Composition Effects in Platforms with Population Heterogeneity*

Michael Sacks

University of Louisville

Abstract

I develop a duopoly model of competition between platforms, incorporating users with heterogeneous preferences over both the platforms' characteristics and the presence of other users. Hence platforms are concerned with both the number and composition of users. The model yields novel representations of heterogeneity, size effects, and composition effects. I use these representations to decompose the relationship between the price and the size and composition of a platform. Prices need not be monotonic in the size of the installed base and profitability can similarly vary inversely. I identify conditions under which prices are increasing, decreasing, or unchanging in platform size. Given that users care about composition, nonpricing strategies to cultivate platform composition emerge when platforms cannot sufficiently leverage price discrimination. Composition effects and cultivation reframe the dominant-firm fringe-firm paradigm and explain the presence of multiproduct firms in platform markets, such as online dating.

January 11, 2020

JEL Classifications: D2, L1, L2, M3, Z1.

Keywords: equilibrium selection, network effects, networks, platforms, price competition, price dispersion, product differentiation.

*Department of Economics, University of Louisville, College of Business, Harry Frazier Hall, Louisville, KY 40292. michael.sacks@louisville.edu.
I am grateful to Jean-Paul Carvalho for invaluable comments and guidance. This paper has benefited greatly from comments by Jiawei Chen, Yong Chao, Linda Cohen, Daniel Grossman, Joshua Hall, Justin Johnson, Christoffer Koch, Paul Lombardi, Michael McBride, Jennifer Muz, Marius Schwartz, Stergios Skaperdas, Angela Vossmeyer, and Timothy Wong. I also thank participants at the 16^{th} annual International Industrial Organization Conference and seminar participants at West Virginia University. All errors are my own.

1 Introduction

As technology evolves and the world becomes increasingly interconnected, a better understanding of network externalities becomes increasingly important. Facebook's market cap is over $500 billion, YouTube's is over $80 billion (*est.*), Match Group Inc.'s is over $20 billion, and the various online videogaming platforms total over $30 billion.[1] A user's willingness to pay for access to a platform depends on the number of other users. While true, this statement is also lacking. A user's willingness to pay for access to a platform also depends on the characteristics of other users. Similarly, willingness to participate in networks such as open source software development communities depends on contributor composition.

Network size effects are well studied.[2] Significantly less attention has been devoted to understanding how the composition of the installed base influences the users' valuations of platforms and thus platforms' responses to user composition effects.[3] The effects of ignoring composition effects are consequential. For one, the pricing strategies of platforms are distorted, creating a wedge between expected and observed prices. Secondly, a second margin of competition between platforms is masked: product differentiation over the network externality. If composition is valuable, then platforms will actively engage in efforts to both influence their composition and shape user preferences. When unable to engage in price discrimination to influence user choice, the platforms utilize in nonpricing strategies to influence their composition. These nonpricing strategies interact with pricing strategies, altering our understanding of platform behavior and competition. This paper develops a tractable representation of heterogeneity, which precisely characterises the changes in predicted behavior when platforms internalize composition effects. Given that there are differences between traditional industries and these digital platforms, new policy protocols must be developed to address the potential market failures associated with platforms and the industries with which they interact.

To better understand composition effects, I develop a model of duopoly competition between *single-sided* platforms with network effects: platforms that charge a single price to all users.

[1]These values are current as of September 2019.

[2]For surveys, see Shy (2001), Farrell & Klemperer (2007), Birke (2009), and Shy (2011).

[3]Exceptions analyzing composition effects include games played on topological networks. See, for example, Goyal (2009) and Jackson (2010). Unlike the typical topological models, I assume neither a binary nor symmetric graph (though many other models relax these assumptions as well). Within the marketing literature, Algesheimer et al. (2005) empirically demonstrates the effects of composition using European automotive clubs. Basu (1989) is also noteworthy in his discussion of status goods, such as awards, modeling the consumption profile of commodities for which "[...] the utility [...] depends on who its other recipients (or consumers) are" (Basu 1989, p. 654).

Consumers are endowed with heterogeneous characteristics and preferences, including heterogeneous preferences over composition. I use the model to illuminate how incorporating composition changes our understanding of platforms and competition in these industries.

The model provides two major methodological contributions. First, I express a gross network valuation function with two arguments: unadjusted network size and adjusted network composition. This representation allows for a clean comparative static analysis that decomposes size and composition effects suitable for empirical estimation. With this decomposition, I precisely characterise the effects of changes in the installed base on prices and the effects of changes in nonpricing strategies on composition and therefore prices. Second, I develop a tractable metric measuring heterogeneity on a platform. This *heterogeneity-weighted network effect* reduces the high-dimensional problem of measuring heterogeneity to a single dimension. E.g. in online dating, users vary by age, race, religion, gender.... The value of adding an additional user depends on both that user's characteristics and the characteristics of current users. The heterogeneity-weighted network effect aggregates this information to a single real value. These two features are built using a structural approach relying only on observable information, allowing for both theoretical and empirical identification of the sign and magnitudes of the comparative statics and marginal effects.

To better understand how platforms operate in the new (digital) economy, a model of platforms should include three features. (i) *Heterogeneity*: individuals are potentially endowed with both heterogeneous preferences and heterogeneous preferences over heterogeneity (composition effects). (ii) *Pricing strategies*: platforms incorporate both a size and composition effect into their pricing strategies. (iii) *Non-pricing strategies*: platforms potentially have non-pricing tools at their disposal to strategically leverage their installed base through coordinating expectations.

Demographic characteristics of current users of an online dating platform affect the platform's value to a prospective user. In online video gaming, a skilled player of a massively multiplayer online video game receives positive utility when there are many other players online, but receives an extra payoff if many of the other players online are skilled as well. The skilled player may even receive disutility if many of the other players are beginners. Traits valued subjectively, such as competitiveness and sportsmanship, also influence value. Before developing the model, I present data from the market for online dating platforms to illustrate how the three aforementioned features affect market outcomes. Including heterogeneity and composition effects dramatically alters the properties of platforms along several margins. A composition effect incentivizes platforms to strategically alter their composition, affecting the pricing statics and the nature of competition.

I show that a platform's price need not be monotonic in the size of its installed base. When composition matters, there is no direct relationship between a platform's price and its size, as its composition can have a greater impact on prices than its size. Moreover, changes in size often induce changes in composition. Platforms with identical features can have identically sized installed bases while charging different prices.[4] Injecting information or decreasing the costs of information acquisition through lower search costs would not resolve price dispersion driven by composition effects. Market share and market dominance are thus no longer analogous to platform success. Instead, profitability can vary inversely with market share rather than moving in conjunction with platform size. Small platforms can both survive and thrive without vertical differentiation. Predicting the effects of changes in installed bases becomes an empirical question rather than a theoretical question. Such an environment reverses our understanding of the dominant-firm fringe-firm paradigm: the fringe can be strategically small to leverage users' composition valuations, while larger platforms rely on size to overcome potential costs of poor composition.

With composition effects, multiple equilibria exist, even with the same market shares. Suppose the market is split evenly between two platforms. If their composition is identical, they will charge identical prices. Shuffling the consumers around can increase each platform's value through composition, without altering size, uniformly raising prices. This point reiterates the idea that market share and traditional measures of "dominance" and anticompetitive activity can be misleading.

Beyond prices, a component of each platform's strategy is to shape the composition of its platform. When platforms can price discriminate, they accomplish this task using pricing strategies. In many cases, such discrimination is impossible, so the platforms rely on a nonpricing strategy to influence composition. I call this (non-price) strategy *cultivation*. Cultivation injects a new feedback loop into the relationship between installed base size and prices: directly influencing consumer expectations. It is a costly expectation-coordination mechanism, which acts as a form of equilibrium selection.[5] Returning to the online dating

[4] This price dispersion is attributable to differences in the composition of the platforms' installed bases, unlike much of the price dispersion literature, which relies on asymmetric/incomplete information. See, for example, Salop & Stiglitz (1977), Varian (1980), Burdett & Judd (1983), Stahl (1989), Sorensen (2000), and Baylis & Perloff (2002).

[5] It is similar to an endogenous process of coalitional rationalizability, a Nash refinement outlined in Ambrus (2006).

platforms example, platforms such as ChristianMingle and JDate cultivate religion. SilverSingles cultivates over age. AttractiveWorld and BeautifulPeople cultivate over physical attractiveness. In video game markets, Nintendo has cultivated a smaller but more valuable and tight-knit network than its competitors (Shankar & Bayus 2003).

Platforms can influence consumers' initial purchasing decisions through a costly investment, cultivating across the characteristics for which the consumers hold preferences.[6] Together, composition effects and cultivation are able to explain the presence of multiproduct firms in platform markets. Without composition, a firm cannibalizes its demand by splitting a single platform into multiple platforms. With composition, these splits can be demand-enhancing, increasing each user's willingness to pay even though the platform is smaller. Cultivation allows for better user segmentation, further enhancing the value of multiple platforms.

