of users on the other side. This does not rule out some users interacting with a larger random and unpriced sample of users on the other side; this will just magnify their interaction valuations.

iv) Finally, it assumes that users on each side of the market are of equal value to those on the other side. This rules out, for example, high advertising-value readers of a newspaper and video games that are especially valued by gamers (Robin S. Lee 2009).

Few of these assumptions are necessary for my analysis; in fact, assumptions ii–iv can be substantially relaxed or eliminated entirely. However, doing so complicates the exposition. Furthermore, given its greater parsimony, I suspect the RT (2006) model will continue to be the most attractive framework in many applications, including those I focus on: the payments and newspaper industries. Most importantly, its assumptions fit many industries quite well.

Consider the case of the newspaper industry as an example. The ways in which advertisers gain from readers viewing their ads, or why exactly users dislike ads (Gary S. Becker and Kevin M. Murphy 1993), seems fairly exogenous to industrial policy in the newspaper industry. While advertisements sometimes compete within a paper for user attention, it seems fairly reasonable to assume that advertisers are close to indifferent as to the number of other advertisements included in a paper, and readers are indifferent to the number of other readers of the paper. Readers usually read all or a fairly random selection of advertisements in a paper, and certainly it seems difficult to charge users (or advertisers) differentially based on the number of advertisements viewed. Finally, some advertisements are certainly more annoying than others and some readers more valuable than others to advertisers. However, I follow many top past applied papers (Stephen T. Berry and Joel Waldfogel 1999; Matthew Gentzkow and Jesse Shapiro 2010; Yin Fan 2010) on industries with advertising in viewing this as of second-order importance.

Therefore I develop most of my analysis in the context of the RT (2006) application, treating the general case only in Section III. There I show that my basic message in the RT (2006) applies generally. Therefore little is lost by focusing on the RT (2006) model, and Section III will likely be of most interest to theoretically inclined readers.

B. *User Preferences and Heterogeneity*

There is a continuum of potential users on each side $\mathcal{I} = \mathcal{A}, \mathcal{B}$ of the market, with mass normalized to 1. Thus the number of users participating on each side represents the fraction of potential participants choosing to do so. All quantities are scaled accordingly as discussed below. $\mathcal{I}$ refers to a generic side of the market and $\mathcal{A}$ and $\mathcal{B}$ to refer to specific sides in examples.

A typical user i on side $\mathcal{I}$ has an inherent *membership* benefit or cost $B_i^{\mathcal{I}}$ from participating in the service if no users participate on the other side. For example, developers must pay fixed costs even if no users own the operating system the software runs on. Given my normalization of a unit mass of users, $B_i^{\mathcal{I}}$ must be measured in terms of the total value all users on side $\mathcal{I}$ would derive if they participated given that they have the same preferences as user i. Suppose a town has

[5] As in all models with market power, the impossibility of price discrimination plays a crucial role in normative conclusions. I believe price discrimination is probably neither systematically easier nor more difficult in two-sided markets than in standard markets. Even when some discrimination is possible, I believe the discrimination-free model gives some insight, as long as the discrimination is imperfect.

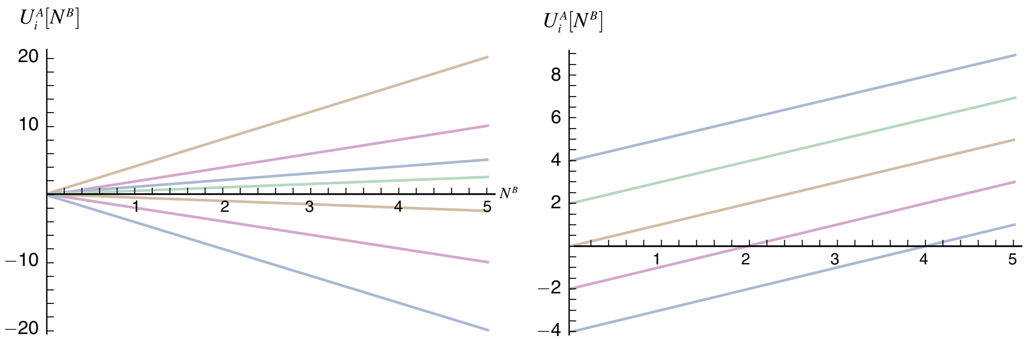

Figure 1. $U_i^{\mathcal{A}}(N^{\mathcal{B}})$ for Various RT (2003) (left) and Armstrong (right) Preferences.

Note: This Illustrates the Two Dimensions of Heterogeneity: Interaction and Membership Values, Respectively

100 possible newspaper subscribers and user i values reading her paper at \$500 a year; then her membership benefit would be $B_i^{\mathcal{I}} = \$50{,}000$.

Each user also derives an *interaction*[6] benefit or cost of participation $b_i^{\mathcal{I}}$ for every user that participates on the other side. Again this must be appropriately scaled.[7] One of 1,000 credit card holders who makes 100 purchases every year deriving a 50 cent surplus from credit over cash would have a interaction benefit of \$50,000 per year, as this would be the value to all users on her side if all cards were accepted and all users on her side had the same preferences as she does. If there were 1,000 possible advertisers and 100 readers in a town and the disutility to a user i of each advertiser being included in a local newspaper were ten cents, then the interaction cost of that user would be $b_i^{\mathcal{I}} = \$10{,}000$. I follow most of the industrial organization literature in assuming that user utility is quasi-linear in money.

Formally the utility derived by user i on side $\mathcal{I}$ from participating is

$$U_i^{\mathcal{I}} = B_i^{\mathcal{I}} + b_i^{\mathcal{I}} N^{\mathcal{J}} - P^{\mathcal{I}}(N^{\mathcal{J}})$$

where $N^{\mathcal{J}}$ is the number of users participating on side $\mathcal{J} = -\mathcal{I}$, the other side than $\mathcal{I}$. $P^{\mathcal{I}}(N^{\mathcal{J}})$ is the tariff set by the platform (independence of i disallows price discrimination), prescribing how much users must pay (or will be paid) to participate conditional on a given size of the platform on side $\mathcal{J}$. Users on each side can therefore be heterogeneous along two dimensions: interaction and membership values. Two natural special cases involve only one dimension of heterogeneity. RT (2003) assume that $B_i^{\mathcal{I}} \equiv 0$ and that users have heterogeneous interaction values. Armstrong (2006) assumes homogeneous interaction values ($b_i^{\mathcal{I}} \equiv b^{\mathcal{I}}$) and allows heterogeneous membership values. Figure 1 shows the difference between these specifications. Utility is graphed as a function of participation on the other side of the market for various RT (2003) preferences (left) and Armstrong preferences (right). When, in general, there are both dimensions of heterogeneity, even fixing $N^{\mathcal{J}}$ and $P^{\mathcal{I}}$, many different types of users may be just on the margin between participating and not (have $U_i^{\mathcal{I}} = 0$): some may have high interaction benefits but large membership costs; others may have low interaction benefits and no membership costs. This is pictured in Figure 2, where all users lying along the lines are marginal. The implications of these different

[6] RT (2006) refers to this as the user's usage valuation; I eschew this terminology to avoid confusion, as users have no choice over how intensively to use the service in the RT (2006) model.

[7] Of course these scales can be renormalized as suits a given application, so long as this is done consistently.

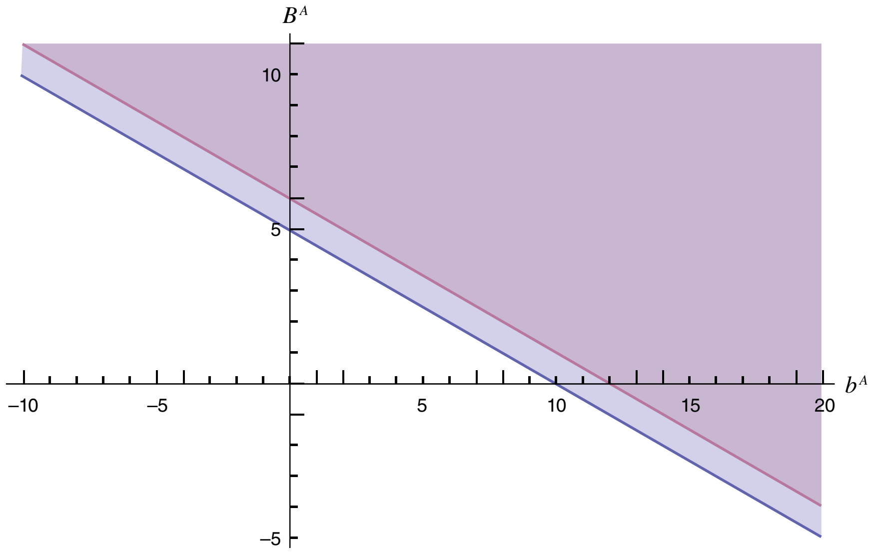

FIGURE 2. THE SET OF USERS PARTICIPATING ON SIDE $\mathcal{A}$ WHEN HALF OF USERS PARTICIPATE ON SIDE $\mathcal{B}$ AND $p\mathcal{A} = 5$ AND 6 RESPECTIVELY

sources of user heterogeneity are the primary focus of this paper. Formally I assume that the user parameters are distributed according to some massless, twice continuously differentiable[8] bi-variate distribution with probability density function $f^{\mathcal{I}}(B_i^{\mathcal{I}}, b_i^{\mathcal{I}})$ and a full support.

C. *Coordination and Insulating Tariffs*

Once the tariff is set, users on the two sides of the market play a game. A user i on side $\mathcal{A}$ will choose to participate if and only if

$$B_i^{\mathcal{A}} + b_i^{\mathcal{A}} N^{\mathcal{B}} > P^{\mathcal{A}}(N^{\mathcal{B}}).$$

However, this typically depends on the decisions of user on side $\mathcal{B}$, which $\mathcal{A}$ users take as given. There may be multiple equilibria for some tariffs and distributions of user tastes. For example, suppose $b_i^{\mathcal{I}} \equiv 1$, $B_i^{\mathcal{I}} \equiv 0$ and $P^{\mathcal{I}} \equiv ½$ for $\mathcal{I} = \mathcal{A}, \mathcal{B}$.[9] Then it is clearly an equilibrium for either all or none of the users to participate. In the former case the utility of participation on either side (taking the other as given) is ½; in the latter it is −½. This is the classic "chicken and egg" problem in two-sided markets (Caillaud and Jullien 2003).

Yet, in a sense, equilibrium multiplicity is inessential to the analysis of two-sided markets. A given pair of participation rates $\widetilde{N^{\mathcal{A}}}$ and $\widetilde{N^{\mathcal{B}}}$ leads to a unique profit and social welfare. To see this, note that, given a fixed side $\mathcal{B}$ participation rate $\widetilde{N^{\mathcal{B}}}$, there is a well-defined demand function determining the number of users who participate on side $\mathcal{A}$ as a function of $P^{\mathcal{A}}$, the equilibrium price to side $\mathcal{A}$. Visually, this is depicted in Figure 2, where the set of side $\mathcal{A}$ users participating

[8] Note that the assumption of twice-continuous differentiability actually rules out all of the one-dimensional cases. As the online Appendix shows, the only assumption needed for the analysis is smoothness with regard to movements of the boundary of a set; that is sufficient conditions for the Leibnitz Integral Rule to apply.

[9] Note this example does not fit my full support and massless assumption, but an analogous example that does can be constructed by perturbing it.

when $N^{\mathcal{B}} = 0.5$ and $P^{\mathcal{A}} = 5$ or 6 is shaded. Clearly participation on side $\mathcal{A}$, holding fixed side $\mathcal{B}$ participation, declines in $P^{\mathcal{A}}$. Formally for either $\mathcal{I}$

$$(1) \qquad N^{\mathcal{I}}(P^{\mathcal{I}},\widetilde{N^{\mathcal{J}}}) \equiv \int_{-\infty}^{\infty}\int_{P^{\mathcal{I}}-b^{\mathcal{I}}\widetilde{N^{\mathcal{J}}}}^{\infty} f^{\mathcal{I}}(B^{\mathcal{I}},b^{\mathcal{I}})dB^{\mathcal{I}}db^{\mathcal{I}}.$$

Clearly[10] $N_1^{\mathcal{I}} \equiv \partial N^{\mathcal{I}}/\partial P^{\mathcal{I}} < 0$. Therefore inverting it with respect to its first argument yields a well-defined function $P^{\mathcal{I}}(N^{\mathcal{I}},N^{\mathcal{J}})$. Thus there is a unique pair of prices, and therefore profit and welfare, $P^{\mathcal{A}}(\widetilde{N^{\mathcal{A}}},\widetilde{N^{\mathcal{B}}})$ and $P^{\mathcal{B}}(\widetilde{N^{\mathcal{B}}},\widetilde{N^{\mathcal{A}}})$ consistent with $\widetilde{N^{\mathcal{A}}}$ users participating on side $\mathcal{A}$, $\widetilde{N^{\mathcal{B}}}$ users participating on side $\mathcal{B}$, and uniform pricing.

Thus multiplicity plays no role if one thinks of the platform as simply choosing an allocation[11] to maximize some objective function. The only concern is that the platform may struggle to consistently implement its desired allocation; there could be a "failure to launch" as a result of a "critical mass problem," in the terminology of David S. Evans and Richard Schmalensee (2009).

This can be avoided, however, by a conscientious platform. The platform can lower (raise) its price[12] on side $\mathcal{I}$ when hoped-for (undesirable) side $\mathcal{J}$ users that are valuable (harmful) to marginal users on side $\mathcal{I}$ fail to show up. This insulates the platform's allocation on side $\mathcal{I}$ from the influence of side $\mathcal{J}$ participation. At the logical extreme the platform can ensure that $\widetilde{N^{\mathcal{I}}}$ users participate on side $\mathcal{I}$ regardless of side $\mathcal{J}$ decisions if it charges the *insulating tariff*[13] $P^{\mathcal{I}}(N^{\mathcal{J}}) \equiv P^{\mathcal{I}}(\widetilde{N^{\mathcal{I}}},N^{\mathcal{J}})$. Then the unique equilibrium is the platform's target allocation.[14]

Insulating tariffs are intuitive in many applications. With homogeneous interactions values (Armstrong), the insulating tariff is an insurance scheme, as originally proposed by Phillip H. Dybvig and Chester S. Spatt (1983) for a one-sided market. The platform charges users a price $b^{\mathcal{I}}$ per user on side $\mathcal{J}$ and charges an "hedonic" price (in Armstong's language) which determines participation. Therefore side $\mathcal{I}$ users are indifferent to $N^{\mathcal{J}}$. With no membership values (RT 2003), the insulating tariff is a pure interaction price $p^{\mathcal{I}}$ so that any side $\mathcal{I}$ user earns utility $(b_i^{\mathcal{I}} - p^{\mathcal{I}})N^{\mathcal{J}}$ from participating. Thus participating side $\mathcal{I}$ users prefer high side $\mathcal{J}$ participation and thus are not insured but, because the sign of their utility is independent of side $\mathcal{J}$ participation, still choose to participate independent of the decisions of side $\mathcal{J}$ users. In general, the composition, but not level, of participants may shift with participation on the other side: a rise in $N^{\mathcal{J}}$ selects $\mathcal{I}$ users with high interaction values.

[10] This follows from my assumption of full support.

[11] This approach, which is the key method used throughout the paper to simplify the complexities of pricing in two-sided markets, was first suggested to me in the context of the RT (2003) model by Jeremy Bulow, to whom I am tremendously grateful. Because of the single dimensionality of user heterogeneity in that model, the allocation approach is not much simpler than the price approach there. This led me, much to my later regret, to ignore Jeremy's advice until after having wasted months trying to implement the price approach. On this, as many other matters, I have come round to seeing the elegance of his perspective. My approach was also inspired by the fulfilled expectations equilibrium of Michael L. Katz and Carl Shapiro (1985) and more broadly by the contract theory literature, starting with Roger B. Myerson (1981). It is surprising that, given the long history of the allocation approach in related literatures, it has not to my knowledge been applied previously to the general analysis of network industries.

[12] A further assumption of the RT (2006) model, not discussed extensively above, is that a (uniform) price can be set at any desired level on both sides of the market costlessly. This assumption fails in broadcast media, as argued by Anderson and Coate (2005). They explore, in a Hybrid model (see Section V) with RT (2003) preferences for advertisers and Armstrong preferences for consumers, the consequences of this price rigidity. A more general analysis of restrictions on pricing is, as discussed in Subsection VIA, an important direction for future research.

[13] I am grateful to Bruno Jullien for helping to guide me towards this name.

[14] Note that the platform can charge an insulating tariff on just one side $\mathcal{I}$ of the market and achieve the same guarantee, as this assures that any equilibrium must have $\widetilde{N^{\mathcal{I}}}$ users, removing expectations from the decision making of $\mathcal{J}$ users. This is what makes possible Anderson and Coate's (2005). analysis: they assume the platform chooses quantity, rather than price, to advertisers (effectively assuming an insulating tariff to one side). See footnote 27 for a more general discussion.

Schemes resembling insulating tariffs are used explicitly in many industries: Web site ad rates are typically per click and credit card fees/incentives per transaction. In fact, in broadcasting, as Anderson and Coate (2005) argue, the structure of programming often allows platforms to commit to a quantity of advertising directly. However, even when such explicit schemes are not used, the static RT (2006) model can reasonably be thought of as a reduced form for a dynamic model, in the spirit of Evans and Schmalensee (2009). In this case insulating tariffs simply require that the platform provides subsidies at early stages of product development which it recoups once its desired allocation is achieved. This pattern is commonly observed in video games, operating systems and Web sites.

However, there may be some circumstances under which firms would refrain or be constrained from employing them; see Section VII for further discussion. In these cases the critical mass problem binds and the coordination problems considered by Glenn Ellison and Drew Fudenberg (2003); Atilla Ambrus and Rossella Argenziano (2009); and Evans and Schmalensee (2009) become important.[15] However, in most mature industries, the focus of the RT (2006) model, price flexibility is sufficient to avoid these problems.[16]

Many other tariffs achieve the platform's desired allocation, even uniquely. In fact, as argued by RT (2003), RT (2006), and Armstrong, any tariff with $P^{\mathcal{I}}(\widetilde{N^{\mathcal{J}}}) = P^{\mathcal{I}}(\widetilde{N^{\mathcal{I}}},\widetilde{N^{\mathcal{J}}})$ for both I has the pair $(\widetilde{N^{\mathcal{A}}},\widetilde{N^{\mathcal{B}}})$ as *an* equilibrium. Thus none of my analysis, except a brief discussion of competition in Subsection VIC, assumes any particular tariff. Rather, this subsection is meant to justify my approach of ignoring the specifics of tariffs and coordination and to show, perhaps surprisingly, that adding optimization *simplifies* the analysis. Thus even a reader skeptical of the possibility of insulating tariffs but willing to focus, exogenously, on a given equilibrium, should accept my analysis in the monopoly case I focus on.

II. Pricing

Industrial policy typically aims to alleviate the social harms caused by market power. The first step towards formulating such policy is therefore understanding the nature of those harms. Towards that goal, this section develops and compares the socially optimal and profit maximizing allocation rules, emphasizing the prices that support these allocations.

A. *Pigouvian Pricing*

The value created by the platform is the benefits it brings to users less the costs of providing the service. RT (2006) assumes marginal costs constant in both participation rates, taking the other participation rate as given. Thus there may be two types of cost: membership costs $C^{\mathcal{I}}N^{\mathcal{I}}$ and interaction costs $cN^{\mathcal{A}}N^{\mathcal{B}}$. The benefits the platform brings to users on side $\mathcal{I}$ are

$$V^{\mathcal{I}}(N^{\mathcal{I}},N^{\mathcal{J}}) = \int_{-\infty}^{\infty}\int_{P^{\mathcal{I}}(N^{\mathcal{I}},N^{\mathcal{J}})-b^{\mathcal{I}}N^{\mathcal{J}}}^{\infty}\left[B^{\mathcal{I}} + b^{\mathcal{I}}N^{\mathcal{J}}\right]f^{\mathcal{I}}\left(B^{\mathcal{I}},b^{\mathcal{I}}\right)dB^{\mathcal{I}}db^{\mathcal{I}}. \tag{2}$$

[15] However, I suspect that explicitly modeling why critical mass problems exist in a given application, ignored in previous work on this problem, would be crucial to understanding their welfare implications. For example, if imperfect information is the cause, platforms and social welfare might actually be harmed by attempts to "solve" the deliberately designed critical mass problem!

[16] An identical argument clearly applies to one-sided networks, or any coordination game. I thus believe that the importance of coordination has been exaggerated in situations when an optimizing agent with the ability to make transfers can regulate coordination. Even in the cases when it is relevant, I believe it is more a choice than a constraint. However, this is obviously a controversial view. See, for example, Joseph Farrell and Paul D. Klemperer (2007) for a well-argued contrasting view and the last paragraph of the paper for further discussion.

Thus the total social value of the platform is

$$V(N^{\mathcal{A}},N^{\mathcal{B}}) = V^{\mathcal{A}}(N^{\mathcal{A}},N^{\mathcal{B}}) + V^{\mathcal{B}}(N^{\mathcal{B}},N^{\mathcal{A}}) - C^{\mathcal{A}}N^{\mathcal{A}} - C^{\mathcal{B}}N^{\mathcal{B}} - cN^{\mathcal{A}}N^{\mathcal{B}}. \tag{3}$$

A benevolent social planner equates marginal social benefits to their marginal social costs:

$$V_1^{\mathcal{I}} + V_2^{\mathcal{J}} = C^{\mathcal{I}} + cN^{\mathcal{J}}$$

where $V_i^{\mathcal{I}}$ is the derivative of $V^{\mathcal{I}}$ with respect to the ith argument. $V_1^{\mathcal{I}} = P^{\mathcal{I}}$ as the user added on side $\mathcal{I}$ must be marginal and therefore earn zero net surplus from participating. $V_2^{\mathcal{J}}$ is the value an additional user on side $\mathcal{I}$ brings to users on side $\mathcal{J}$: $\overline{b^{\mathcal{J}}}N^{\mathcal{J}}$, where

$$\overline{b^{\mathcal{J}}} = \frac{\int_{-\infty}^{\infty}\int_{P^{\mathcal{J}}(N^{\mathcal{J}},N^{\mathcal{I}})-b^{\mathcal{J}}N^{\mathcal{I}}}^{\infty} b^{\mathcal{J}} f^{\mathcal{I}}\left(B^{\mathcal{J}},b^{\mathcal{J}}\right) dB^{\mathcal{J}} db^{\mathcal{J}}}{\int_{-\infty}^{\infty}\int_{P^{\mathcal{J}}(N^{\mathcal{J}},N^{\mathcal{I}})-b^{\mathcal{J}}N^{\mathcal{I}}}^{\infty} f^{\mathcal{I}}\left(B^{\mathcal{J}},b^{\mathcal{J}}\right) dB^{\mathcal{J}} db^{\mathcal{J}}}$$

is the *average interaction value of participating users* on side $\mathcal{J}$. Thus the optimal price is

$$P^{\mathcal{I}} = \underbrace{C^{\mathcal{I}} + cN^{\mathcal{J}}}_{\text{marginal private cost}} - \underbrace{\overline{b^{\mathcal{J}}}N^{\mathcal{J}}}_{\text{marginal external benefit}}. \tag{4}$$

This is the standard Pigouvian[17] condition: the price of an activity should equal its private cost less any external benefits. This last term is the essential difference between optimal pricing in two-sided markets and standard multiproduct pricing: because network effects are external to individual decisions, price should diverge from cost. Thus positive network effects should be subsidized and negative ones taxed.

Newspapers offer a simple example. Optimal pricing calls for readers to be subsidized, below the cost of providing the news by the value they bring to advertisers, and for advertisers to be taxed, above the cost of printing their ads, by the amount readers dislike them.

In the Armstrong model, interaction values are homogeneous ($b_i^{\mathcal{I}} \equiv b^{\mathcal{I}}$) and interaction costs are disallowed ($c = 0$) so (4) becomes Alex Gaudeul and Bruno Jullien's (2008) formula

$$P^{\mathcal{I}} = C^{\mathcal{I}} - b^{\mathcal{J}}N^{\mathcal{J}}.$$

RT (2003) rule out membership values/costs so user prices and surplus are all from interactions. Letting $p^{\mathcal{I}} \equiv P^{\mathcal{I}}/N^{\mathcal{I}}$ be the *per-interaction price* and $\overline{s^{\mathcal{I}}} \equiv (V^{\mathcal{I}}/N^{\mathcal{J}}) - p^{\mathcal{I}}$ the average per-interaction surplus on side $\mathcal{I}$ gives Bedre-Defolie and Calvano's (2010) and Weyl's (2009b) optimal pricing rule

$$p^{\mathcal{A}} + p^{\mathcal{B}} - c = -\overline{s^{\mathcal{A}}} = -\overline{s^{\mathcal{B}}}\,.$$

I now compare this classical rule to that which a profit-maximizing monopolist would adopt.

[17] First-best pricing has traditionally been known in the literature as Lindahl pricing (Özlem Bedre-Defolie and Emilio Calvano 2010; Weyl 2009b). However, because price discrimination is ruled out in the RT (2006) model, pricing follows Pigou (1920) rather than Lindahl (1919).