This paper is not the first to broadly consider heterogeneity and composition, though it is the first to consider horizontal composition effects within a single side of a platform.[7] It is also among the first to incorporate composition, pricing strategies, and non-pricing strategies into a single framework. For example, An & Kiefer (1995), Damiano & Li (2007), Chandra & Collard-Wexler (2009), Athey & Ellison (2011), Henkel & Block (2013), and Marx & Schummer (2019) incorporate aspects of the first two features. Through compatibility, Katz & Shapiro (1985), Farrell & Saloner (1985, 1986), Markovich (2008), Markovich & Moenius (2009), and Chen et al. (2009) incorporate aspects of the third feature. Besides compatibility, the industrial organization literature has largely ignored the third feature.[8] Many models in the economics of identity, culture, and religion have incorporated subsets of the first and third features, including Iannaccone (1992, 1994), Berman (2000), McBride (2008, 2015), Carvalho et al. (2017), and Carvalho & Sacks (2019). Similarly, the literature on the provision of local public goods has incorporated the first feature (Easterly & Levine 1997, Alesina & La Ferrara

[6]In a dynamic environment, which I discuss in Section 6.2, such cultivation creates path dependence, which the firms can leverage by augmenting the dynamics such that the probability of entering a disadvantageous state in the long run becomes vanishingly small. The firms can also use cultivation to either reinforce or eliminate preferences over composition. Thus incorporating cultivation alters many established long run predictions, such as those generated in Mitchell & Skrzypacz (2006) and Cabral (2011).

[7]Chandra & Collard-Wexler (2009) and Athey & Ellison (2011) incorporate heterogeneity on each side of a two-sided market, but composition plays no role within a side. Damiano & Li (2007) considers unidimensional heterogeneity on each side and price discrimination. White & Weyl (2016) develops a monopoly model of two-sided platforms along the lines of Laffont et al. (1998*a*) and Laffont et al. (1998*b*) but incorporating vertical heterogeneity of users (income). Marx & Schummer (2019) study two-sided monopoly matching markets with unidimensional vertical heterogeneity.

[8] One noteworthy exception is in format selection by radio stations. In two-sided markets, radio stations select a genre / format that appeals to a particular group of listeners, which influences the class of potential advertisers. See, e.g., Waldfogel (2003), Sweeting (2010), and Jeziorski (2014).

2000). Hagiu & Spulber (2013) and Veiga et al. (2017) incorporate all three features. Pricing strategies occur through purchasing access to the platform. Heterogeneity is vertical in each paper and composition plays no role within a side. In Hagiu & Spulber (2013), the platform developer can invest in content, which increases its value uniformly to all consumers and in Veiga et al. (2017), a monopoly platform developer chooses the platform characteristics according to the distribution of potential users.

Much of the previous literature assumes that (i) network benefits are homogeneous among consumers and (ii) a function only of the size of the installed base of consumers. This assumption is unrealistic, as changes in size often induce changes in composition. Exceptions to (i) are found in only a few areas. An & Kiefer (1995), Henkel & Block (2013) analyze local networks where individuals receive different network benefits. de Palma & Leruth (1996) and Janssen & Mendys-Kamphorst (2007) study network goods where individuals have heterogeneous (vertical) valuations for the networks. Ambrus & Argenziano (2009), Weyl (2010), Gomes & Pavan (2011), and White & Weyl (2016) consider heterogeneous (vertical) valuations in two-sided markets of platform competition. I consider only single-sided platforms. This article, along with Marx & Schummer (2019) present the first exceptions to both (i) and (ii), though Marx & Schummer (2019) considers only monopolistic two-sided platforms.[9] Without jettisoning (i) and (ii), multiplatform firms cannot be explained without product differentiation across the platforms, which is not always found in practice.[10]

The standard models of platforms, such as Katz & Shapiro (1985), Fudenberg & Tirole (2000), Doganoglu (2003), those presented in the surveys Farrell & Klemperer (2007) and Shy (2011), Cabral (2011), and Chen & Sacks (2018) show that absent external forces, e.g. non-linear pricing regulations à la Laffont et al. (1998*a*,*b*) (Cabral 2011), prices are monotonic in network size. Larger platforms either always set higher prices or always set lower prices on a given side. It is the horizontal composition effects influencing platform value that introduce the non-monotonicities.

Though new to the industrial organization literature, cultivation has, to an extent, been studied in the literature on the economics of identity, culture, and religion. Iannaccone (1992, 1994), Berman (2000), and Carvalho & Sacks (2019) illustrate how small, strict religions producing club goods both survive and thrive, sometimes even more so than their larger

[9]Although there is a large literature on matching platforms and heterogeneity, this body of work has not integrated platform pricing. See Marx & Schummer (2019) and the citations therein for details.

[10]Ambrus & Argenziano (2009) illustrate this possibility with vertical differentiation of users: some users are willing to pay more than others for a given sized network, allowing segmentation and two platforms to persist.

counterparts. I apply these principles to the platforms literature and shows that, even in the case of private goods, where free-riding is not an issue, similar incentives apply. Unlike other non-pricing strategies in the network and platforms literature, such as compatibility, third-party content, and first-party content, cultivation only affects the network externality offered by the platform. It is solely driven by user responsiveness.

The remainder of the paper is structured as follows. In Section 2, I describe the economic problem discussed in the introduction in the context of online dating services. Section 3 develops the model. I conduct the equilibrium analysis in Section 4. I offer antitrust implications and discuss multiproduct monopolists in Section 5. Section 6 discusses multisided platforms and dynamic extensions. Section 7 concludes.

2 Heterogeneity, Cultivation, and Dating Platforms

Online dating platforms are the quintessential example of a marketplace in which (i) users hold heterogeneous preferences over heterogeneity, (ii) platforms are willing and able to cultivate such heterogeneity, and (iii) users face identical prices.[11] The subset of users for which an individual has no interest (and are thus often ignored) still influences this individual's valuation, as these users are visible to and influence potential matches (e.g. through competition). Hence the network value of a dating platform depends on both the number of users and the traits of all other users, not just potential matches. Hitch et al. (2010*b*) show that users prefer homogeneity over a fairly large subset of characteristics such as race and religion. A preference for homogeneity along racial lines is also shown in Fisman & Iyengar (2008). Hitch et al. (2010*a*,*b*) and Fisman & Iyengar (2008) collectively show that there are differences between men and women in terms of how various characteristics, such as income and education, are valued. Internalizing that both size and composition effects are significant, platforms leverage their networks to increase profits through selective membership and matching algorithms. Therefore, heterogeneous preferences over heterogeneity is an empirical regularity that should be incorporated.

Table 1 lists the number of visitors per month and single-month prices of ten of the highest-rated online dating platforms (as of March, 2018). I aggregate two data sources to compare information from 2018 to 2008.[12] The correlation between price and the number of users per

[11] Price discrimination occurs only through rate of time preferences.

[12] The 2018 information was retrieved from http://www.10bestonline.com/top_10_best_online_dating_reviews/, accessed March 3, 2018. The 2008 data were collected by the now defunct service http://www.compete.com/ and were retained and aggregated by the blog

Table 1: Dating Websites: Prices and Number of Active Users*

	2018 Data		2008 Data		Δ Data	
Dating Website	Users	Price	Users	Price	ΔUsers	ΔPrice
BigChurch.com	85,000	$29.99	96,852	$15.95	-16,852	$14.04
Chemistry.com	366,000	$39.99	744,448	$49.95	-378,448	-$9.96
ChristianMingle.com	1,000,000	$29.99	363,919	$29.95	636,081	$0.04
eHarmony	2,300,000	$59.95	4,167,975	$59.95	-1,867,975	$0.00
Match.com	15,900,000	$35.99	5,376,096	$39.99	10,523,904	-$4.00
MatchMaker.com	215,000	$34.95	651,003	$34.95	-436,003	$0.00
PerfectMatch.com	210,000	$59.95	483,174	$59.95	-273,174	$0.00
SilverSingles.com	85,000	$59.95	·	·	·	·
SingleParentMeet.com	250,000	$14.99	366,861	$29.95	-116,861	-$14.96
Spark.com	63,000	$24.99	·	·	·	·

* Users measured in active users per month. Price is the one-month membership price. The sources for the data are provided in footnote 12. Note that minor variations in active users per month may follow from the differing sources of data. Dots correspond to missing information.

month for these ten dating platforms in 2018 is -0.05. This small and negative relationship affords the possibility that a smaller network with a "better" composition of consumers can be more valuable than a larger network with suboptimal matching. Comparisons within each individual platform across time yields a similar result. There is no consistent relationship between the price and the number of users. Composition appears to be relevant.

By creating platforms such as ChristianMingle or SingleParentMeet, developers are providing a coordination device for users. Such coordination decreases the search cost of finding a compatible match, the probability of encountering an incompatible match, and the number of competing members. The developers are engaging in endogenous product differentiation over the network externality. This mechanism is a non-pricing strategy that feeds back into prices.