B. *Profit-Maximizing Pricing*

Often the operators of platforms are concerned with their profits rather than with social welfare. Because price discrimination is typically imperfect, these differ. To make as clear as possible the distortions introduced by imperfect price discrimination it is useful to take them to their logical extreme, ruling out all discrimination.[18] Profits are then

$$\pi(N^{\mathcal{A}},N^{\mathcal{B}}) = (P^{\mathcal{A}}[N^{\mathcal{A}},N^{\mathcal{B}}] - C^{\mathcal{A}})N^{\mathcal{A}} + (P^{\mathcal{B}}[N^{\mathcal{B}},N^{\mathcal{A}}] - C^{\mathcal{B}})N^{\mathcal{B}} - cN^{\mathcal{A}}N^{\mathcal{B}}. \tag{5}$$

A profit-maximizing monopolist equates marginal revenues of participation to marginal cost:

$$\underbrace{P^{\mathcal{I}} + P_1^{\mathcal{I}}N^{\mathcal{I}} + P_2^{\mathcal{J}}N^{\mathcal{J}}}_{\text{marginal revenue}} = \underbrace{C^{\mathcal{I}} + cN^{\mathcal{J}}}_{\text{marginal cost}}.$$

The first two terms of marginal revenue are classical: price minus the inverse hazard rate of demand (or *market power*) $\mu^{\mathcal{I}} \equiv -P_1^{\mathcal{I}}N^{\mathcal{I}} = P^{\mathcal{I}}/\epsilon^{\mathcal{I}}$ where $\epsilon^{\mathcal{I}}$ is the elasticity of demand. The final term is special to two-sided markets: it is the revenue that can be extracted from side $\mathcal{J}$ by adding an additional user on side $\mathcal{I}$. Letting $\widetilde{b^{\mathcal{J}}}$ be the *average interaction value of marginal users* (AIVMU) on side $\mathcal{J}$, by the implicit function theorem and equation (1)

$$P_2^{\mathcal{J}} = -\frac{N_2^{\mathcal{J}}}{N_1^{\mathcal{J}}} = \frac{\int_{-\infty}^{\infty} b^{\mathcal{J}} f^{\mathcal{J}}\Big(P^{\mathcal{J}}[N^{\mathcal{J}},N^{\mathcal{I}}] - b^{\mathcal{J}}N^{\mathcal{I}}, b^{\mathcal{J}}\Big)db^{\mathcal{J}}}{\int_{-\infty}^{\infty} f^{\mathcal{J}}\Big(P^{\mathcal{J}}[N^{\mathcal{J}},N^{\mathcal{I}}] - b^{\mathcal{J}}N^{\mathcal{I}}, b^{\mathcal{J}}\Big)db^{\mathcal{J}}} \equiv \widetilde{b^{\mathcal{J}}}.$$

The platform can extract only the value *marginal* users on side $\mathcal{J}$ place on an additional side $\mathcal{I}$ user joining. This is an example of the general tendency, emphasized by Spence (1975) and discussed extensively below, of monopolists to serve the preferences of marginal, rather than all participating, users. The platform's side $\mathcal{J}$ marginal revenue from a side $\mathcal{I}$ user is therefore $\widetilde{b^{\mathcal{J}}}N^{\mathcal{J}}$. Privately optimal pricing follows a simple extension of Lerner's formula[19]

$$\frac{P^{\mathcal{I}} - (C^{\mathcal{I}} + cN^{\mathcal{J}} - \widetilde{b^{\mathcal{J}}}N^{\mathcal{J}})}{P^{\mathcal{I}}} = \frac{1}{\epsilon^{\mathcal{I}}}. \tag{6}$$

In the Armstrong case this immediately simplifies to Armstrong's pricing condition

$$P^{\mathcal{I}} = C^{\mathcal{I}} - b^{\mathcal{J}}N^{\mathcal{J}} + \mu^{\mathcal{I}}.$$

In RT (2003), only interaction benefits exist so $\widetilde{b^{\mathcal{J}}} = p^{\mathcal{J}}$. Therefore the pricing condition is

$$p^{\mathcal{A}} + p^{\mathcal{B}} - c = m^{\mathcal{A}} = m^{\mathcal{B}}$$

where $m^{\mathcal{I}} \equiv \mu^{\mathcal{I}}/N^{\mathcal{J}}$. This is the formula that RT (2003) derives.

[18] For an analysis of distortions that can arise even under perfect price discrimination and with a single group of homogeneous users, when there are externalities to nonparticipating consumers or other frictions, see Ilya Segal (1999).

[19] RT (2006) states the general condition for optimal two-product pricing in terms of derivatives of $N^{\mathcal{A}}$ and $N^{\mathcal{B}}$, determined as fixed points of an equilibrium among users. However, as a function of the allocation $(N^{\mathcal{A}}, N^{\mathcal{B}})$, profits are just the simple explicit function above. This is what allows me to express the first-order condition for optimal in terms of the primitive properties of preferences in two-sided markets.

Comparing private to socially optimal pricing,

$$(7)\qquad P^{\mathcal{I}} = \underbrace{C^{\mathcal{I}} + cN^{\mathcal{J}} - \overline{b^{\mathcal{J}}}N^{\mathcal{J}}}_{\text{socially optimal price}} + \underbrace{\mu^{\mathcal{I}}}_{\text{classical market power distortion}} + \underbrace{(\overline{b^{\mathcal{J}}} - \widetilde{b^{\mathcal{J}}})N^{\mathcal{J}}}_{\text{Spence distortion}}.$$

Thus there are two distortions in a two-sided market. First, classical marginal revenue lies below price by the amount of the market power $\mu^{\mathcal{I}}$. Second, if $\widetilde{b^{\mathcal{J}}} \neq \overline{b^{\mathcal{J}}}$, the average interaction values of marginal users differ from those of loyal users; the platform will either over- or undersubsidize (tax) users on side $\mathcal{I}$. Like the classical market power distortion, this *Spence distortion* is a consequence of the platform's inability to price discriminate. The platform internalizes network externalities but does so imperfectly (see Subsection IIIC).

The tendency to truckle to marginal users is familiar to anyone living in, or observant while visiting,[20] a tourist destination: the city government and businesses tend to cater to mobile tourists rather than to locked-in residents.[21] This Spence distortion is likely more important in two-sided markets than the contexts for which it was originally conceived. A platform is unlikely to partially ameliorate inefficiency (while introducing other distortions) by offering multiple products (Michael Mussa and Sherwin Rosen 1978; Mary O'Keeffe, W. Kip Viscusi, and Richard J. Zeckhauser 1984; David Besanko, Shabtai Donnenfeld, and Lawrence J. White 1987) as this would require inefficiently wasting potential interactions.[22] Once "quality" is provided to some users on one side of the market, it is free to provide to others.

The existence and sign of the Spence distortion depend crucially on the source of user heterogeneity.[23] $\overline{b^{\mathcal{I}}}$ will tend to exceed $\widetilde{b^{\mathcal{I}}}$ if users differ primarily in their interaction values. For example, in the extreme case of only interaction heterogeneity (RT 2003), the Spence distortion is exactly the per-interaction surplus on side $\mathcal{I}$, while when there is only membership heterogeneity (Armstrong), there is no Spence distortion. The Spence distortion may even be downward, as in the newspaper example above. If heterogeneity in willingness to pay for content dominates and is correlated with willingness to pay to avoid advertising, then loyal users dislike advertising more than marginals, and the Spence distortion is downwards.

[20] Of course in real life, as in the RT (2006) model, marginal users (tourists) are a heterogeneous bunch, and many, including the author, have preferences more similar to natives than to those of other tourists.

[21] Readers living in less-frequented climes may find a joke instructive. I tell a variant of a classic Israeli joke, first told to me by David Hariton, to whom I am grateful. In the original joke, Smith is replaced by David Ben-Gurion.

Adam Smith dies and, for his service to economics, is given a choice of where to spend eternity. He requests to see each option before deciding. First he is shown Hell, which, full of decadent food, French wines and beautiful women, seems a merry way to spend the rest of time. Heaven, on the other hand, is an unending stream of presentations of leading research in economics and philosophy. Having spent his life in contemplation on these topics, Smith decides he has earned a bit of relaxation in the afterlife and opts for Hell. Immediately he is thrown onto the rack, whipped, waterboarded, and subjected to other "enhanced" methods of entertainment. Astonished, he says, "I was just here a few minutes ago and things were so much nicer. What happened?" Lucifer replies, "Then you were a tourist."

[22] If the incentive for price discrimination is sufficiently large the platform might "throw away" quality. While such strategies are common in standard markets, in two-sided markets they seem to occur only when justified by other concerns outside this model, such as optimal matching (e.g., targeted ads). However, this is an important question for future research.

[23] Another, perhaps more general way to put this follows the language of Spence more closely. Spence argued quality would be undersupplied ($P^{\mathcal{J}}$ distorted upward) when $P^{\mathcal{I}}_{12} < 0$ and oversupplied ($P^{\mathcal{J}}$ distorted downward) when $P^{\mathcal{I}}_{12} > 0$. If, as in the RT (2006) model, each user can participate at most once, the former is equivalent to users with high utility (lower reservation values) having high sensitivity to quality and users with low utility (high reservation values) being less so; the latter conversely. Note that $P^{\mathcal{I}}_{12} = -\mu^{\mathcal{I}}_{b}/N^{\mathcal{I}}$, the measure of local interaction heterogeneity I develop in Section IV. Thus there is a one-to-one correspondence between Spence's cross-partial of the price function and my focus on user heterogeneity.

Thus the harms of market power depend crucially on the source of heterogeneity. If, as is typically assumed, the costs of price distortions are convex, then market power over card accepters is particularly pernicious as it compounds the Spence distortion from cardholders. However, it may actually be beneficial that the *Times* has market power over advertisers, as this offsets the Spence distortion potentially leading to a better second-best level of advertising. Even with market power, therefore, it is possible ad rates may be too low.

C. *Ramsey Pricing*

Achieving first-best prices may be infeasible in practice at it would require subsidies whose granting, given the cost of raising public funds, political economy constraints, and imperfect information, would be more costly than the monopoly distortions they seek to address. When granting subsidies is infeasible, second-best pricing requires maximizing social welfare subject to some constraint, such as allowing the firm a rate of return (possibly 0) on its variable or fixed costs. Because of the externalities in two-sided markets, this Ramsey solution must be extended as proposed by Tae Hoon Oum and Michael M. Tretheway (1988) to take these into account.

I consider three formulations of the Ramsey problem, all of which are equivalent if the required level of profit is 0. First, in the text, I consider the classic Ramsey problem: social welfare is maximized subject to achieving a minimum absolute profit. In the Appendix, I consider a modified version of the Ramsey problem that RT (2003)[24] uses in a two-sided market where the rate of return is required on variable costs. As I argued in Weyl (2009b), there are two possible social objectives: maximizing user or social surplus subject to the rate-of-return constraint. The first approach addresses externalities more completely, while the second comes closer to the monopolist's constrained goals.

THEOREM 1: *Interior Ramsey prices maximizing user or social surplus subject to the constraint that the platform makes a profit of at least K must solve*

$$\frac{P^{\mathcal{I}} - \left(\overbrace{C^{\mathcal{I}} + cN^{\mathcal{J}} - \widetilde{b^{\mathcal{J}}}N^{\mathcal{J}}}^{\textit{private marginal opportunity cost}} \quad - \quad \overbrace{[1-\lambda]}^{\textit{Lagrangian weighting}} \quad \overbrace{[\overline{b^{\mathcal{J}}} - \widetilde{b^{\mathcal{J}}}]}^{\textit{Spence distortion}}\right)}{P^{\mathcal{I}}} = \lambda\frac{1}{\epsilon^{\mathcal{I}}} \tag{8}$$

where

$$\lambda \equiv \frac{\overbrace{K}^{\textit{target profit}} + \overbrace{(\overline{b^{\mathcal{A}}} + \overline{b^{\mathcal{B}}} - c)N^{\mathcal{A}}N^{\mathcal{B}}}^{\textit{subsidy required for (local) Pigouvian prices}}}{\underbrace{N^{\mathcal{A}}\mu^{\mathcal{A}} + N^{\mathcal{B}}\mu^{\mathcal{B}} + (\overline{b^{\mathcal{A}}} + \overline{b^{\mathcal{A}}} - \widetilde{b^{\mathcal{A}}} - \widetilde{b^{\mathcal{B}}})N^{\mathcal{A}}N^{\mathcal{B}}}_{\textit{(local) profit gain moving to monopoly from Pigouvian prices}}}.$$

[24] Rochet and Tirole use this modified Ramsey set-up to consider whether firms distort the "balance" of prices as separate from their level, a major focus of mine in Weyl (2009c).

PROOF:

See the Ramsey Pricing portion of the Appendix.

Thus the Ramsey pricing condition is just a simple weighted average of the Pigouvian and profit-maximizing prices. These, again, diverge in their attention to both the Spence and classical market power distortions. Prices are closer to profit maximization i) the higher is the target profit, ii) the larger is the subsidy called for by Pigouvian prices and iii) the further one must move towards monopoly to achieve a given gain in net profits. Just as first-best prices take a classic Pigouvian form, Ramsey prices take OT's Pigou-Ramsey form.

III. Generalization

The primary aim of this paper is to understand the price theory of and proper policy towards industries such as payment cards and newspapers. After a brief interlude in this section, I continue toward this goal in Section IV, to which a casual reader may wish to skip directly. However, the general character of my basic ideas thus far suggests they may help analyze a broader class of models than that RT (2006) specifically adapted to those industries. In fact with any number of groups of users and essentially arbitrary heterogeneous preferences, the same principles developed above apply. Insulating tariffs exist, allowing a simple analysis of the platform's choice of allocation showing in general that the Spence distortion is the key element added by network externalities. This section considers such a generalization.

I maintain four important assumptions of the RT (2006) model:

i) (Quasi-linear) user preferences are taken as exogenous (RT 2006 assumption 1).

ii) All groups of users can be explicitly (third-degree) price discriminated and all users within each group differ only in their preferences.[25]

iii) No price discrimination is possible, but prices to any given group can take any positive or negative value. Users interact with an exogenous collection of other users (in their own and other groups); any marginal price for such interactions is exogenous to the model and enters only to the extent that it determines preferences.

iv) Externalities are only to participating users.[26]

A. *The Model*

There are M groups $\mathcal{I} = \mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ and users may value participation by members not only of other groups, but of their own. A typical user i on side $\mathcal{I}$ is characterized by a vector $\theta_i^{\mathcal{I}}$ of characteristics drawn according to a smooth and massless distribution with probability density function

[25] Andre Veiga and Weyl (2010) have made significant progress in relating this assumption.

[26] Unlike the others, this assumption can easily be dispensed with. This generalizes Segal's (1999) basic model of contracting with externalities to allow asymmetric information (on reservation values) and asymmetric agents ("sides of the market"). However, the increase in notational complexity and distance from a realistic model of network industries (most nonparticipant externalities arise in contracting, rather than uniform pricing, settings) led me to this assumption. The intuition of that model should be clear from Segal and the general model here: the social planner internalizes all externalities, while a profit-maximizer internalizes the *reverse sign* of externalities to *marginal* nonparticipating consumers, scaled by the number of participating users. Effectively, to a profit maximizer, negative out-group externalities are equivalent to positive in-group externalities, while to a social planner they are opposite. Details are available on request. This extension and the more general connections between the theory of multi-sided platforms (network industries) and contracting with externalities are promising areas for future research.