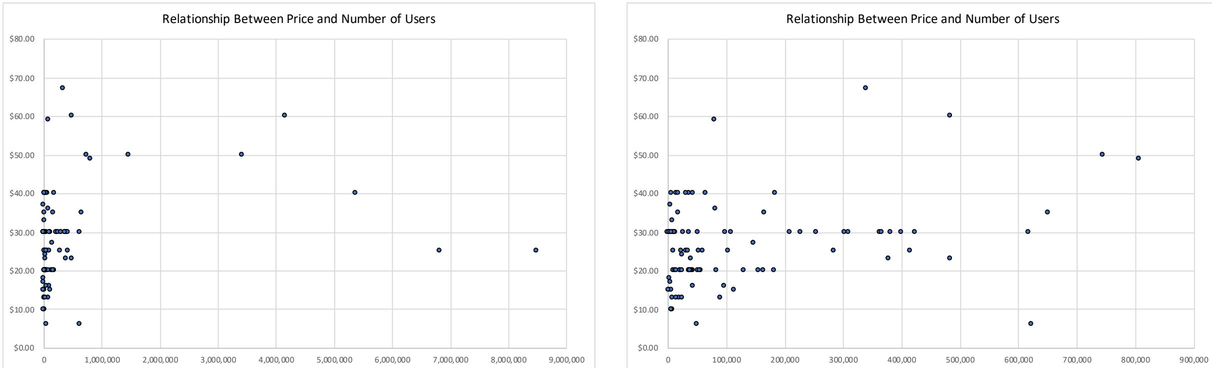
Figure 1 presents the raw relationship between prices and the number of users using the 2008 data. To formally test this relationship, I compare four models:

$$(1)\quad \ln{(price)}_i = \alpha + \beta_1 \ln{(size)}_i + \varepsilon_i$$

$$(2)\quad \ln{(price)}_i = \alpha + \beta_1 \ln{(size)}_i + \beta_2 female_i + \sum_j \beta_j race_{ji} + \varepsilon_i$$

$$(3)\quad \ln{(price)}_i = \alpha + \beta_1 \ln{(size)}_i + \beta_2 female_i + \sum_j \beta_j race_{ji} + \sum_k \beta_k age_{ki} + \varepsilon_i$$

http://www.giveyourhandabreak.com/price/#all, accessed March 3, 2018. The 2008 data consist of 104 for-pay platforms after dropping free (to the end-user) multisided platforms.



(a) All dating sites.

(b) All sites with fewer than 1,000,000 users.

Figure 1: Relationship between number of users and price - 2008 Data.

Table 2: Dating Platforms: Size Effect - 2008 Data*

	Number of users (in thousands)						
	All Websites	$\leq 1,000$	≤ 500	≤ 400	≤ 300	≤ 200	≤ 100
Size Effect (1)	0.072	0.060	0.063	0.055	0.032	0.022	0.018
SE	(0.023)	(0.030)	(0.027)	(0.029)	(0.032)	(0.036)	(0.047)
Size Effect (2)	0.097	0.088	0.089	0.080	0.056	0.046	0.034
SE	(0.023)	(0.030)	(0.027)	(0.028)	(0.034)	(0.040)	(0.048)
Wald p-val	**0.000**	**0.001**	**0.003**	**0.010**	**0.022**	**0.030**	0.167
Size Effect (3)	0.097	0.087	0.087	0.077	0.051	0.035	0.019
SE	(0.023)	(0.030)	(0.026)	(0.028)	(0.034)	(0.040)	(0.048)
Wald p-val	**0.001**	**0.005**	**0.010**	**0.025**	0.073	0.208	0.479
Size Effect (4)	0.096	0.088	0.088	0.079	0.055	0.041	0.020
SE	(0.022)	(0.029)	(0.025)	(0.027)	(0.033)	(0.041)	(0.050)
Wald p-val	**0.002**	**0.005**	**0.009**	**0.012**	**0.045**	0.142	0.916
N	104	98	93	88	81	77	67

$*$ "SE" corresponds to the robust standard error. "Wald" corresponds to the p-value of the test statistic of the one-sided Wald test of the null hypothesis $\beta_1(j) - \beta_1(1) \leq 0$ for models $j = 2, 3, 4$. Bold values identify specifications in which the difference is significant at the 5% level.

$$(4)\quad \ln{(price)}_i = \alpha + \beta_1 \ln{(size)}_i + \beta_2 female_i + \sum_j \beta_j race_{ji} + \sum_k \beta_k age_{ki} + \sum_\ell \beta_\ell income_{\ell i} + \varepsilon_i.$$

Model (1) captures the total relationship (the combined size and composition effect). Each subsequent model increasingly controls for various aspects of composition, better decomposing the size and composition effects. The size effect is given by each β_1. The variable $price_i$ is

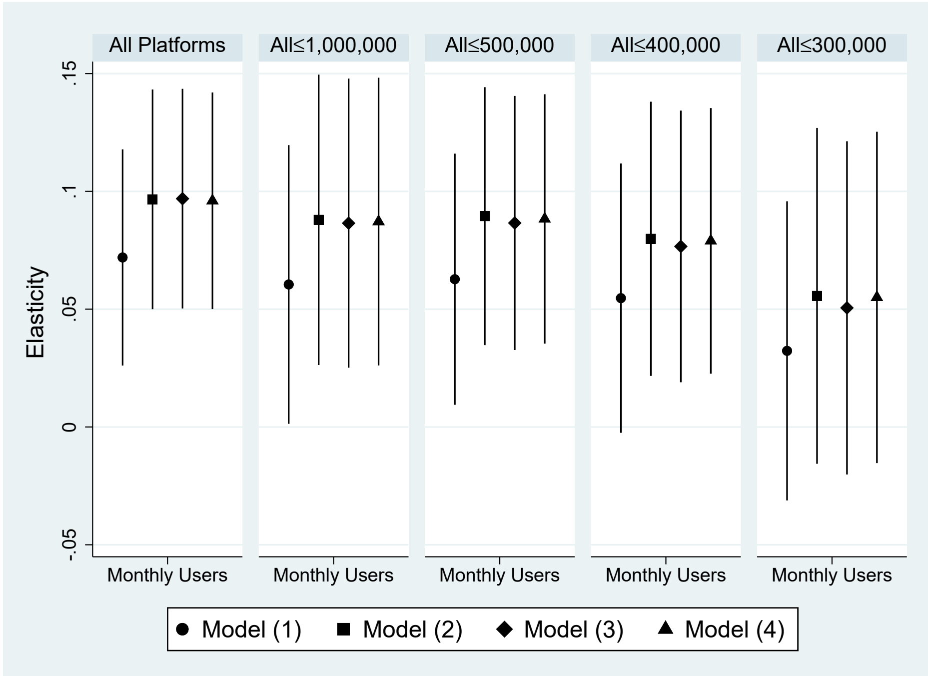


Figure 2: Size effects across various specifications.

platform i's monthly price, $size_i$ the number of users per month, $female_i$ the proportion of females, $race_i$ the proportion of Black, Asian, Hispanic, and other (nonwhite) users, age_i the proportion of users over 49, and $income_i$ the proportion of users with income above $60,000. Figure 2 offers a diagrammatic representation for several specifications. Table 2 presents the regression results for each of these specifications. After controlling for composition, the size effect increases in both size and statistical significance. To formally test the increase, I run a Wald test using the null hypothesis $\beta_1(j) - \beta_1(1) \leq 0$ for models $j = 2, 3, 4$. Figure 2 provides a diagrammatic representation of the coefficients.

Comparing the size effect in these models to the total effect in model (1), the size effect is uniformly larger in when controlling for composition, with the Wald test indicating statistical significance in most specifications.[13] This relationship indicates that the size effect is indeed positive as expected in networks and composition is an important factor in understanding both variation and changes in price. Under model (1) using the full sample, an increase in membership by one percent leads to a price increase of only 0.07%. With an average price of $27.41 and average size of 406,036 users, this elasticity corresponds to an average price increase of only 1.6 cents for every 4,060 new users. When excluding the largest sites (e.g. Ashley Madison, eHarmony, Match, OKCupid, Plenty of Fish, and Yahoo Personals), the

[13]The exceptions are because of a lack of variation due to small sample sizes.

effects are eliminated entirely and there is no relationship, statistical or economic, between the price of a dating site and the number of active users when ignoring composition.

In what follows, I decompose the size effect and composition effect so it can be represented by a single value rather than numerous controls, allowing sharp identification, both theoretically and empirically, of both the size and composition effects.

3 Model

This section develops a model of platforms competing in prices. Users are both heterogeneous and endowed with heterogeneous preferences over heterogeneity. The platforms are able to cultivate this heterogeneity to their advantage, though such cultivation is costly.

3.1 Preliminaries

There are two risk-neutral profit-maximizing platforms (A and B), indexed by j, and n utility-maximizing users, indexed by i with large n. Each platform j produces a single network good (the platform itself) at zero marginal cost, selling access at price p_j. Denote by $\mathbf{p} = (p_A, p_B)$ the vector of platforms' prices. Each user is endowed with a set of traits $y = (y_1, \ldots, y_T) \in Y$. Each y_τ is a trait, which itself an element of Y_τ. Hence $Y = Y_1 \times \cdots \times Y_T$ (the cartesian product). There are $\overline{Y} = \prod_\tau |Y_\tau|$ types in the population and $\hat{Y} = \sum_\tau |Y_\tau|$ distinct traits.[14]

Example 1. *Suppose that* $T = 3$. Y_1 *is the set of ages,* Y_2 *the set of genders, and* Y_3 *the set of religions. For simplicity, suppose that* $Y_1 = \{20, 30, 40\}$, $Y_2 = \{\mathit{male}, \mathit{female}\}$ *and* $Y_3 = \{\mathit{Christian}, \mathit{Jewish}, \mathit{Muslim}, \mathit{other}\}$. *There are* $\overline{Y} = 3 \times 2 \times 4 = 24$ *sets of traits and* $\hat{Y} = 3 + 2 + 4 = 9$ *traits, with a* 20 *year old female Christian denoting one specific type.*

To avoid confusion, I use ℓ and m to index arbitrary types (rather than y and y') and y_r to index arbitrary traits. Denote by $n_\ell = |\ell| > 0$ the number of type-ℓ users. Specifying the model by linking traits to types provides a structured approach to empirically modeling heterogeneity and composition.[15]

[14]When $\overline{Y} > 2$, there need not be a one-to-one mapping between traits and types, qualitatively affecting the results in Section 4.2. Hence I consider a general $\overline{Y}$ with a one-to-one mapping as a special case.

[15]In cases where there is correlation across types, unsupervised machine learning methods, e.g. k-means/modes clustering or Gaussian mixture models, can be used to partition the $\overline{Y}$ trait profiles into L bins, with clustering at the bin-level.

Each type ℓ places a value on the presence of a type m possessing trait $y_r \in Y_\tau$ for each Y_τ: $x_{\ell m y_r} \in \{-1, 0, 1\}$, where 1 represents a desirable trait, 0 a neutral trait, and -1 an undesirable trait.[16] E.g. a type may assign a 1 to all types with the same religion and a -1 to all types with a different religion. These valuations create, for each trait Y_τ, a $\overline{Y} \times \overline{Y}$ directed graph. The graphs are stacked to build an $\overline{Y} \times \overline{Y} \times T$ tensor X. As not all traits are equally valued, each layer of the tensor is assigned a weight α_τ. The weight signifies the importance of the trait. E.g. religion is likely more important than a type's number of pets, so the α associated with religion will be larger than the α associated with pets. The tensor is then collapsed into an $\overline{Y} \times \overline{Y} \times 1$ directed graph H with elements $h_{\ell m}$ given by

$$h_{\ell m} = \sum_{\tau=1}^{T} \alpha_\tau x_{\ell m y_r}. \tag{1}$$

If $h_{\ell m} = 0$, then a type-ℓ user receives neither a premium nor a penalty from the presence of a type-m user. Positive values convey premiums while negative values convey penalties. Figure 3 offers an example for $\overline{Y} = 4$.

Remark 1. *When the $h_{\ell m}$ are observed, the α_{y_r} can be estimated using $h_{\ell m} = \sum_{y_r} \alpha_{y_r} x_{\ell m y_r} + \varepsilon_{\ell m}$. In many applications the $h_{\ell m}$ are unobserved latent values, in which case a binary representation*

$$h^*_{\ell m} = \begin{cases} 1 & \text{if } h_{\ell m} > 0 \\ 0 & \text{if } h_{\ell m} \leq 0 \end{cases}$$

often is observed. In online dating, whether or not a specific type chooses to match with another is observed, while the underlying latent weight $h_{\ell m}$ is not. Using the discrete $h^*_{\ell m}$ as the dependent variable, the α_{y_r} are estimated using a Gibbs Sampler with data augmentation. The latent $h_{\ell m}$ are pulled from the data augmentation step.[17]

Users hold inherent idiosyncratic preferences $\zeta_j \gg 0$ for each platform j. I assume that ζ_j is sufficiently large to ensure market coverage. Hence the relevant decision is the relative preference for platform A: $\xi_A \equiv \zeta_A - \zeta_B$, which is distributed according to the CDF $\Phi(\xi)$ with PDF $\phi(\xi)$.

Assumption 1. *(i) $\Phi(\cdot)$ is continuously differentiable. (ii) $\phi(\xi) > 0$ for all ξ and (iii) $\phi(\xi) = \phi(-\xi)$. (iv) $\Phi(\xi)/\phi(\xi)$ is strictly increasing in ξ. (v) $\phi(\xi|z) = \phi(\xi)$.*

[16] The results generalize when allowing any value in $\mathbb{R}$.

[17] For details, see Albert & Chib (1993) and Vossmeyer (2014).

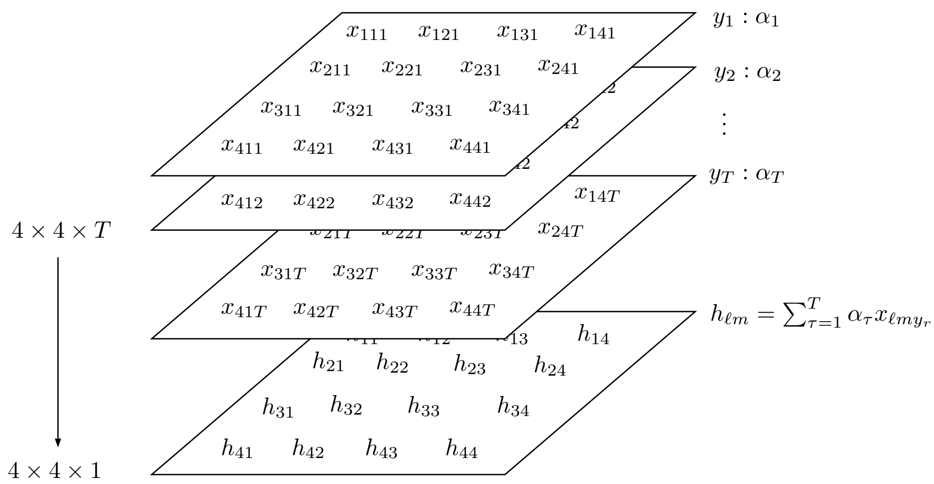


Figure 3: Visual representation flattening of X to H with $\overline{Y} = 4$.

Items (i) and (ii) ensure that the demand curves are well behaved and the platforms' profit functions are quasiconcave, while item (iii) ensures that asymmetry occurs only through pricing and market shares. Many continuous distributions satisfy item (iv), which is standard in the literature.[18] Two consumers of the same type may, and often do, possess different preferences ξ_A, allowed by item (v).

3.2 Platform Behavior

The platforms make two sequential choices: cultivation followed by pricing. Cultivation is a costly coordination mechanism that operates by altering consumer expectations prior to the pricing subgame. Through costly investment, platforms are able to target and attract traits, influencing all types possessing those targeted traits. For example, by targeting Jewish singles, a dating platform attracts both Jewish males and females, those with and without children, and so on.

The platforms simultaneously and independently set cultivation vectors $\mathbf{c}_j = (c_{j1}, \ldots, c_{j\hat{Y}}) \in [0,1]^{\hat{Y}}$, where c_{jy_r} corresponds to the investment made by platform j in cultivating trait y_r. Denote by $\mathbf{c} = (\mathbf{c}_A, \mathbf{c}_B)$ the collection of cultivation decisions.

Costs take the form $\kappa(\mathbf{c}_j)$, which may consist of both monetary and nonmonetary components.

Assumption 2. *(i)* $\partial\kappa(\mathbf{c}_j)/\partial c_{jy_r} > 0$ *for all* y_r *and* $c_{jy_r} \in (0,1)$. *(ii)* $\partial^2\kappa(\mathbf{c}_j)/\partial c_{jy_r}^2 > 0$ *for*

[18] Assumptions 1(i)-(iv) collectively mirror Assumption 1 of Cabral (2011, p. 88).

all y and $c_{jy_r} \in (0,1)$. (iii) $\partial^2\kappa(\mathbf{c}_j)/\partial c_{jy_r}\partial c_{jy_s} > 0$ for $y_s \neq y_r$. (iv) $\lim_{c_{jy_r}\to 0} \partial\kappa(\mathbf{c}_j)/\partial c_{jy_r} = 0$ and $\lim_{c_{jy_r}\to 1} \partial\kappa(\mathbf{c}_j)/\partial c_{jy_r} = \infty$ for all y_r.

Again, cultivation operates by influencing consumer expectations via coordination, which is denoted by the superscript 'e'. Prior to cultivation, the consumers' (prior) expectations are given by $q^e_{j\ell}$. Consumers then receive a signal: the observed cultivation vector. The (posterior) expected number of type-ℓ users of platform j is given by $q^e_{j\ell}(\mathbf{c})$. The expected installed base of platform j is $q^e_j(\mathbf{c}) = \sum_\ell q^e_{j\ell}(\mathbf{c})$. I assume that the expectations are updated according to the function $q^e_{j\ell}(\mathbf{c}) = f(q^e_{j\ell}, \mathbf{c})$, which is increasing in c_{jy_r}, decreasing in c_{-jy_r} for all y_r present in type ℓ, and satisfies $f(\mathbf{q}^e_\ell, 0) = q^e_{j\ell}$. I also require traits with a greater α to have a greater marginal effect of an increase in c_{jy_r}.[19] For tractability, suppose that f takes the piecewise-linear form:

$$f(q^e_{j\ell}, \mathbf{c}) = \min\left\{n_\ell, \max\left\{0, q^e_{j\ell} + \frac{\sum_{y_\tau \in \ell} \alpha_\tau (c_{jy_\tau} - c_{-jy_\tau})}{\sum_{y_\tau \in \ell} \alpha_\tau}\right\}\right\}; \tag{2}$$

however, all results hold for any f satisfying the above properties. This representation formalizes cultivation as a coordination mechanism. Symmetry in cultivation is imposed across both platforms. Denote by $\mathbf{q}^e(\mathbf{c})$ the collection of expectations.