$f^{\mathcal{I}}$ with full support on $R^{K^{\mathcal{I}}}$ where $K^{\mathcal{I}} \in N$. Let $\mathbf{N} \equiv (N^{\mathcal{A}}, N^{\mathcal{B}}, N^{\mathcal{C}}, \ldots)$ be an *allocation*, a vector of participation rates on each side.[27] The utility of user i on side $\mathcal{I}$ from participating is

$$U_i^{\mathcal{I}} = u^{\mathcal{I}}(\mathbf{N}; \theta_i^{\mathcal{I}}) - P^{\mathcal{I}}$$

where $P^{\mathcal{I}}$ is the price a user on side $\mathcal{I}$ must pay to participate. I assume that $u^{\mathcal{I}}$ is smooth in and finite for all values of the allocation and characteristics.

Note that this general model has a few special cases of particular interest:

i) $M = 1$ is a one-sided network with arbitrary utility and heterogeneity. I do not believe this model has ever been considered, but given the substantial interest in one-sided network monopolists (Economides 1996) it seems a natural general model.

ii) $M = 2$ is RT (2006) with arbitrary heterogeneous utilities and within-side network effects.

iii) Suppose M sides can be split into two groups $\mathbf{A}$ and $\mathbf{B}$ such that $u^{\mathcal{I}}$ is independent of $N^{\mathcal{J}}$ whenever either $\mathcal{I}, \mathcal{J} \in \mathbf{A}$ or $\mathcal{I}, \mathcal{J} \in \mathbf{B}$. This is (ii) without within-side effects but with groups of discriminable, heterogeneously valuable users on each side.

For a particular allocation $\mathbf{N}$ and price $P^{\mathcal{I}}$ the set of $\mathcal{I}$ users weakly benefiting from participating is $\overline{\Theta^{\mathcal{I}}}(\mathbf{N}, P^{\mathcal{I}}) \equiv \{\theta^{\mathcal{I}} : u^{\mathcal{I}}(\mathbf{N}; \theta^{\mathcal{I}}) \geq P^{\mathcal{I}}\}$ and the set of all marginal users is $\widetilde{\Theta^{\mathcal{I}}}(\mathbf{N}, P^{\mathcal{I}}) \equiv \{\theta^{\mathcal{I}} : u^{\mathcal{I}}(\mathbf{N}; \theta^{\mathcal{I}}) = P^{\mathcal{I}}\}$. Then the fraction of users interested in participating on side $\mathcal{I}$ given an allocation $\mathbf{N}$ and a price $P^{\mathcal{I}}$ is

$$\widetilde{N^{\mathcal{I}}}(P^{\mathcal{I}}, \mathbf{N}) = \int_{\overline{\Theta^{\mathcal{I}}}(P^{\mathcal{I}}, \mathbf{N})} f^{\mathcal{I}}(\theta^{\mathcal{I}}) d\theta^{\mathcal{I}}.$$

Because the set $\overline{\Theta^{\mathcal{I}}}$ is clearly contracting in $P^{\mathcal{I}}$, $\widetilde{N^{\mathcal{I}}} < 0$ and[28] $\widetilde{N^{\mathcal{I}}}$ can be inverted to yield $P^{\mathcal{I}}(\widetilde{N^{\mathcal{I}}}, \mathbf{N})$, the price needed to attract $\widetilde{N^{\mathcal{I}}}$ of users who anticipate allocation $\mathbf{N}$.

Note that the potential multiplicity problem here is far worse than in the two-sided case, as utility functions have arbitrary structure, and there can be an arbitrary number of sides. However, this enormous coordination problem can again be avoided by careful tariffs. In particular the platform may charge an insulating tariff, which is here a price to side $\mathcal{I}$ depending on the full realized allocation that ensures the desired allocation is realized. Formally the insulating tariff for a desired participation rate $\widetilde{N^{\mathcal{I}}}$ is $P^{\mathcal{I}}(\mathbf{N}) \equiv P^{\mathcal{I}}(\widetilde{N^{\mathcal{I}}}, \mathbf{N})$. As with RT (2006), if the platform charges the insulating tariff associated with its desired allocation on all sides, then the unique equilibrium is its desired allocation.[29] Thus once again the platform's problem can be viewed as

[27] In particular, I assume every utility level is achieved by same type, given any $\mathbf{N}$.

[28] When participation is positive, but not total, from my assumption of smooth $f^{\mathcal{I}}$'s and full support.

[29] Here, again, insulating *every* side *from every side* can be dispensed with. Imagine drawing a graph where each node represents a side of the market and a directed edge is drawn between each side and those sides whose participation affects their utility, but against whose participation they are not insulated. I would conjecture, but have only the sketch of a proof that, so long as this graph is acyclic there is a unique equilibrium. Intuitively if the graph is acyclic, one can trace back from its sinks to tie down the unique participation rate on each side. Furthermore other tariffs than the insulating tariff may do the trick for particular (distributions of) user preferences. However, I believe that the "simplest" approach to "robustly" ensuring uniqueness is fully insulating every side of the market from every other side. A formal analysis of all this will likely appear in joint work in progress with Alex White, as also referred to in footnote 53.

one of choosing an allocation $\widetilde{\mathbf{N}}$ to maximize some objective, eliminating the need to consider derivatives of complex, multi-sided fixed points.

B. *Pricing*

Let $P^{\mathcal{I}}(\mathbf{N}) = P^{\mathcal{I}}(N^{\mathcal{I}}, \mathbf{N})$, where $N^{\mathcal{I}}$ is the $\mathcal{I}$th entry of $\mathbf{N}$, $\overline{\Theta^{\mathcal{I}}}(\mathbf{N}) \equiv \overline{\Theta^{\mathcal{I}}}(\mathbf{N}, P^{\mathcal{I}}[\mathbf{N}])$ and $\widetilde{\Theta^{\mathcal{I}}}(\mathbf{N}) \equiv \widetilde{\Theta^{\mathcal{I}}}(\mathbf{N}, P^{\mathcal{I}}[\mathbf{N}])$. The gross value created on side $\mathcal{I}$ by an allocation $\mathbf{N}$ is simply

$$V^{\mathcal{I}}(\mathbf{N}) = \int_{\theta^{\mathcal{I}} \in \overline{\Theta^{\mathcal{I}}}(\mathbf{N})} u^{\mathcal{I}}(\mathbf{N}; \theta^{\mathcal{I}}) f^{\mathcal{I}}(\theta^{\mathcal{I}}) d\theta^{\mathcal{I}}.$$

I allow for arbitrary smooth, positive cost functions $C(\mathbf{N})$. Thus the (net) surplus created by the service as a function of the allocation is

$$V(\mathbf{N}) = \sum_{\mathcal{I}} V^{\mathcal{I}}(\mathbf{N}) - C(\mathbf{N}).$$

Maximizing the surplus created by the service requires equating marginal social value to marginal cost. Let $X_{\mathcal{J}} \equiv \partial X / \partial N^{\mathcal{J}}$. A socially optimal allocation then requires that for each $\mathcal{I}$

$$\sum_{\mathcal{J}} V_{\mathcal{I}}^{\mathcal{J}} = C_{\mathcal{I}}.$$

The following theorem states that these conditions can again be written in a Pigouvian form.

THEOREM 2: *The first-order conditions for a socially optimal allocation are*

$$P^{\mathcal{I}} = \underbrace{C_{\mathcal{I}}}_{\textit{marginal cost}} - \underbrace{\sum_{\mathcal{J}} \overline{u_{\mathcal{I}}^{\mathcal{J}}} N^{\mathcal{J}}}_{\textit{marginal externalities}} \tag{9}$$

where the average marginal interaction value of participating side $\mathcal{I}$ users for side $\mathcal{J}$ users is

$$\overline{u_{\mathcal{J}}^{\mathcal{I}}}(\mathbf{N}) \equiv \frac{\int_{\theta^{\mathcal{I}} \in \overline{\Theta^{\mathcal{I}}}(\mathbf{N})} u_{\mathcal{J}}^{\mathcal{I}}(\mathbf{N}; \theta^{\mathcal{I}}) f^{\mathcal{I}}(\theta^{\mathcal{I}}) d\theta^{\mathcal{I}}}{\int_{\theta^{\mathcal{I}} \in \overline{\Theta^{\mathcal{I}}}(\mathbf{N})} f^{\mathcal{I}}(\theta^{\mathcal{I}}) d\theta^{\mathcal{I}}}.$$

PROOF:

See the Generalization portion of the Appendix.

Thus the Pigouvian formula (4) extends in the most natural way possible: interaction values are replaced by the *marginal* value of users who have potentially nonlinear utility and all externalities, to those within side $\mathcal{I}$ and on other sides $\mathcal{J} \neq \mathcal{I}$, are included.

On the other hand revenues on side $\mathcal{I}$ are $R^{\mathcal{I}}(\mathbf{N}) = P^{\mathcal{I}}(\mathbf{N}) N^{\mathcal{I}}$ and profits

$$\pi(\mathbf{N}) = \sum_{\mathcal{I}} R^{\mathcal{I}}(\mathbf{N}) - C(\mathbf{N}).$$

Profit maximization requires equating marginal revenue of an additional side $\mathcal{I}$ user, from all sides of the market, to the marginal cost of serving that user:

$$\sum_{\mathcal{J}} R_{\mathcal{I}}^{\mathcal{J}} = C_{\mathcal{I}}.$$

This yields a similarly intuitive extension of the RT (2006) profit maximizing pricing.

THEOREM 3: *The first-order conditions for a profit maximizing allocation are*

$$\frac{P^{\mathcal{I}} - \left(C_{\mathcal{I}} - \sum_{\mathcal{J}} \widetilde{u_{\mathcal{I}}^{\mathcal{J}}} N^{\mathcal{J}}\right)}{P^{\mathcal{I}}} = \frac{1}{\epsilon^{\mathcal{I}}} \tag{10}$$

or equivalently

$$P^{\mathcal{I}} = \underbrace{C_{\mathcal{I}} - \sum_{\mathcal{J}} \overline{u_{\mathcal{I}}^{\mathcal{J}}} N^{\mathcal{J}}}_{\textit{socially optimal price}} + \underbrace{\mu^{\mathcal{I}}}_{\textit{classical market power distortion}} + \underbrace{\sum_{\mathcal{J}} \left(\overline{u_{\mathcal{I}}^{\mathcal{J}}} - \widetilde{u_{\mathcal{I}}^{\mathcal{J}}}\right) N^{\mathcal{J}}}_{\textit{Spence distortion}(s)} \tag{11}$$

where the average marginal interaction value of marginal side $\mathcal{I}$ users for side $\mathcal{J}$ *users is*

$$\widetilde{u_{\mathcal{J}}^{\mathcal{I}}}(\mathbf{N}) \equiv \frac{\int_{\theta^{\mathcal{I}} \in \widetilde{\Theta^{\mathcal{I}}}(\mathbf{N})} u_{\mathcal{J}}^{\mathcal{I}}(\mathbf{N}; \theta^{\mathcal{I}}) f^{\mathcal{I}}(\theta^{\mathcal{I}}) d\theta^{\mathcal{I}}}{\int_{\theta^{\mathcal{I}} \in \widetilde{\Theta^{\mathcal{I}}}(\mathbf{N})} f^{\mathcal{I}}(\theta^{\mathcal{I}}) d\theta^{\mathcal{I}}}.$$

PROOF:

See the Appendix.

Thus again profit maximization distorts the allocation in two ways. First it raises prices (lowers participation) as classical marginal revenue falls below price. Second it imperfectly internalizes network externalities, as preferences of marginal rather than all participating users determine the marginal revenues generated by an additional side $\mathcal{I}$ user. Thus there are now M classical distortions and M^2 Spence distortions.

C. *Discussion*

Conceptually little has changed from the RT (2006) model. Insulating tariffs exist and the platform can therefore achieve any desired allocation. The platform thus maximizes its objectives over possible allocations, making its problem simple. Profit maximization leads to classical and Spence distortions. The nature of these latter distortions depends on how the preferences of loyal and marginal users diverge, that is, on the source of user heterogeneity.

This suggests three interesting conclusions. First, while most of this paper focuses on affine user preferences, this is without significant loss of generality. While affine preferences allow only two dimensions of heterogeneity, these are two dimensions which generally matter. This extends even to my comparative statics analysis below, as none of the terms governing these include the curvature of utility (none involve $V_{\mathcal{J}\mathcal{J}}^{\mathcal{I}}$). Of course the irrelevance of further dimensions of heterogeneity depends crucially on the impossibility of price discrimination. If user utility is not affine, platforms may use a marginal price, such as differential charges for viewing certain Web sites, to discriminate among users. In this case social value and profits depend not only on participation rates, but also on marginal prices. This important and largely open[30] problem is well beyond the scope of this paper.

Second, it provides a simple and general strategy for analyzing monopoly networks: the allocation approach. While my results here constitute only the most superficial of first passes, having no comparative static or policy analysis, they suggest a path for future research.

Finally, it answers perhaps the oldest open question in network economics: the general validity of the (Liebowitz and Margolis 1994) conjecture that optimizing networks will internalize (and thereby neutralize) network externalities. Liebowitz and Margolis were partially correct, but only up to a point. While platforms do internalize externalities, they do so imperfectly as they take into account only the preferences of marginal users. This Spence distortion will be smallest, and therefore network externalities most nearly neutralized, when loyal and marginal users place a similar value on the participation of a marginal network user. In this case, the only distortions caused by market power are the classic, familiar ones of any multiproduct monopolist. In this case actions of users influence the welfare of other users only indirectly, through platform optimization (see Subsection IVB). On other hand, when loyal and marginal users have sharply different preferences, network monopolies have much more complex distortions with large direct network externalities persisting in equilibrium.