Definition 1. *Define a type ℓ's expected valuation of the heterogeneity-weighted network effects provided by platform j as*

$$\tilde{q}^e_{j\ell}(\mathbf{c}) = \sum_{m=1}^{\overline{Y}} q^e_{jm}(\mathbf{c}) h_{\ell m}. \tag{3}$$

Remark 2. *While cultivation is, at its core, a location decision, there are substantial differences between cultivation and the typical Hotelling-style location game. Firstly, user "locations" are not primitive, but instead depend on both the distribution of users in the population and their expectations via $\{n_\ell\}_\ell$ and $\tilde{q}$. Similarly, the ideal location choice of the platform depends on both, adding a feedback loop not present in typical location models.*

As in $\mathbf{q}^e(\mathbf{c})$, denote by $\tilde{\mathbf{q}}^e(\mathbf{c})$ the collection of expected heterogeneity-weighted market shares. After observing the cultivation decisions, the platforms simultaneously and independently set their prices, taking user expectations as given.

[19] This assumption is natural. Cultivation has a greater impact on more important traits.

3.3 User Behavior

Every individual has unit demand and chooses between one of the two platforms. By purchasing access to platform j, a type-ℓ user with preferences ζ_j receives expected utility

$$u^e(j,\ell) = \zeta_j + v\left(q_j^e(\mathbf{c}), \theta\tilde{q}_{j\ell}^e(\mathbf{c})\right) - p_j, \tag{4}$$

where $v\left(q_j^e(\mathbf{c}), \theta\tilde{q}_{j\ell}^e(\mathbf{c})\right)$ is the value of the installed base (network effect) of platform j to this individual and $\theta \geq 0$ is the marginal value of composition. If $\theta = 0$, then user composition is irrelevant. The relative importance of composition to size is increasing in θ. The realized utility is given by the same functional, substituting the realized values in for the expected values and dropping the superscripts 'e'.

3.4 Network Effects

The expected (and realized) value of the network associated with each platform depends on two factors: the platform's size and its heterogeneity-weighted network effect.

Assumption 3. *(i) For all $q_j^e(\mathbf{c}) > 0$ and $\theta > 0$, $\partial v\left(q_j^e(\mathbf{c}), \theta\tilde{q}_{j\ell}^e(\mathbf{c})\right)/\partial\tilde{q}_{j\ell} > 0$. (ii) For all fixed $\tilde{q}_{j\ell}^e(\mathbf{c})$, $\partial v\left(q_j^e(\mathbf{c}), \theta\tilde{q}_{j\ell}^e(\mathbf{c})\right)/\partial q_j(\mathbf{c}) > 0$. (iii) For fixed $q_j^e(\mathbf{c})$ and $q_j^e(\mathbf{c})'$, $v\left(q_j^e(\mathbf{c}), \theta\tilde{q}_{j\ell}^e(\mathbf{c})\right) - v\left(q_j^e(\mathbf{c})', \theta\tilde{q}_{j\ell}^e(\mathbf{c})'\right)$ is increasing in $\theta[\tilde{q}_{j\ell}^e(\mathbf{c}) - \tilde{q}_{j\ell}^e(\mathbf{c})']$.*

These same properties carry over to the realized values $q_j(\mathbf{c})$ and $\tilde{q}_{j\ell}(\mathbf{c})$.

4 Results

I utilize the subgame-perfect Nash equilibrium (SPE) solution concept. In the pricing stage, I find the coordination problem inducing multiple equilibria discussed in the literature to be more severe than suggested, as there are multiple dimensions to coordinate over. Those equilibria that, upon an increase in size yield a less desirable composition, may fail to satisfy the monotone comparative static property. The set of prices obtainable in equilibrium is significantly wider than that without composition.

In the cultivation stage, I illustrate endogenous product differentiation over the network externality. The product differentiation acts as a form of equilibrium selection. Depending on the nature of heterogeneity and the desirability of composition, cultivation can either soften or strengthen competitive forces. Such competition can lead to higher prices and a less competitive market, lower prices and a more competitive market, or a mixture with increased price dispersion. A platform may opt for a smaller sized network if it corresponds to a more valuable composition, allowing its competitor to grow large, but with a less valuable composition. Both platforms may benefit from such a scenario.

4.1 Pricing Stage

Fix the cultivation profile at $\mathbf{c}$. By (4) and market coverage, for all user expectations $\mathbf{q}^e$, $\overline{Y}$ cutoff values can be defined:

$$\underbrace{\zeta_A - \zeta_B = \xi_A}_{\equiv \omega_\ell} = \underbrace{\left[p_A - v\left(q_A^e(\mathbf{c}), \theta \tilde{q}_{A\ell}^e(\mathbf{c})\right)\right]}_{\substack{\textbf{expected hedonic price} \\ \textbf{of platform } A}} - \underbrace{\left[p_B - v\left(q_B^e(\mathbf{c}), \theta \tilde{q}_{B\ell}^e(\mathbf{c})\right)\right]}_{\substack{\textbf{expected hedonic price} \\ \textbf{of platform } B}},$$

with each cutoff corresponding to the type-ℓ user indifferent between platforms A and B.[20] When there is no confusion, I drop the cultivation profile $\mathbf{c}$. Using these cutoff values, platform A's demand is $D_A(p_A, p_B, \mathbf{q}^e) = n - \sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell)$, and platform B's demand is $D_B(p_B, p_A, \mathbf{q}^e) = \sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell)$.[21] Denote by $p_A^* = p_A(\mathbf{q}^e)$, $p_B^* = p_B(\mathbf{q}^e)$, and $\mathbf{p}(\mathbf{q}^e) = (p_A(\mathbf{q}^e), p_B(\mathbf{q}^e))$ the prices in the equilibrium of the subgame induced by cultivation profile $\mathbf{c}$, where in each equilibrium $\mathbf{q}^e = \mathbf{q}$ (expectations are fulfilled):

$$\begin{aligned} p_A^* &\equiv \arg\max \left\{p_A D_A(p_A, p_B^*, \mathbf{q}^e)\right\} \\ p_B^* &\equiv \arg\max \left\{p_B D_B(p_B, p_A^*, \mathbf{q}^e)\right\}. \end{aligned} \tag{5}$$

The results of the pricing stage are also interpretable as outcomes of a single-stage game in which there is no cultivation, only composition effects.

Lemma 1. *If $\mathbf{p}(\mathbf{q}^e)$ and $\hat{\mathbf{p}}(\mathbf{q}^e)$ both satisfy* (5)*, then $\mathbf{p}(\mathbf{q}^e) = \hat{\mathbf{p}}(\mathbf{q}^e)$.*

Lemma 1 does not imply that there is a unique SPE, rather that there is a unique equilibrium of each subgame induced by cultivation profile $\mathbf{c}$. Define

$$\omega_\ell^* = \left[p_A(\mathbf{q}^e) - v\left(q_A^e, \theta \tilde{q}_{A\ell}^e\right)\right] - \left[p_B(\mathbf{q}^e) - v\left(q_B^e, \theta \tilde{q}_{B\ell}^e\right)\right]$$

as the cutoff value for a type-ℓ user in the equilibrium of the subgame.

Lemma 2. *For every set of expectations $\mathbf{q}^e$, there exists a representative user endowed with $\xi_A = \omega_0$, defined as the individual whose demand corresponds to the representative demand: $\Phi(\omega_0) = n^{-1} \sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell)$.*

[20] If there type-ℓ users in platform A's installed base, then $\Phi(\omega_\ell) = 1$.

[21] I omit the $\tilde{\mathbf{q}}^e$ from the demand expressions, as the exogenous $h_{\ell m}$ and the expectations $\mathbf{q}^e$ are sufficient in characterising $\tilde{\mathbf{q}}^e$.

This individual possesses traits that are a weighted representation of all individuals in the economy. For convenience, I define this user's expected heterogeneity-weighted network effect as $\tilde{q}^e_{j0}$ and type as 0, which is interpreted as an $\overline{Y}+1^{th}$ type (if there is no type ℓ such that $0=\ell$). Lemma 2 offers an alternative interpretation to the pricing stage where there is only a single type-0 user on the market, with the remaining users locked in. The static model is then analogous to the stage game of a single-mover overlapping generations framework.

Definition 2. *If $\tilde{q}_{j0} > 0$, then network j exhibits a heterogeneity premium [$H+$ (H-plus)]. If $\tilde{q}_{j0} < 0$, then network j exhibits a heterogeneity penalty [$H-$ (H-minus)]. If $\tilde{q}_{j0} = 0$, then network j is heterogeneity neutral [$H0$ (H-naught)].*

The typical comparative static of pricing with respect to the installed base in network industries considers $\partial p_j(\mathbf{q}^e)/\partial q^e_j$.

Remark 3. *The complete comparative static of pricing with respect to the installed base is*

$$\underbrace{\frac{\partial p_j(\mathbf{q}^e)}{\partial q^e_j}}_{\textbf{size effect}} + \underbrace{\frac{\partial p_j(\mathbf{q}^e)}{\partial \tilde{q}^e_{j0}}\frac{\partial \tilde{q}^e_{j0}}{\partial q^e_j}}_{\textbf{composition effect}} . \tag{6}$$

This statement follows structurally from the theorem of the maximum. The traditional models of platforms often only incorporate size effects ($H0$ networks) and can therefore be interpreted as special cases of this model by either taking $\theta \to 0$ or assuming $H = \mathbf{0}_{\overline{Y}\times\overline{Y}}$. The models of local network effects, e.g. An & Kiefer (1995), Henkel & Block (2013), and the citations therein, can also be derived as special cases by assuming that local (known) contacts are given zero weight and nonlocal (unknown) contacts are given negative weight ($H-$ networks), discounting the network effect. Remark 3 is formalized as follows.

Proposition 1. *Under Lemma 2,*

$$\frac{dp_A(\mathbf{q}^e)}{dq^e_{A\ell}} = \frac{1+\frac{\phi'(\omega^*_0)[1-2\Phi(\omega^*_0)]}{\phi(\omega^*_0)^2}}{3+\frac{\phi'(\omega^*_0)[1-2\Phi(\omega^*_0)]}{\phi(\omega^*_0)^2}} \left[\underbrace{\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial q^e_{A\ell}}}_{\textbf{size effect}} + \underbrace{\theta\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial \tilde{q}^e_{A0}}h_{0\ell}}_{\textbf{composition effect}}\right] \tag{7}$$

$$\frac{dp_B(\mathbf{q}^e)}{dq^e_{A\ell}} = \frac{\frac{\phi'(\omega^*_0)\Phi(\omega^*_0)}{\phi(\omega^*_0)^2}-1}{3+\frac{\phi'(\omega^*_0)[1-2\Phi(\omega^*_0)]}{\phi(\omega^*_0)^2}} \left[\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial q^e_{A\ell}} + \theta\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial \tilde{q}^e_{A0}}h_{0\ell}\right]. \tag{8}$$

Moreover, there exists a cutoff value $\theta^*(\mathbf{q}^e)$ such that for all expectations $q^e_A \le q^e_B$:

(i) $q^e_A \leq q^e_B$ *but* $p_A(\mathbf{q}^e) > p_B(\mathbf{q}^e)$,

(ii) *if* $\theta > \theta^*(\mathbf{q}^e)$ *and* $h_{0\ell} < 0$, *then* $dp_A(\mathbf{q}^e)/dq^e_{A\ell} < 0$.

Proposition 1 possesses two noteworthy implications. First, the effect of a platform's price with respect to a change in the size of its own installed base [equation (7)] cannot be generically signed. The sign of the effect depends on the type of user being added and the relationship between the user's type and those of the current installed base. Note that

$$\left(1 + \frac{\phi'(\omega^*_0)[1 - 2\Phi(\omega^*_0)]}{\phi(\omega^*_0)^2}\right) \Bigg/ \left(3 + \frac{\phi'(\omega^*_0)[1 - 2\Phi(\omega^*_0)]}{\phi(\omega^*_0)^2}\right) > 0 \ ,$$

as is the size effect and the first term of the composition effect: $\theta\partial v(q^e_A, \theta\tilde{q}^e_{A0})/\partial\tilde{q}^e_{A0}$. A direct relationship between a platform's price and the size of its installed base is guaranteed only if either $\theta = 0$ or $h_{0\ell} \geq 0$. If $\theta > 0$ and $h_{0\ell} < 0$, then the relationship can be non-monotonic depending on the current state and the magnitude of θ. The effect of a change in the installed base is an empirical question.

Second, when composition effects are present, the effect of a change in the size of a platform's installed base can have an indeterminate effect on a competing platform's price. When composition effects are absent, if one platform grows relative to another, then all else equal, the competing platform must compensate by lowering its price. When composition effects are present, if a platform increases its size, but that increased size comes at the cost of a poorer composition, then the relative value of the unchanged competing platform increases. A platform can increase its size but worsen its composition, which can induce a decrease in price while allowing a competing platform to increase its price: price dispersion without asymmetric information, search, or other frictions.

The importance of composition effects can be illustrated with a simple example. Suppose platform A currently has Q users and there are three potential users to add: a type-1 user, a type-2 user, and a type-3 user. In each case, adding a user increases the size to $Q + 1$; however, adding the type-1 user yields $dp_A/dq > 0$ and $dp_B/dq < 0$, adding the type-2 user yields $dp_A/dq = 0$ and $dp_B/dq = 0$, and adding the type-3 user yields $dp_A/dq < 0$ and $dp_B/dq > 0$. To see this pattern clearly, suppose that $\Phi \sim U$. Then $\phi'(\cdot) = 0$ and the first term in dp_A/dq becomes $1/3$ and the first term in dp_B/dq becomes $-1/3$. The outcome follows immediately from (7) and (8).

The differentiability of $\phi(\cdot)$ and $v(q_j, \theta\tilde{q}^e_{j0})$ are not necessary for the patterns described in the proposition to hold, but are necessary for analytically decomposing the size and composition effects.[22] Moreover, (7) and (8) can also be derived without invoking Lemma 2. Define

$$\overline{\Phi} = \sum_m \frac{n_m}{n} \Phi(\omega_m), \qquad \overline{\phi} = \sum_m \frac{n_m}{n} \phi(\omega_m), \qquad \overline{\phi}' = \sum_m \frac{n_m}{n} \phi'(\omega_m).$$

Then (7) and (8) are represented by the solution to

$$\begin{pmatrix} -2 - \frac{\overline{\phi}'\left(1-\overline{\Phi}\right)}{\overline{\phi}^2} & 1 + \frac{\overline{\phi}'\left(1-\overline{\Phi}\right)}{\overline{\phi}^2} \\ 1 - \frac{\overline{\phi}'\overline{\Phi}}{\overline{\phi}^2} & -2 + \frac{\overline{\phi}'\overline{\Phi}}{\overline{\phi}^2} \end{pmatrix} \begin{pmatrix} \frac{dp_A(\mathbf{q}^e)}{dq^e_{A\ell}} \\ \frac{dp_B(\mathbf{q}^e)}{dq^e_{A\ell}} \end{pmatrix} = \begin{pmatrix} -\frac{\sum_m \frac{n_m}{n}\left(\phi(\omega_m)\overline{\phi}+\phi'(\omega_m)[1-\overline{\Phi}]\right)\Omega_m}{\overline{\phi}^2} \\ \frac{\sum_m \frac{n_m}{n}\left(\phi(\omega_m)\overline{\phi}-\phi'(\omega_m)\overline{\Phi}\right)\Omega_m}{\overline{\phi}^2} \end{pmatrix},$$

where

$$\Omega_m = \frac{\partial v(q^e_A, \theta \tilde{q}^e_{Am})}{\partial q^e_{A\ell}} + \theta \frac{\partial v(q^e_A, \theta \tilde{q}^e_{Am})}{\partial \tilde{q}^e_{Am}} h_{m\ell}.$$

Ω_m is analogous to the composition and size effects illustrated in Proposition 1, weighted and aggregated across all types.

The data presented in Table 1 of Section 2 illustrates both the importance of the second term in Remark 3 and the above example. The size effect is only sufficient to explain variation in price when there is either no change in composition or no effect of composition. Take the dating website SingleParentMeet as an example. Single-parenthood is the relevant characteristic for composition. Such composition is unlikely to change as the number of users changes. Equations (6) and (7) suggest a direct relationship between the price and the number of users, which is consistent with the data. A decrease of 116,861 users corresponds to a per-month price decrease of \$14.96 (an arc elasticity of 0.64). When observing dating sites with more heterogeneity among users, such as Chemistry, eHarmony, or Match, no consistent patterns emerge. Between these three sites, there is a positive relationship between the price and the number of users, no relationship, and a negative relationship. The size effect is not sufficient in deriving the comparative static, owning the need for accounting for the composition effect in (6) and (7). Those sites illustrating no relationship between price and number of users indicate a potentially balanced trade-off between value derived from size and value derived from composition. Although PerfectMatch lost approximately 275,000 users, the potential change in the distribution of users (potentially) lead to a more perfect match for remaining users, yielding a net effect of zero on price.