IV. Comparative Statics

A primary motivation for the theory of two-sided markets is that conditions on each side affect participation and welfare on the other. As with any comparative statics exercise, understanding these indirect cross-effects relies on the second-order conditions for optimization (Paul A. Samuelson 1941) and specifically, because of the multiproduct (Weyl 2009a) monopoly (Weyl and Michal Fabinger 2009) context, on pass-through rates and the cross partial of the allocation in profits. I begin by formally developing these closely related concepts.

The *pass-through rate* on side $\mathcal{I}$,

$$\rho^{\mathcal{I}} \equiv \left.\frac{dP^{\mathcal{I}}}{dC^{\mathcal{I}}}\right|_{N^{\mathcal{J}}} = -\frac{\mu^{\mathcal{I}}}{N^{\mathcal{I}}\frac{\partial^2 \pi}{\partial N^{\mathcal{I}^2}}}$$

is the amount a private platform finds it optimal to increase $P^{\mathcal{I}}$ in response to an increase in $C^{\mathcal{I}}$ if $N^{\mathcal{J}}$ is held fixed. The cross partial

$$\chi \equiv \frac{\partial^2 \pi}{\partial N^{\mathcal{A}} \partial N^{\mathcal{B}}}$$

measures the complementarity/substitutability (if positive/negative) of participation rates.

For traditional comparative static analysis, it is important that the first-order conditions used actually represent the optimal allocation for the platform. To ensure this, a convenient assumption is that the platform's profit function is concave. However, it is important to avoid overly restrictive assumptions that are sufficient, but unnecessary, for the purpose as these can bias analysis; log-concavity is a typical such assumption. To add tractability without undue restrictions, I

[30] See Bedre-Defolie and Calvano (2008) for a first pass, in the context of the RT (2003) model.

propose a "weak" (in a sense formalized by Theorem 4) second-order condition. As far as I know this *two-sided contraction* (2SC) is the first second-order condition to be proposed for the general RT (2006) model.

If for all $(N^{\mathcal{A}}, N^{\mathcal{B}}) \in (0,1)^2, \rho^{\mathcal{A}}, \rho^{\mathcal{B}} > (\geq) 0$ and $(\mu^{\mathcal{A}}\mu^{\mathcal{B}}/N^{\mathcal{A}}N^{\mathcal{B}}) > (\geq)\rho^{\mathcal{A}}\rho^{\mathcal{B}}\chi^2$, I will say that f satisfies strict (weak) *two-sided contraction* (2SC) given interaction cost c.

THEOREM 4: *If $f^{\mathcal{A}}, f^{\mathcal{B}}$, and c exhibit strict* 2*SC then for any $C^{\mathcal{A}}$ and $C^{\mathcal{B}}$ a solution to equation* (6) *for both I is the unique platform's optimal price. If they violate weak* 2*SC then there exists a pair* $(C^{\mathcal{A}}, C^{\mathcal{B}})$ *for which there is a solution to equation* (6) *which is not an optimum.*

PROOF:

See Appendix.

In the RT (2003) case,[31] $\chi = \mu^{\mathcal{I}}/N^{\mathcal{J}}$ for both *I* so the condition becomes my (Weyl 2009c) "cross-subsidy contraction" condition $\rho^{\mathcal{A}}\rho^{\mathcal{B}} < 1$. The Comparative Statics portion of the Appendix gives Pigouvian SOCs. These could be extended to the general model of Section III by deriving conditions for the Hessian matrix of cross partials of profits with respect to the allocation to be negative definite.

A. *Complements versus Substitutes*

The most famous, supposedly robust result on the comparative statics of two-sided markets is what RT (2006) calls the "simple 'seesaw principle': a factor that is conducive to a high price on one side, to the extent that it raises the platform's margin on that side, tends also to call for a low price on the other side as attracting members on that other side becomes more profitable." While intuitive, this result faces two challenges. First, the appropriate notion of "price" is unclear. In the only (RT 2003) model where the seesaw principle has been demonstrated formally (RT 2003; Weyl 2009c), the price is per interaction. In other models this price has little special significance. However, as discussed above, holding fixed the number of users on side $\mathcal{I}$ the price (in any sense) on side $\mathcal{J}$ is decreasing in the number of users on side $\mathcal{J}$. Therefore RT (2006)'s seesaw principle can be reformulated as follows: factors leading the platform to choose higher $N^{\mathcal{I}}$ lead it to choose lower $N^{\mathcal{J}}$. That is $\partial^2\pi/\partial N^{\mathcal{A}}\partial N^{\mathcal{B}} < 0$ or participation on the two sides are substitutes *for the platform*.[32] In RT (2003), this holds, and the two formulations are equivalent. However, this formulation can be examined beyond the context of the RT (2003) model.

The more serious challenge to the seesaw principle is that this broader formulation is not generally true but instead depends on the source of user heterogeneity.

To see this formally, it is useful to construct a general measure of the local importance of the two dimensions of heterogeneity. A natural such measure is how interaction and membership benefits of marginal users increase with price. Price is, by definition, always equal to the total value of marginal users. It is therefore natural to decompose increases in price into changes in interaction and membership values. From Subsection IIB $P_1^{\mathcal{I}} = -\mu^{\mathcal{I}}/N^{\mathcal{I}}$; but the total gross utility of a marginal user is $\widetilde{B^{\mathcal{I}}} + \widetilde{b^{\mathcal{I}}}N^{\mathcal{J}}$ so

$$\widetilde{B_1^{\mathcal{I}}} + \widetilde{b_1^{\mathcal{I}}}N^{\mathcal{J}} = -\frac{\mu^{\mathcal{I}}}{N^{\mathcal{I}}}.$$

[31] See below. Also see the online Appendix for the Armstrong special case.

[32] Note that the "demand system" does not necessarily exhibit either complements or substitutes: Slutsky symmetry is not obeyed ($\widetilde{b^{\mathcal{I}}} = P_2^{\mathcal{I}} \neq P_2^{\mathcal{J}} = \widetilde{b^{\mathcal{J}}}$) and may even be violated in signs, despite quasi-linearity, because of the externalities between the sides.

We can therefore define natural measures of local heterogeneity along the two dimensions as the projection of market power onto each dimension. $\mu^{\mathcal{I}}_{\tilde{\mathcal{B}}} \equiv -\widetilde{b^{\mathcal{I}}_1} N^{\mathcal{I}}$ is the *membership market power* and $\mu^{\mathcal{I}}_{\tilde{b}} \equiv -\widetilde{b^{\mathcal{I}}_1} N^{\mathcal{A}} N^{\mathcal{B}}$ is the *interaction market power*. The cross partial

$$(12) \qquad \chi = \widetilde{b^{\mathcal{A}}} + \widetilde{b^{\mathcal{B}}} - c - \frac{\mu^{\mathcal{A}}_{\tilde{b}}}{N^{\mathcal{B}}} - \frac{\mu^{\mathcal{B}}_{\tilde{b}}}{N^{\mathcal{A}}}$$

as the effect of side $\mathcal{J}$ participation on $\mathcal{I}$ marginal revenue is the difference between its effect on price $P^{\mathcal{I}}$, $\widetilde{b^{\mathcal{I}}}$, and market power $\mu^{\mathcal{I}}$, as shown formally in the proof of Theorem 5 below. Intuitively, interaction benefits favor complementarity: the value of a side $\mathcal{A}$ user is proportional to the number of users she interacts with on side $\mathcal{B}$. Thus an increase in side $\mathcal{B}$ users makes it more attractive to recruit side $\mathcal{A}$ users. Offsetting this is the fact that when interaction benefits are the main dimension of heterogeneity, increasing the participation on side $\mathcal{B}$ requires recruiting low interaction benefit users. Thus increased side $\mathcal{B}$ participation reduces the AIVMU, eroding the cross-subsidy to, and therefore participation by, side $\mathcal{A}$.

Thus the sign of the cross partial is determined by how the surplus created by marginal interaction benefits compares to their heterogeneity. Perhaps the sharpest way to express this is in terms of the relative importance of interaction benefits in profits compared to their relative importance in heterogeneity. Platform profits are just the sum of (twice) marginal interaction surplus $\tilde{b} N^{\mathcal{A}} N^{\mathcal{B}} \equiv (\widetilde{b^{\mathcal{A}}} + \widetilde{b^{\mathcal{B}}} - c) N^{\mathcal{A}} N^{\mathcal{B}}$ and marginal membership surplus, $\sum_{\mathcal{I}=A,B} (\widetilde{b^{\mathcal{I}}} - C^{\mathcal{I}}) N^{\mathcal{I}}$. It is therefore natural to consider the fraction of profits arising from marginal interaction surplus, *the interaction surplus ratio* $\alpha \equiv \tilde{b} N^{\mathcal{A}} N^{\mathcal{B}} / \pi$. Similarly the *interaction heterogeneity ratio* $\beta \equiv \mu_{\tilde{b}} / \mu_{\tilde{\mathcal{B}}}$, where $\mu_{\tilde{b}} = \sum_{\mathcal{I}=A,B} N^{\mathcal{I}} \mu^{\mathcal{I}}_{\tilde{b}}$ and $\mu_{\tilde{\mathcal{B}}} = \sum_{\mathcal{I}=A,B} N^{\mathcal{I}} \mu^{\mathcal{I}}_{\tilde{\mathcal{B}}}$, measures the relative *aggregate* importance of interaction heterogeneity.

THEOREM 5: *Participation on the two sides of the market are complements if $\mu_{\tilde{\mathcal{B}}} > 0$ and $\alpha > \beta$, substitutes if either $\mu_{\tilde{\mathcal{B}}} \leq 0$ or $\beta > \alpha$, and independent if $\mu_{\tilde{\mathcal{B}}} > 0$ and $\alpha = \beta$.*

PROOF:

See the Scale-Income Model portion of the Appendix.

Thus user heterogeneity ties the Spence distortion to the cross partial of participation rates. Because observing the cross partial requires only marginal shocks to market conditions, it may be easier to study empirically than the Spence distortion directly. Therefore one might measure basic features of user heterogeneity by the sign of the cross-participation effect, for example by observing the effect of a shock to one membership cost. Sadly, this is a coarse instrument, unable to distinguish which side of the market generates the interaction heterogeneity nor anything beyond its magnitude relative to the interaction surplus ratio. Measuring these finer properties requires richer data[33] or stronger assumptions.

The theorem makes clear the source of the seesaw effect in the RT (2003) model: there is no membership heterogeneity, so $\beta = \infty$, implying substitutes. By contrast, in Armstrong's model $\beta = 0$ as there is no interaction heterogeneity, and $\alpha > 0$ as otherwise the firm would end the

[33] A companion paper under preparation, Weyl (2009a), treats identification in multiproduct monopoly, with a focus on two-sided markets. I show that first-order instruments for participation rates reveal elasticities and the AIVMU, while quantitatively observable cost shocks reveal pass-through rates and the cross partial. Some tests of general multiproduct monopoly are also possible, but many of the finer normative features, and tests of the RT (2006) model specifically, require stronger assumptions or higher-order variation.

two-sidedness, separately providing services to the two sides. Thus the Armstrong model always has complements, showing that the seesaw principle is far from general.

B. *Welfare Effects*

In Section I, I argue that cross-group externalities *in the absence of transfers* are a defining feature of two-sided markets. However, others take the view (Hagiu 2007; Hagiu 2009; Rysman 2009) that two-sided markets are ones where, holding fixed some notion of price, each group's welfare depends on the other's participation and thereby *indirectly* (Jeffrey Church and Neil Gandal 1992; Michael Katz and Carl Shapiro 1994) on its own.

Such views are difficult to parse in multi-sided network models because the welfare-irrelevant details of pricing in these models still lead to very different indirect network effects holding fixed prices. For example, if insulating tariffs are charged to both sides then, by construction, such indirect network effects *never* exist. Thus, unless these authors think the canonical models miss the "essential nature" of two-sided markets, which I view as largely being defined by these models, it is difficult to see how such a test can be applied.[34]

Perhaps a more reasonable line of inquiry is therefore the nature of network effects in equilibrium. Suppose that participation on one side of a market rises for a reason, such as membership costs falling or membership values uniformly shifting up, that has no direct effect on the platform's incentives on the other side. I call the effect of such an exogenous increase in side $\mathcal{A}$ participation on side $\mathcal{B}$ welfare the equilibrium network effect.

THEOREM 6: *The equilibrium network effect from side $\mathcal{J}$ to side $\mathcal{I}$ has the same sign as*

$$\overline{b^{\mathcal{I}}} - \widetilde{b^{\mathcal{I}}} + \rho^{\mathcal{I}}\chi \, . \tag{13}$$

PROOF:

See the Scale-Income Model portion of the Appendix.

The first term of expression (13) is the direct effect of $\mathcal{J}$ participation on $\mathcal{I}$ welfare: the Spence distortion from $\mathcal{I}$ to $\mathcal{J}$. Only the distortion matters: $\widetilde{b^{\mathcal{I}}}$ is internalized by the platform as a higher price to side $\mathcal{I}$. One might view this direct effect as the equilibrium network *externality*. The second term is an indirect effect through the platform's optimization: the pass-through of the cross-side pricing effect. For example, if $\chi < 0$ (participation rates are substitutes as in RT 2003), side $\mathcal{I}$ will tend to benefit from an increase in prices on side $\mathcal{J}$ as this will incentivize the firm to obtain greater participation in side $\mathcal{I}$ by reducing prices.

Interaction heterogeneity both enlarges the Spence distortion and makes χ negative, while membership heterogeneity eliminates it or even reverses its sign but makes χ positive. Thus the source of heterogeneity has an ambiguous effect on expression (externalities). However, the first effect is fundamentally inframarginal, while the second depends only on local properties.

For example, in Armstrong's model, which has no Spence distortion, complementarity implies positive equilibrium network effects.[35] In the RT (2003) model, as discussed in Section II, the Spence distortion from side $\mathcal{I}$ is $\overline{s^{\mathcal{I}}}$, the user surplus on side $\mathcal{I}$, and the cross partial can be shown

[34] Liebowitz and Margolis (1994) discuss the dangers of abusing the concept of network externalities.

[35] On the other hand, if interaction surplus is negative and participation rates are substitutes (I do not know of any simple example of this), equilibrium network effects are negative. These conditions do not have any consistent relationship to the *primitive* externalities, the level of interaction values on the two sides.

to be the negative of per-interaction market power on either side of the market $m^{\mathcal{I}}$. Furthermore in Weyl and Fabinger (2009) we show that $\overline{s} = \overline{\rho} m$ where $\overline{\rho}$ is an average of pass-through rates over prices above the equilibrium level, as pass-through measures the log-curvature of demand. Expression (13) therefore becomes, in the RT (2003) case,

$$-m^{\mathcal{I}}\left(\overline{\rho^{\mathcal{I}}} - \rho^{\mathcal{I}}\right),$$

whose sign is determined by the slope of $\rho^{\mathcal{I}}$ with respect to cost/price: increasing pass-through implies average pass-through exceeds local pass-through, decreasing pass-through the reverse. Thus the *third* derivative of log-demand determines equilibrium network effects.

It may seem immediate that an increase in costs on side $\mathcal{I}$ harms side $\mathcal{I}$ users, but in Weyl (2009a) I showed that in the RT (2003) model the average user on one side of the market may actually want her prices increased to encourage a reduction in prices to users on the other side. The following corollary provides general conditions for this counterintuitive result.

COROLLARY 7: *$dV^{\mathcal{I}}/dC^{\mathcal{I}}$ has the sign of*

$$-\left(\frac{\mu^{\mathcal{A}}\mu^{\mathcal{B}}}{N^{\mathcal{A}}N^{\mathcal{B}}} + \chi\rho^{\mathcal{J}}\left[\overline{b^{\mathcal{I}}} - \widetilde{b^{\mathcal{I}}}\right]\right). \tag{14}$$

Thus the counterintuitive effect can occur at either extreme of heterogeneity. When interaction heterogeneity dominates, χ is negative while interaction surplus is strongly positive, so average users on side $\mathcal{I}$ may benefit from higher prices which encourage the platform to bring in more customers on side $\mathcal{J}$. For example in the RT (2003) model, expression (14) becomes $\overline{\rho^{\mathcal{I}}}\rho^{\mathcal{J}} - 1$; second-order conditions require $\rho^{\mathcal{I}}\rho^{\mathcal{J}} < 1$, as shown above, so beneficial own-cost increases require a $\rho^{\mathcal{I}}$ increasing rapidly in price, yielding a large Spence distortion.

On the other hand, when membership heterogeneity is strong enough to give negative interaction heterogeneity, interaction surplus is negative but $\chi > 0$ and average newspaper readers[36] may actually want higher prices to force firms to internalize their distaste for advertising and reduce its quantity. In intermediate cases, such as the Armstrong model, with small Spence distortions, own-cost effects are always negative.

C. *Arbitrary Comparative Statics*

Effects of local shocks to the market can always be expressed as a combination of direct externalities[37] and indirect effects through optimally chosen participation rates. The former can be analyzed through a partial derivative holding fixed firm actions; the second is equivalent to some combination of changes in the firm's (opportunity) cost on each side of the market. Therefore knowing $dN^{\mathcal{I}}/dC^{\mathcal{J}}, N_1^{\mathcal{I}}, N_2^{\mathcal{I}}, V_1^{\mathcal{I}}, V_2^{\mathcal{I}}$ for both $I, J = A, B$ is sufficient to compute arbitrary comparative statics; expressions for these are given in the text and Appendix. The same approach may be taken in the more general model proposed in Section III, though explicit expressions of the relevant derivatives do not appear in this paper.

[36] This makes clear that all the reasoning about surplus is about the *total* user surplus on each side of the market: it integrates over all users. Clearly marginal or near-marginal users are harmed by any increase in prices, even if these benefit loyal users. In some settings we may care about such distributional consequences (is ritzy readers' distaste for advertising reason enough to exclude poorer marginal readers?), but that analysis is beyond the scope of this paper and in fact most standard industrial organization.

[37] An earlier draft of this paper, available on request, provided a variety of such explicit comparative statics.

V. An Example: the Scale-Income Model

A primary contribution of this paper is to simplify the platform's problem to analyze, for the first time, the effect of multiple dimensions of user heterogeneity. The ability to analyze these more general models does not, however, eliminate all motivation for unidimensional models. As shown above, multidimensional heterogeneity leaves substantial ambiguities about the direction of various distortions and comparative statics. In cases when most heterogeneity plausibly lies along a single dimension, making this assumption explicit can help resolve these. Furthermore, from an empirical perspective it may be difficult to identify a two-dimensional model without parametric assumptions; restricting heterogeneity to a single dimension may be a simple and transparent way to impose the necessary additional structure.

Unfortunately, the source of heterogeneity in the most commonly applied model (Rysman 2004; Kaiser and Julian Wright 2006; Elena Argentesi and Lapo Filistrucchi 2007), Armstrong's, seems implausible in most settings where it is applied. A primary dimension of heterogeneity for at least one side of the market is almost certainly the value derived from the other side. The RT (2003) model focuses on this source of heterogeneity but has the unfortunate feature that it rules out any membership costs or benefits, making it implausible in many industries. However, my foregoing analysis emphasized that most results in the RT (2003) model are due to the source of heterogeneity rather than the absence of membership costs and benefits. Thus most of the results of the RT (2003) model extend to a Generalized RT (2003) (henceforth GRT 2003) model that allows for (homogeneous) membership costs and benefits.[38]

However, the GRT (2003) model still seems to fit many markets poorly. Newspaper readers and software producers, to name a few, clearly differ substantially in their membership benefits and costs, respectively, of participating in a platform. One reasonable model[39] of such settings (Anderson and Coate 2005) has GRT (2003)-like users on one side and Armstrong-like users on the other. In this section I propose an alternative that I think is likely to be most fruitful in applications: *Scale-Income* (SI) model. It offers a useful rule of thumb for thinking about sources of heterogeneity, making analysis a bit more concrete.

Users on each side agree on the relative size of membership and interaction values but differ in scale. All newspaper readers (side $\mathcal{A}$) lose a fraction $-\beta^{\mathcal{A}}N^{\mathcal{B}}$ of the value they take from reading if a fraction $N^{\mathcal{B}}$ of advertisers participate; however, they may differ in their total utility. Intuitively, higher income users have greater willingness to pay to gain the utility of reading the newspaper and avoid the disutility of advertising. Advertisers have the same value of circulation as a fraction of the fixed cost $-B_i^{\mathcal{B}}$ they expend to establish a relationship with the newspaper, but differ in the scale of both of these depending on their business size. Thus $b_i^{\mathcal{I}}/B_i^{\mathcal{I}} = \beta^{\mathcal{I}}$ for all $i, \mathcal{I}$, but users differ in the scale of their utility. They are heterogeneous vertically (Spence 1976; Eytan Sheshinski 1976) rather than horizontally (Harold Hotelling 1929).

I believe this model provides a better approximation to many two-sided markets than any of the other unidimensional models.[40] It seems to me a fairly good fit to software platforms[41] (operating

[38] This model was analyzed extensively in a previous draft of this paper and, while omitted here for brevity, is available on request.

[39] This *Hybrid* model was extensively analyzed in a previous draft of this paper, available on request.

[40] Note that the RT (2003) model is the special case of the SI model where $\beta^{\mathcal{I}} = \infty$. An interesting potential extension of the SI model is to extend this in the way the GRT (2003) model extends the RT (2003) model: allow users to lie along any line in R^2.

[41] As an example, I will go into a bit more detail on this case. Users typically derive some value from the platform itself and some proportional to the media (games or programs) on the platform. It seems reasonable to assume that the ratio between these is probably quite homogeneous in the population. Similarly software producers have development costs and average per user (unit profits multiplied by the probability of a purchase). At least in expected terms, this ratio is likely quite homogeneous, as software producers that expend larger fixed costs for the same variable benefit as another

systems, video games etc.), dating clubs, commercial intermediation (supermarkets, stock markets, eBay, etc.) and Internet service provision.

For concreteness, I focus here on a version of the model adapted to newspapers or other advertising platforms. $\beta^{\mathcal{I}} < 0$ on both sides of the market: readers on side $\mathcal{A}$ have positive membership values from reading the paper but negative interaction values from viewing advertising, while side $\mathcal{B}$ advertisers have positive interaction value of circulation but membership costs associated with establishing relationships with the paper. Because

$$P^{\mathcal{I}} = \widetilde{B^{\mathcal{I}}} + \widetilde{b^{\mathcal{I}}}N^{\mathcal{J}} = \widetilde{b^{\mathcal{I}}}\left(N^{\mathcal{J}} + \frac{1}{\beta^{\mathcal{I}}}\right)$$

we have that $\widetilde{b^{\mathcal{I}}} = P^{\mathcal{I}}/(N^{\mathcal{J}} + (1/\beta^{\mathcal{I}})) = P^{\mathcal{I}}\nu^{\mathcal{I}}$ where $\nu^{\mathcal{I}} \equiv 1/(N^{\mathcal{J}} + (1/\beta^{\mathcal{I}}))$. The Spence distortion from side $\mathcal{I}$ is now $\mu^{\mathcal{I}}\rho^{\mathcal{I}}N^{\mathcal{I}}$ as interaction surplus is just interaction market power multiplied by the average pass-through of its distribution (see Subsection IIIB above). Rather than the sign of the Spence distortion's being dictated directly by the model or left entirely ambiguous, it is given in an intuitive way by market conditions that can be reflected upon or econometrically measured. If prices on side $\mathcal{B}$ (advertisers) have the same sign as interaction benefits on that side, then loyal users tend to have higher (or less negative, in the case of negative prices) interaction benefits than marginal users and therefore prices on side $\mathcal{A}$ (readers) are distorted upwards. On the other hand if interaction benefits on side $\mathcal{A}$ have the opposite sign of price, as with a high-quality newspaper whose readers dislike advertising, then loyal users tend to have more negative (or less positive when prices are negative) interaction values than marginal readers and therefore prices on side $\mathcal{B}$ are distorted *downward* (advertisers).

Note that the crucial difference here is *not just the sign of interaction values, but how these compare to the sign of price*. Free tabloids essentially have a negative price, given their aggressive marketing in public transport hubs, and therefore have low scale-income, advertising-insensitive loyal readers, implying an upward Spence distortion *despite negative interaction benefits*.[42] Thus the SI model would have very different predictions about the behavior of tabloids versus high-quality papers as the marginal readers of both desert for the Internet: tabloids will become further laden with advertising and market more aggressively, while quality papers will pare back advertising and raise subscription fees.

Comparative statics are similarly dictated by the market conditions. The Scale-Income Model portion of the Appendix shows participation on the two sides are complements (substitutes) if and only if

$$\sum_{\mathcal{I}=A,B}\left(\nu^{\mathcal{I}}\left[\widetilde{b^{\mathcal{J}}}N^{\mathcal{J}} - C^{\mathcal{I}} - cN^{\mathcal{J}}\right]\right) - c > (<)0\,.$$

For quality newspapers $\nu^{\mathcal{A}} < 0 < \nu^{\mathcal{B}}$. Assuming subscribers are net profitable even in the absence of subscription fees (advertisers are obviously unprofitable in this sense), $\{\widetilde{b^{\mathcal{B}}}N^{\mathcal{B}} - C^{\mathcal{A}} - cN^{\mathcal{B}} > 0$. So long as these effects are large enough to outweigh interaction costs, participation rates are complements. Also intuitively the equilibrium network effect from readers to advertisers is positive by complementarity, but the sign of equilibrium network effect of advertiser participation to

producer will be driven out of the market. However, some games and software are clearly much more prominent and higher impact than others, having larger fixed costs and variable benefits. Thus the SI model seems a sensible fit.

[42] Similarly if programs for an operating system are subsidized, as with Macintosh in the 1990s, low scale programs will tend to be served and thus there will be a downward Spence distortion (potentially underpriced operating systems) despite positive interaction benefits.

readers is ambiguous (the harm to loyal readers may outweigh the benefits to marginal readers or not).

Empirical data become particularly useful in the SI model as it has substantial identifying power even when little can be observed,[43] especially when explicit links can be made to observable income or size distributions,[44] as is common in structural empirical work (Berry, James Levinsohn, and Ariel Pakes 1995). Thus in cases where the source of user heterogeneity is unknown, relevant policy implications are theoretically ambiguous, and empirical data for full identification are missing, the SI model provides a reasonable way to structure policy intuitions. Furthermore, it can easily be extended to the general model of Section III: $u^{\mathcal{I}}(\mathbf{N};\theta^{\mathcal{I}}) \equiv \theta^{\mathcal{I}} u^{\mathcal{I}}(\mathbf{N})$ where $u^{\mathcal{I}}(\cdot)$ is an arbitrary smooth function of the allocation.

VI. Applications

This section briefly discusses three policy-oriented applications of my results, designed to demonstrate how the tools developed help address longstanding applied questions.

A. *Measuring Market Power and Predation*

In applied antitrust analysis, price-cost margins are used to measure market power or as a screen for predatory pricing. It has long been argued (David S. Evans 2003; Wright 2004) that pricing below cost is not indicative of predatory behavior[45] in a two-sided market as users may be subsidized on one side to reflect the benefits of users on the other side. Similarly pricing significantly above cost need not indicate large market power, as users on one side may be taxed if users on the other side have interaction costs. Measuring market power and predation in two-sided markets is therefore an old open question. My framework provides a simple answer: a general Lerner index for two-sided markets, which encompasses and unifies previous Lerner indices proposed for special models, such as Armstrong's and RT (2003).