Proposition 1 relates the findings of Section 2 to the model via H and θ. While the size effect is positive as the previous literature (and Section 2) indicates, the composition effect

[22] This statement is proven in the proof of Proposition 1.

varies in both sign and magnitude, leading to an indeterminate total effect, as illustrated by the data in Tables 1 and 2 and Figures 1 and 2.[23]

Proposition 1 implicitly assumed that platform B's installed base was unchanged (the addition of a user to the market). It is also worthwhile to augment the analysis to understand the effect of users switching platforms.

Proposition 2. *Under Lemma 2 and assuming a transfer of users from platform B to A,*

$$\frac{dp_A(\mathbf{q}^e)}{dq^e_{A\ell}} = \frac{1+\frac{\phi'(\omega_0^*)[1-2\Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}}{3+\frac{\phi'(\omega_0^*)[1-2\Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}} \left[\underbrace{\left(\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial q^e_{A\ell}} + \frac{\partial v(q^e_B,\theta\tilde{q}^e_{B0})}{\partial q^e_{B\ell}}\right)}_{\textbf{size effect}} + \theta \underbrace{\left(\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial \tilde{q}^e_{A0}} + \frac{\partial v(q^e_B,\theta\tilde{q}^e_{B0})}{\partial \tilde{q}^e_{B0}}\right) h_{0\ell}}_{\textbf{composition effect}}\right] \tag{9}$$

$$\frac{dp_B(\mathbf{q}^e)}{dq^e_{A\ell}} = \frac{\frac{\phi'(\omega_0^*)\Phi(\omega_0^*)}{\phi(\omega_0^*)^2}-1}{3+\frac{\phi'(\omega_0^*)[1-2\Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}} \left[\left(\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial q^e_{A\ell}} + \frac{\partial v(q^e_B,\theta\tilde{q}^e_{B0})}{\partial q^e_{B\ell}}\right) + \theta \left(\frac{\partial v(q^e_A,\theta\tilde{q}^e_{A0})}{\partial \tilde{q}^e_{A0}} + \frac{\partial v(q^e_B,\theta\tilde{q}^e_{B0})}{\partial \tilde{q}^e_{B0}}\right) h_{0\ell}\right]. \tag{10}$$

Moreover, there exists a cutoff value $\theta^{**}(\mathbf{q}^e)$ such that for all expectations $q^e_A \leq q^e_B$:

(i) $q^e_A \leq q^e_B$ *but* $p_A(\mathbf{q}^e) > p_B(\mathbf{q}^e)$,

(ii) *if* $\theta > \theta^{**}(\mathbf{q}^e)$ *and* $h_{0\ell} < 0$, *then* $dp_A(\mathbf{q}^e)/dq^e_{A\ell} < 0$.

Unfortunately, the relationship between $\theta^*(\mathbf{q})$ and $\theta^{**}(\mathbf{q})$ cannot be determined without placing further assumptions on $v(q^e_j, \theta\tilde{q}^e_{j0})$.[24] For a given θ, the size effect is compounded, placing upward pressure on the platform's own price and downward pressure on the competing platform's price. An increase in the size of platform A's installed base necessitates a decrease in the size of platform B's installed base. The composition effect is compounded as well under a transfer of users, though this compounding effect can place either upward or downward pressure on price. When $h_{0\ell} > 0$, the composition effect compounds the price effect, as platform A's composition value increases while platform B's decreases. When $h_{0\ell} < 0$ the opposite occurs, putting downward pressure on the price.

The marginal value of composition θ bounds the composition effect. When θ is sufficiently small, the network is approximately $H0$. In $H0$ networks, the size effect will always dominate the composition effect. The monotone-pricing relationship described in Katz & Shapiro (1985), Cabral (2011), and those in Farrell & Klemperer (2007) and Shy (2011) can thus be expected to hold in network industries such as software and telecommunications. Empirical

[23]In dynamic settings, the size effect can potentially be negative, as in Cabral (2011, p. 84) and Fudenberg & Tirole (2000). Nonetheless, a positive composition effect still implies indeterminacy.

[24]The relationship between $\theta^*(\mathbf{q})$ and $\theta^{**}(\mathbf{q})$ is determined by the relative differences between the size and composition effects in Propositions 1 and 2.

evidence for monotone pricing in software is found in Gandal (1994) and Brynjolfsson & Kemerer (1996). Yet in many other network industries, such as internet dating and video games, composition effects can represent a strong factor and these effects exert significant force over the pricing.[25] When θ is non-negligible, signing the effect of an increase in the size of the installed base requires knowledge of the type being added and how the various types interact with one-another ($h_{0\ell}$).

Proposition 3. *The effects of H and θ on pricing are as follows.*

(i) *If network j is $H+$, then $\frac{p_j(\mathbf{q}^e)}{d\theta} > 0$.*

(ii) *If network j is $H-$, then $\frac{p_j(\mathbf{q}^e)}{d\theta} < 0$.*

(iii) *If network j is $H0$, then $\frac{p_j(\mathbf{q}^e)}{d\theta} = 0$.*

When θ increases, platforms have a greater incentive to positively affect composition in order to cultivate an $H+$ network. If both platforms are able to cultivate an $H+$ network, then both platforms benefit. For large enough θ, the composition effect exceeds the size effect and a small increase (or decrease) in market share can lead to large swings, both positive and negative, in price. These are $H+$ and $H-$ networks. Such markets include online dating websites, social networks, and video games. Shankar & Bayus (2003) provides empirical evidence in the video game industry. A change in the size of a platform's installed base need not correspond to a direct change in the price. Nearly every change in market share is accompanied by a corresponding change in composition, except for the rare case in which the proportions of all types remains unchanged.

An increase in installed base, when coupled with an increase in the relative heterogeneity premium (or decrease in the heterogeneity penalty), implies compounding effects and monotone price changes. An increase in market share, when coupled with a decrease in the relative heterogeneity premium (or increase in the heterogeneity penalty), implies countering effects, which can lead to lower prices for larger networks.

Let $\pi_j(p(\mathbf{q}^e))$ denote platform j's profits in the equilibrium of the subgame.

Corollary 1. *For all expectations $q^e_A \leq q^e_B$ with $\tilde{q}^e_{A0} > \tilde{q}^e_{B0}$, there exists a cutoff value $\overline{\theta}(\mathbf{q}^e)$ such that $\pi_A(p_A(\mathbf{q}^e)) > \pi_B(p_B(\mathbf{q}^e))$ if and only if $\theta > \overline{\theta}(\mathbf{q}^e)$.*

[25]For more details on the empirical networks literature, see Birke (2009).

Corollary 1 is a direct implication of Propositions 1-3. Composition effects can become important enough that the smaller platform with a better composition can set a higher price than a larger competitor. This price difference can be so severe that the smaller platform is also more profitable than its larger competitor.

Proposition 4. *Under Lemma 2 and assuming a transfer of users from platform B to A,*

$$\frac{\partial \pi_A(\mathbf{p}(\mathbf{q}^e))}{\partial q^e_{A\ell}} = np_A(\mathbf{q}^e)\phi(\omega^*_0)\left[\frac{\partial v(q^e_A, \theta\tilde{q}^e_{A0})}{\partial q^e_{A\ell}} + \frac{\partial v(q^e_B, \theta\tilde{q}^e_{B0})}{\partial q^e_{B\ell}} + \theta\left(\frac{v(q^e_A, \theta\tilde{q}_{A0})}{\partial \tilde{q}^e_{A0}} + \frac{v(q^e_B, \theta\tilde{q}_{B0})}{\partial \tilde{q}^e_{B0}}\right)h_{0\ell}\right] \quad (11)$$

$$\frac{\partial \pi_B(\mathbf{p}(\mathbf{q}^e))}{\partial q^e_{A\ell}} = -np_B(\mathbf{q}^e)\phi(\omega^*_0)\left[\frac{\partial v(q^e_A, \theta\tilde{q}^e_{A0})}{\partial q^e_{A\ell}} + \frac{\partial v(q^e_B, \theta\tilde{q}^e_{B0})}{\partial q^e_{B\ell}} + \theta\left(\frac{v(q^e_A, \theta\tilde{q}_{A0})}{\partial \tilde{q}^e_{A0}} + \frac{v(q^e_B, \theta\tilde{q}_{B0})}{\partial \tilde{q}^e_{B0}}\right)h_{0\ell}\right]. \quad (12)$$

Therefore, $\text{sign}\left\{\partial\pi_A(\mathbf{p}(\mathbf{q}^e))/\partial q^e_{A\ell}\right\} \neq \text{sign}\left\{\partial\pi_B(\mathbf{p}(\mathbf{q}^e))/\partial q^e_{A\ell}\right\}$.

If, instead of a transfer from platform B to A, platform B's installed base remains unchanged, then $\partial v(q^e_B, \theta\tilde{q}^e_{B0})/\partial q^e_{B\ell} = v(q^e_B, \theta\tilde{q}_{B0})/\partial\tilde{q}^e_{B0} = 0$. Like Propositions 1 and 2, the signs of (11) and (12) depend explicitly on θ and $h_{0\ell}$. The potential non-monotonicities described in Propositions 1 and 2 carry over to platform profits. If a platform adds users that do not mesh well with the existing base, the platform becomes less valuable to its existing users and thus less profitable as it grows. Equation (11) illustrates the potential for a niche construction strategy: by purging users that are not a good fit, the platform can shrink in size but increase its value to existing users, increasing profits.[26] Unlike the Groucho Marx adage, "I don't want to belong to any club that will accept people like me as a member," there is no reliance on some vertical measure of differentiation, but rather a horizontal measure. Small $H+$ networks with highly-favorable compositions can thus be more valuable to consumers than significantly larger $H0$ or $H-$ networks, allowing for niche, boutique, and exclusive platforms to both survive and thrive.