One approach to such a Lerner index is to construct them for each side of the market individually, in which case they are given by equation (6) and require a measurement of the AIVMU, as well as costs.[46] Measuring the AIVMU may be difficult, but it's not much harder than observing costs. These measures can then be used, as any Lerner index, as a test for market power[47] and predation. Because prices are often near or below zero in two-sided markets, absolute market power $\mu^{\mathcal{I}}$, perhaps normalized by something other than price, may be a more attractive metric as it is guaranteed to be positive and finite for a statically optimizing firm. This may be calculated

[43] A decomposition of price into interaction and membership benefits and identification of market power, which is feasible simply based on first-order instruments for participation or price on both sides of the market, suffice to identify interaction market power.

[44] These predict higher-order properties of demand, allowing pass-through rates and cross partials to be predicted, and the size of interaction surplus, and therefore normative comparative statics, to be estimated.

[45] While there is much dislike about requiring below true cost pricing as a necessary condition for predation (Aaron S. Edlin 2002), most legal (Frank H. Easterbrook 1981; Brooke Group Ltd. 1993) and economic (Patrick Bolton, Joseph F. Brodley, and Michael H. Riordan 2000) doctrine holds that allegations of predation must establish in Easterbrook's words "a sacrifice of today's profits for tomorrow's." This means that, in practice, to the extent predation cases arise at all in two-sided markets, the argument that below-cost pricing does not establish that prices are below "true" costs is likely to be a potent one.

[46] Market power may also be estimated structurally (Argentesi and Filistrucchi 2007). I discuss this approach, which also allows costs to be estimated rather than observed, extensively in Weyl (2009a).

[47] It is not immediately clear why market power, and not market power combined with the Spence distortion, is the right thing to measure. For the purposes of my discussion here, I just take as given the policymakers' interest in measuring market power.

just as easily: $P^{\mathcal{I}} - C^{\mathcal{I}} - cN^{\mathcal{J}} + \widetilde{b}^{\mathcal{J}} N^{\mathcal{J}}$. Then a natural test for predation on one side individually is that this be negative.

If, instead, an aggregate measure of market power is desired, weighting by participation on the two sides is a natural way to aggregate. The aggregate Lerner index is then

$$N^{\mathcal{A}}\mu^{\mathcal{A}} + \frac{N^{\mathcal{B}}\mu^{\mathcal{B}}}{N^{\mathcal{A}}P^{\mathcal{A}}} + N^{\mathcal{B}}P^{\mathcal{B}} = (1 + \alpha)\frac{\pi}{R}$$

where R is revenue and α is the interaction surplus ratio of Subsection IVA. Intuitively if two-sidedness makes up a large part of profits, one should expect relatively low prices for a given amount of market power, as the platform will tend to subsidize users for participation. Therefore even a small profit to revenue ratio indicates significant market power if two-sidedness is a main source of profits. The test for predation is the natural extension of the standard test: profits are negative if and only if the aggregate Lerner index is. My formulae, side-specific or aggregated, extend intuitively to the general multi-sided model of Section III.

B. *Regulation*

Regulation of two-sided markets has been a topic of substantial recent interest. Two prominent examples are the policy debates over interchange fee caps on prices to card-accepting merchants, and net neutrality regulations, interpreted variously as price caps on fees Internet service providers (ISPs) can charge Web sites or a limit on their price discrimination. As with merger analysis, designing regulation in two-sided markets is beyond the scope of this paper. However, I believe the paper does provide three issues for future research to consider.

First it emphasizes that, to the extent that regulation aims to emulate the optimal benchmarks of Pigouvian or Ramsey pricing, it should solve distortions on both sides. In fact Pigouvian and Ramsey pricing require solving a constant fraction of distortions on each side, rather than only one side, as with net neutrality or interchange fee regulations. In considering the size of these distortions it suggests two factors are crucial: the size of classical market power and the Spence distortion *on the other side of the market*.

Thus the novel element in two-sided markets is that regulators should focus most on reducing price opposite a side with a large Spence distortion. Thus regulators of ISPs should focus on limiting prices to Web sites (net neutrality) if there is more (interaction) surplus among loyal users than among highly profitable Web sites. But if the situation is reversed, forcing ISPs to reduce prices and build more line to consumer homes may be a higher priority.

Second, implementing Ramsey-type regulation requires a detailed knowledge of demand[48] that may not be available to a regulator. If so it may be more attractive to regulate only one side of the market, especially if market power is thought to particularly distort that side's prices. However a price cap on side $\mathcal{I}$ can create further distortions, especially with positive interaction benefits, as the platform can lower side $\mathcal{I}$'s price either by increasing participation on side $\mathcal{I}$ (which the regulator wants) or by decreasing participation on side $\mathcal{J}$ (which she likely does not want). Thus Sheshinski's (1976) argument that price regulation tends to reduce quality provision is even stronger. In two-sided markets "quality reduction" comes from further distorting prices charged to users on the other side of the market. Of course, when interaction benefits are negative, especially if the Spence distortion is upward, this may be desirable: price caps on newspaper readers

[48] When information is more limited, the appropriate response is to explicitly incorporate these informational constraints into a model of policy design (David P. Baron and Myerson 1982; Jean-Jacques Laffont and Jean Tirole 1993). This is an important open problem in two-sided markets and is certainly beyond the scope of this paper.

may lead to more ads,[49] but this could well be an efficient counterbalance to their market power over advertisers especially if, as with tabloids, loyal readers dislike advertising less than marginal readers.

In the positive interaction benefit and Spence distortion case, when price regulation is particularly unattractive, Sheshinski's suggestion of quantity regulation may be more attractive as it does not change pricing incentives on the other side of the market. The simplest way to see this[50] is to note that the privately optimal pricing condition on side $\mathcal{I}$ takes as given participation on side $\mathcal{J}$, and thus the first-order condition on side $\mathcal{I}$ is not (directly) affected by a constraint on participation on side $\mathcal{J}$. A regulator might require an ISP to have a certain fraction of Web sites available on its service, rather than prohibiting the charging of Web sites. This might well encourage the recruitment of more Internet users, as a natural way to increase Web site participation without lowering price is to increase the number of subscribers. Of course, as in any market where quantity regulation is proposed, implementation would require care, to ensure that the ISP does not cheat by signing up the smallest Web sites. Given the diversity of Web sites, the practical complexity of implementing such a policy may outweigh its theoretical benefits. Furthermore, even at a theoretical level, more detailed analysis would be needed to establish the cases in which, once all indirect effects are taken into account, participation regulations are truly preferable to price regulations, and for whom. Nonetheless, such *allocation regulation* at least merits further investigation in multi-sided networks.

Finally, the analysis above seems to provide further rationale for allowing price discrimination in two-sided markets, at least when Spence distortions are positive. In this case discrimination has the additional benefit (Weyl 2009c; Rysman 2009) of increasing the subsidy to users on the other side of the market, helping ameliorate both the market power (on the other side) and Spence (on the discriminated side) distortions. Because Spence distortions are likely upward among Web sites (incumbents like Google make greater profits from a marginal surfer than entrants), this seems to lean in favor of allowing price discrimination,[51] that is, repealing net neutrality. However, if the Spence distortion is negative, as among *Times* readers, price discrimination may be more harmful than usual as it may lead to higher advertiser prices exacerbating market power.[52] Again, more detailed analysis of price discrimination would be required to formalize such arguments.

C. *Mergers*

Merger analysis requires a general model of competition, which is beyond the scope of this paper. Nonetheless my results make three small contributions towards this goal.

First, the approach taken here is likely to be useful in analyzing such merger models. To illustrate this, I show in the Applications portion of the Appendix how it can be used to analyze mergers in a nonparametric, market-expanding version of Armstrong's model of symmetrically differentiated single-homing duopoly, the *generalized Armstrong single-homing model* (GASH). A companion paper (Weyl 2008) uses the same techniques to analyze mergers between

[49] Some of these issues are analyzed, under particular assumptions about user heterogeneity (see footnote 12) by Anderson and Coate (2005).

[50] A formalization is available on request.

[51] Of course, as in any vertical moral hazard/double marginalization problem, transferring incentives to the platform is not all good; this may hold up Web sites, extracting surplus from their investment in producing quality content if contracts are not sufficiently rich (Bengt Holmström 1982). For an analysis that emphasizes the effects on Web site investment see Jay Pil Choi and Byung-Cheol Kim (forthcoming).

[52] Of course this depends on whether prices are initially too high or too low to advertisers; in the latter case, the effect is ambiguous.

a platform and a non–two-sided firm producing a good that is a substitute (broadcast TV merger with advertising-free cable) or a complement (operating systems and hardware manufacturers) for users on one side of the market, considering the second case in substantial detail.

Second, the insulating tariff offers an approach to overcoming a problem[53] plaguing the analysis of competition in multi-sided networks. As Armstrong points out, a tremendous multiplicity of equilibria are possible in competition between platforms depending on the tariffs $P^{\mathcal{A}}(\cdot)$ and $P^{\mathcal{B}}(\cdot)$ the other firm uses *at participation levels other than the equilibrium*. For example, if one payment card firm uses a fixed fee, this will encourage the other to steal its cardholders as a means of attracting merchants who now have fewer partners, while if it uses a negative fixed fee and a large per-interaction fee this softens competition as stealing cardholders actually discourages merchant participation.

However, if one assumes firms choose insulating tariffs, these cross-side participation stealing effects are reduced and, in the GASH case, entirely eliminated. This *insulated equilibrium* greatly simplifies the analysis of competition. It also seems at least as reasonable an assumption as the more basic Nash-in-prices (Bertrand) assumption universal in the multi-sided networks literature,[54] given that this tariff is both intuitive and plausible, as well as robustly ensuring good equilibria are uniquely selected. It is a simple extension of the common assumption in price-quality competition that firms take as given other firms' choice of quality (Avner Shaked and John Sutton 1982) when choosing price, as the number of users participating on side $\mathcal{J}$ is effectively the quality of the platform's "product" on side $\mathcal{I}$.

Third, my results suggest that in any model of competition, the source of user heterogeneity will be central to determining the positive and normative effects of mergers. Mergers largely affect firm market power, and potentially the size of network effects, both of which act to shift platform (opportunity) costs (Farrell and Carl Shapiro 2008). Because the positive and normative effects of changes in costs and network effects are determined by the sources of user heterogeneity, so too will be the effects of mergers. Furthermore, whether market power is more or less harmful in a two-sided market depends on the source of heterogeneity.

This is confirmed by the two most prominent existing models of competition in two-sided markets. In Weyl (2009c) I show that a merger (with no efficiencies) in the RT (2003) model of competition is tantamount to an increase in market power on one or more sides of the market. It will therefore *increase* participation, and potentially benefit users, on one side if competition is much more intense on one side (participation on the two sides are substitutes). On the other hand, in the online Appendix I show that, at least when competitors use insulating tariffs and regardless of the relative intensity of competition, a merger (without efficiencies) in the GASH model increases market power and therefore reduces participation and welfare on both sides, as

[53] An alternative approach to making a specific assumption about conduct, as I suggest here, is to search for results that are robust across various solution concepts or to attempt to explicitly identify the solution concept. The first approach seems reasonable, if challenging, and is an interesting direction for future research. A simple example of this strategy was a result, included in a previous version of this paper and available on request, that in many reasonable cases, even without an insulating tariff, mergers from GASH lead to lower participation on both sides. The second approach is in the spirit of the classic contributions of Timothy F. Bresnahan (1982) but has proven difficult to implement empirically given its data demands (Aviv Nevo 1998). Nonetheless there has been some recent interest in identifying solution concepts in other contexts, such as vertical relations (Sofia Villas-Boas and Rebecca Hellerstein 2006), so asking how one would go about identifying the two-sided markets solution concept (what sort of price schedules do firms take as given) would be an interesting topic for future research. Finally, one might use demand uncertainty to tie down a unique optimal tariff (Klemperer and Margaret A. Meyer 1989), though this approach has proved challenging to implement in applications in the simpler context of one-sided supply function equilibrium. Nonetheless I think operationalizing uncertainty-based refinements of oligopoly equilibria is an exciting direction for future research.

[54] Bruno Jullien proposed to me, in a private conversation, a model of undifferentiated Cournot-style competition. However, this model has symmetric equilibria only when there is a single dimension of user heterogeneity, making it difficult to analyze more generally. A proof is available on request.

participation rates are complements. Thus merger models in two-sided markets must show care in their assumptions about the sources of user heterogeneity.

VII. Conclusion

This paper makes two contributions. First, by formulating the platform's problem in terms of its choice of allocation, rather than prices, I simplify and generalize the analysis of network industries. Second, I show that the key normative properties and comparative statics of two-sided markets depend on the source of user heterogeneity, which previous analysis has restricted. The modesty of these contributions makes clear the early stage of the literature. I therefore conclude by discussing directions for future research.

On the empirical side a number of questions are suggested quite directly by my arguments above. Does the SI model fit well in some market where ex ante the sources of user heterogeneity seem unclear? How well do the predictions of the RT (2003) model fit actual payment card data? Do newspapers actually exhibit complements? Comparing market power to the Spence distortions, are there overall too many or too few ads in most papers? Applications will largely be driven by the data available, so I will not dwell on them excessively here.

On the theoretical side, much remains to be done to understand pricing in networks more generally. For example, my approach so far allows only extremely stylized models of competition of limited direct empirical relevance. I consider formulating a workhorse, general empirical model[55] of two-sided markets, and practical means for identifying it, to be the most important open question in this area. For regulatory policy the monopoly model is likely to be of greater use, but a more careful analysis of price discrimination and regulatory design (Baron and Myerson 1982; Laffont and Tirole 1993) are needed.

A number of fundamental theoretical problems remain open, three of which I will mention.

First, the exploding literature on matching design, surveyed by Alvin E. Roth (2002), has thus far had limited interaction with the literature on pricing in two-sided markets; see Glenn Ellison, Drew Fudenberg and Markus Möbius (2004); Susan Athey and Glenn Ellison (2008); Ettore Damiano and Hao Li (2008); Andrei Hagiu and Bruno Jullien (2008); Weyl and Tirole (2010) for notable, if early, exceptions. These literatures have much in common, though market design has largely focused on efficiency and paid little attention to prices, while the two-sided markets literature largely ignores, as this paper does, the possibility of designing platforms to increase surplus. I suspect optimal pricing interacts importantly with platform design and therefore that such "revenue maximizing matching" is a fruitful direction for future research.

Finally, as my discussions in Subsection IC and IIIA emphasize, the coordination problems that have long been thought central to networks can generally be overcome by appropriate tariffs. This does not seem to always occur in practice, however. Insulating tariffs might be difficult to implement if demand is not known exactly to the platform; they might, in fact, be unprofitable if demand is uncertain as the critical mass problem might be an effective screen for high demand states. Standard capital market imperfections could also play a role in limiting the platform's ability to borrow, which might be necessary in the true dynamic process of network formation swept under the static model here. A platform might signal to its financiers that it knows it will succeed by overcoming the critical mass problem without subsidies. These are all interesting topics for future theoretical research.

Regardless of the precise explanation for imperfect insulation, my discussion suggests that coordination problems may be a choice, rather than a constraint. If correct, this would imply, for

[55] Alex White and Weyl (2010) make a first attempt at this, extending the insulating tariff to oligopoly.

example, that coordination is, on its own, an important basic source of market power and possible coordination failures are not a reasonable rationale for a merger or collusion. More careful evaluation of this controversial claim is an important theoretical challenge.

REFERENCES

Ambrus, Attila, and Rossella Argenziano. 2009. "Asymmetric Networks in Two-Sided Markets." *American Economic Journal: Microeconomics*, 1(1): 17–52.

Anderson, Simon P., and Stephen Coate. 2005. "Market Provision of Broadcasting: A Welfare Analysis." *Review of Economic Studies*, 72(4): 947–72.

Argentesi, Elena, and Lapo Filistrucchi. 2007. "Estimating Market Power in a Two-Sided Market: The Case of Newspapers." *Journal of Applied Econometrics*, 22(7): 1247–66.

Armstrong, Mark. 2006. "Competition in Two-Sided Markets." *RAND Journal of Economics*, 37(3): 668–91.

Athey, Susan, and Glenn Ellison. 2008. "Position Auctions and Consumer Search." http://kuznets.harvard.edu/~athey/position.pdf.

Baron, David P., and Roger B. Myerson. 1982. "Regulating a Monopolist with Unknown Costs." *Econometrica*, 50(4): 911–30.

Becker, Gary S., and Kevin M. Murphy. 1993. "A Simple Theory of Advertising as a Good or Bad." *Quarterly Journal of Economics*, 108(4): 941–64.

Bedre-Defolie, Özlem, and Emilio Calvano. 2010. "Pricing Payment Cards." Unpublished.

Berry, Steven T., and Joel Waldfogel. 1999. "Free Entry and Social Inefficiency in Radio Broadcasting." *RAND Journal of Economics*, 30(3): 397–420.

Berry, Steven T., James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–90.

Besanko, David, Shabtai Donnenfeld, and Lawrence J. White. 1987. "Monopoly and Quality Distortion: Effects and Remedies." *Quarterly Journal of Economics*, 102(4): 743–67.

Bolton, Patrick, Joseph F. Brodley, and Michael H. Riordan. 2000. "Predatory Pricing: Strategic Theory and Legal Policy." *Georgetown Law Journal*, 88(8): 2239–330.

Bresnahan, Timothy F. 1982. "The Oligopoly Solution Concept Is Identified." *Economics Letters*, 10(1–2): 87–92.

Caillaud, Bernard, and Bruno Jullien. 2001. "Competing Cybermediaries." *European Economic Review*, 45(4–6): 797–808.

Caillaud, Bernard, and Bruno Jullien. 2003. "Chicken & Egg: Competition among Intermediation Service Providers." *RAND Journal of Economics*, 34(2): 309–28.

Choi, Jay Pil, and Byung-Cheol Kim. Forthcoming. "Net Neutrality and Investment Incentives." *RAND Journal of Economics*.

Church, Jeffrey, and Neil Gandal. 1992. "Network Effects, Software Provision, and Standardization." *Journal of Industrial Economics*, 40(1): 85–103.

Damiano, Ettore, and Hao Li. 2008. "Competing Matchmaking." *Journal of the European Economic Association*, 6(4): 789–818.

Dybvig, Philip H., and Chester S. Spatt. 1983. "Adoption Externalities as Public Goods." *Journal of Public Economics*, 20(2): 231–47.

Easterbrook, Frank H. 1981. "Predatory Strategies and Counterstrategies." *University of Chicago Law Review*, 48(2): 263–337.

Economides, Nicholas. 1996. "The Economics of Networks." *International Journal of Industrial Organization*, 14(6): 673–99.

Edlin, Aaron S. 2002. "Stopping above-Cost Predatory Pricing." *Yale Law Journal*, 111(4): 941–91.

Ellison, Glenn, and Drew Fudenberg. 2003. "Knife-Edge or Plateau: When Do Market Models Tip?" *Quarterly Journal of Economics*, 118(4): 1249–78.

Ellison, Glenn, Drew Fudenberg, and Markus Mobius. 2004. "Competing Auctions." *Journal of the European Economic Association*, 2(1): 30–66.

Evans, David S. 2003. "The Antitrust Economics of Multi-sided Platform Markets." *Yale Journal on Regulation*, 20(2): 325–81.

Evans, David S., and Richard Schmalensee. 2009. "Failure to Launch: Critical Mass in Platform Businesses." http://www.interic.org/Conference/Schmalensee.pdf.

Fan, Ying. 2010. "Ownership Consolidation and Product Quality: A Study of the U.S. Daily Newspaper Market." http://www.personal.umich.edu/~yingfan/DailyNewspaper_Fan.pdf.

Farrell, Joseph, and Paul Klemperer. 2007. “Coordination and Lock-in: Competition with Switching Costs and Network Effects.” In *Handbook of Industrial Organization*, Vol. 3, ed. Mark Armstrong and Robert H. Porter, 1967–2072. Amsterdam: North-Holland.
Farrell, Joseph, and Carl Shapiro. 2008. “Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition.” http://faculty.haas.berkeley.edu/shapiro/alternative.pdf.
Gaudeul, Alex, and Bruno Jullien. 2008. “E-Commerce, Two-Sided Markets and Info-Mediation.” In *Internet and Digital Economics: Principles, Methods and Applications*, ed. Eric Brousseau and Nicolas Curien, 268–90. Cambridge, UK: Cambridge University Press.
Gentzkow, Matthew, and Jesse Shapiro. 2010. “What Drives Media Slant? Evidence from U.S. Newspapers.” *Econometrica*, 78(1): 35–71.
Guthrie, Graeme, and Julian Wright. 2007. “Competing Payment Schemes.” *Journal of Industrial Economics*, 55(1): 37–67.
Hagiu, Andrei. 2006. “Pricing and Commitment by Two-Sided Platforms.” *RAND Journal of Economics*, 37(3): 720–37.
Hagiu, Andrei. 2007. “Merchant or Two-Sided Platform?” *Review of Network Economics*, 6(2): 115–33.
Hagiu, Andrei. 2009. “Two-Sided Platforms: Product Variety and Pricing Structures.” *Journal of Economics and Management Strategy*, 18(4): 1011–43.
Hagiu, Andrei, and Bruno Jullien. 2008. “Why Do Intermediaries Divert Search?” http://www.people.hbs.edu/ahagiu/research.html.
Holmstrom, Bengt. 1982. “Moral Hazard in Teams.” *Bell Journal of Economics*, 13(2): 324–40.
Hotelling, Harold. 1929. “Stability in Competition.” *Economic Journal*, 39(153): 41–57.
Kaiser, Ulrich, and Minjae Song. 2009. “Do Media Consumers Really Dislike Advertising?: An Empirical Assessment of the Role of Advertising in Print Media Markets.” *International Journal of Industrial Organization*, 27(2): 292–301.
Kaiser, Ulrich, and Julian Wright. 2006. “Price Structure in Two-Sided Markets: Evidence from the Magazine Industry.” *International Journal of Industrial Organization*, 24(1): 1–28.
Katz, Michael L., and Carl Shapiro. 1985. “Network Externalities, Competition, and Compatibility.” *American Economic Review*, 75(3): 424–40.
Katz, Michael L., and Carl Shapiro. 1994. “Systems Competition and Network Effects.” *Journal of Economic Perspectives*, 8(2): 93–115.
Klemperer, Paul D., and Margaret A. Meyer. 1989. “Supply Function Equilibria in Oligopoly under Uncertainty.” *Econometrica*, 57(6): 1243–77.
Laffont, Jean-Jacques, and Jean Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.
Lee, Robin S. 2009. “Vertical Integration and Exclusivity in Platform and Two-Sided Markets.” http://pages.stern.nyu.edu/~rslee/.
Liebowitz, S. J., and Stephen E. Margolis. 1994. “Network Externality: An Uncommon Tragedy.” *Journal of Economic Perspectives*, 8(2): 133–50.
Lindhal, Erik. 1919. “Positive Lösung.” *Die Gerechtigkeit der Besteuerung.*
Ltd., Brooke Group. 1993. “Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.” (92–466), 509 U.S.(209).
Mussa, Michael, and Sherwin Rosen. 1978. “Monopoly and Product Quality.” *Journal of Economic Theory*, 18(2): 301–17.
Myerson, Roger B. 1981. “Optimal Auction Design.” *Mathematics of Operations Research*, 6(1): 58–73.
Nevo, Aviv. 1998. “Identification of the Oligopoly Solution Concept in a Differentiated-Products Industry.” *Economics Letters*, 59(3): 391–95.
O’Keeffe, Mary, W. Kip Viscusi, and Richard J. Zeckhauser. 1984. “Economic Contests: Comparative Reward Schemes.” *Journal of Labor Economics*, 2(1): 27–56.
Oum, Tae Hoon, and Michael W. Tretheway. 1988. “Ramsey Pricing in the Presence of Externality Costs.” *Journal of Transport Economics and Policy*, 22(3): 307–17.
Pigou, Arthur. 1920. *The Economics of Welfare.* London: MacMillan.
Rochet, Jean-Charles, and Jean Tirole. 2003. “Platform Competition in Two-Sided Markets.” *Journal of the European Economic Association*, 1(4): 990–1029.
Rochet, Jean-Charles, and Jean Tirole. 2006. “Two-Sided Markets: A Progress Report.” *RAND Journal of Economics*, 37(3): 645–67.
Roth, Alvin E. 2002. “The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics.” *Econometrica*, 70(4): 1341–78.
Rysman, Marc. 2004. “Competition between Networks: A Study of the Market for Yellow Pages.” *Review of Economic Studies*, 71(2): 483–512.

Rysman, Marc. 2009. "The Economics of Two-Sided Markets." *Journal of Economic Perspectives*, 23(3): 125–43.
Samuelson, Paul A. 1941. "The Stability of Equilibrium: Comparative Statics and Dynamics." *Econometrica*, 9(2): 97–120.
Segal, Ilya. 1999. "Contracting with Externalities." *Quarterly Journal of Economics*, 114(2): 337–88.
Shaked, Avner, and John Sutton. 1982. "Relaxing Price Competition through Product Differentiation." *Review of Economic Studies*, 49(1): 3–13.
Sheshinski, Eytan. 1976. "Price, Quality and Quantity Regulation in Monopoly Situations." *Economica*, 43(17): 127–37.
Spence, A. Michael. 1975. "Monopoly, Quality, and Regulation." *Bell Journal of Economics*, 6(2): 417–29.
Spence, A. Michael. 1976. "Product Differentiation and Welfare." *American Economic Review*, 66(2): 407–14.
Veiga, Andre, and E. Glen Weyl. 2010. "Platforms with Heterogeneous Externalities." Unpublished.
Villas-Boas, Sofia, and Rebecca Hellerstein. 2006. "Identification of Supply Models of Retailer and Manufacturer Oligopoly Pricing." *Economics Letters*, 90(1): 132–40.
Weyl, E. Glen. 2008. "Double Marginalization in Two-Sided Markets." http://www.people.fas.harvard.edu/~Eweyl/Genius_JMP.pdf.
Weyl, E. Glen. 2009a. "Slutsky Meets Marschak: First-Order Identification of Multi-Product Production." http://www.people.fas.harvard.edu/~weyl/SlutMars_12_09_pdf.
Weyl, E. Glen. 2009b. "Monopoly, Ramsey and Lindahl in Rochet and Tirole (2003)." *Economics Letters*, 103(2): 99–100.
Weyl, E. Glen. 2009c. "The Price Theory of Two-Sided Markets." http://www.people.fas.harvard.edu/~weyl/pt2sms_3_09.pdf.
Weyl, E. Glen, and Michal Fabinger. 2009. "Pass-Through as an Economic Tool." http://www.people.fas.harvard.edu/~weyl/Pass-through_10_09.pdf.
White, Alexander, and E. Glen Weyl. 2010. "Imperfect Platform Competition: A General Framework." Unpublished.
Wright, Julian. 2004. "The Determinants of Optimal Interchange Fees in Payment Systems." *Journal of Industrial Economics*, 52(1): 1–26.