Proposition 5. *The elasticity of demand for platform j is decreasing in both* $\tilde{q}^*_j$ *(ceteris paribus) and* $\tilde{q}^*_j - \tilde{q}^*_{-j}$.

When delineating the market by type and increasing the value of composition for a given θ, consumers' responsiveness to prices diminishes, implying a less elastic demand.

Such competition for composition opens up an avenue for product differentiation over the network externality. This avenue is the primary mode of competition observed in online dating markets as features become standardized across platforms.[27] Dating platforms such as

[26] In a dynamic setting, the current state will influence future user behavior.

[27] Most dating websites employ similar algorithms built upon a singular value decomposition (Klemens 2006, p. 61).

ChristianMingle, JDate, and SilverSingles each target a distinct trait over which to differentiate, creating natural market segmentation. Given that these platforms no longer directly compete with one-another, the platforms have greater flexibility in pricing and can thus generate greater surpluses through higher pricing. In practice, competition for the market often emerges within each segment, as multiple politics-based, religious-based, race-based, and age-based dating platforms have emerged.

4.2 Cultivation Stage

To isolate the effects of cultivation, I make the following assumption regarding expectations.

Assumption 4. *For each type ℓ, $q^e_{A\ell} \sim U(0, n_\ell)$. Let $\mathbf{Q}^e$ denote the random variable.*[28]

Each draw is stochastic, with a symmetric average. Without cultivation, each platform has an expected $n/2$ users, with the types split evenly across platforms (perfect symmetry).

Remark 4. *The comparative static of pricing with respect to composition is*

$$\left[\underbrace{\frac{\partial p_j(\mathbf{q}^e(\mathbf{c}))}{\partial \tilde{q}^e_{j0}(\mathbf{c})}}_{\textbf{direct composition effect}} + \underbrace{\frac{\partial p_j(\mathbf{q}^e(\mathbf{c}))}{\partial q^e_{j\ell}(\mathbf{c})}\frac{\partial q^e_{j\ell}(\mathbf{c})}{\partial \tilde{q}^e_{j0}(\mathbf{c})}}_{\textbf{indirect composition effect}}\right]\frac{\partial \tilde{q}^e_{j0}(\mathbf{c})}{\partial c_{jy_r}}. \tag{13}$$

This decomposition is analogous to the pricing stage. It uncouples the direct effect of cultivation on prices through composition and the indirect effect of cultivation on prices through size.

Given $p_A(\mathbf{q}^e(\mathbf{c}))$ and $p_B(\mathbf{q}^e(\mathbf{c}))$, the platforms' objectives are

$$\max_{\mathbf{c}_A} \left\{p_A(\mathbf{q}^e(\mathbf{c}))[1 - \Phi(\omega^*_0)] - \kappa(\mathbf{c}_A)\right\} \tag{14}$$

$$\max_{\mathbf{c}_B} \left\{p_B(\mathbf{q}^e(\mathbf{c}))\Phi(\omega^*_0) - \kappa(\mathbf{c}_B)\right\}, \tag{15}$$

with ω^*_0 evaluated at $\mathbf{p}(\mathbf{q}^e(\mathbf{c}))$. The arbitrary first-order necessary condition (FOC) with

[28] Any truncated marginal distribution centered about $n_\ell/2$ is sufficient, I use the uniform distribution for simplicity.

respect to cultivating trait y_r by platform A is given by

$$p_A(\mathbf{q}^e(\mathbf{c}^*))\phi(\omega_0^*)\sum_{\ell:y_\tau\in\ell}\left[\frac{\partial v(q_A^e,\theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e}+\frac{\partial v(q_B^e,\theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e}\right.$$
$$\left.+\theta\left(\frac{\partial v(q_A^e,\theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e}+\frac{\partial v(q_B^e,\theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e}\right)h_{0\ell}\right]\frac{\partial f(q_{A\ell},\mathbf{c}^*)}{\partial c_{Ay_r}}=\frac{\partial\kappa(\mathbf{c}_A^*)}{\partial c_{Ay_r}},$$

where

$$\frac{\partial f(q_{j\ell},\mathbf{c}^*)}{\partial c_{jy_r}}=\begin{cases}\frac{\alpha_\tau}{\sum_{y_s\in\ell}\alpha_s} & \text{if } q_{A\ell}^e(\mathbf{c}^*)=q_{j\ell}^e+\frac{\sum_{y_s\in\ell}\alpha_s(c_{jy_s}-c_{-jy_s})}{\sum_{y_s\in\ell}\alpha_s}\\ 0 & \text{otherwise.}\end{cases}$$

Although not the true comparative static, the two effects are visible, augmented by the change in cultivation $[f(q_{j\ell},\mathbf{c})]$.

Lemma 3. *There exists at least one SPE.*

While existence is guaranteed, uniqueness is not. The SPE cannot be completely characterized without placing further restrictions on $\mathbf{q}^e$, H, and θ; however, analyzing the above FOC still provides insights.

Proposition 6. *In all SPE* $(\mathbf{c}^*,\mathbf{p}(\mathbf{q}^e(\mathbf{c})))$ *and for all* H *and* $\mathbf{q}^e$*:*

(i) *There exists a* $\underline{\theta}(\mathbf{q}^e)$ *such that* $\mathbf{c}_j^*\neq\mathbf{0}$ *for all* j *whenever* $\theta\leq\underline{\theta}(\mathbf{q}^e)$*.*

(ii) $c_{jy_r}^*>0$ *if and only if*

$$\sum_{\ell:y_\tau\in\ell}\left[\frac{\partial v(q_A^e,\theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e}+\frac{\partial v(q_B^e,\theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e}+\theta\left(\frac{\partial v(q_A^e,\theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e}+\frac{\partial v(q_B^e,\theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e}\right)h_{0\ell}\right]\frac{\partial f(q_{A\ell},\mathbf{c}^*)}{\partial c_{Ay_r}}>0.$$

(iii) *Suppose that* $\{\ell:y_r\in\ell\}=\{\ell:y_s\in\ell\}$ *and* $c_{jy_r}^*,c_{jy_s}^*>0$ *for some traits* y_r,y_s*. Then* $c_{jy_r}>(=)\,c_{jy_s}$ *if and only if* $\alpha_r>(=)\,\alpha_s$*.*

When θ is small, the rewards to efforts aimed at increasing the size of the installed base are significantly more than efforts aimed at improving the composition. As both platforms are able to employ strategic cultivation, a prisoner's dilemma emerges, much like the typical advertising game. The platforms' efforts counteract one-another.

Trait y_r is cultivated whenever the total marginal effect of cultivating that trait is positive. The total effect, given in (ii), considers every type ℓ in which trait y_r is present. The size effect (appearing in the term) is positive for each ℓ, though the composition effect (also appearing) can take any sign. The net effect is the sum of the two. Summing the net effect across all ℓ affected by trait y_r gives the total marginal effect, which when positive implies that trait y_r is revenue-increasing. When negative, it is revenue-decreasing and when zero, revenue-neutral. In the latter two cases, not cultivating trait y_r is a best response.

For any positive θ, those traits with a greater α carry a greater weight according to the marginal effect of cultivation on expectations, $\partial f(q_{j\ell}, \mathbf{c}^*)/\partial c_{jy_r}$, which is nondecreasing in that α. If there are two traits that affect the same set of types, then the more dominant trait, defined by the larger α, is cultivated more intensely.

Definition 3. *Trait y_r (weakly) dominates y_s if $\alpha_r >\ (\geq)\ \alpha_s$ Trait y_r is (weakly) dominant if $\alpha_r >\ (\geq)\ \max_{y_s \neq y_r}\{\alpha_s\}$.*

Dominance is unclear across two traits that affect non-identical sets of types, even if those sets have a non-empty intersection.

Regardless of the size effect, specific traits may be cultivated so long as they are revenue-increasing. Thus cultivation always occurs when the size effect is dominant, as any increase in the size of the installed base is valuable. This outcome of Proposition 6 resembles the traditional literature, although the strategy pursued is distinct. While useful, it is more interesting to consider investments increasing the installed base when not all increases are treated equally. Some increases can even harm the platform. Unfortunately, precise characterizations of the SPE are unavailable without placing further restrictions on H. For the remainder of this section, suppose that θ is sufficiently large so that the composition effect dominates the size effect. In what follows, I consider several specifications of H that resemble real-world occurrences, such as in-group bias, out-group bias, and grouping.

4.2.1 In-Group Bias

In-group bias occurs whenever an individual prefers an own-type to an other-type. This broad definition implies that $h_{\ell\ell} > h_{\ell m} > 0$ for $\ell \neq m$ constitutes in-group bias. I use a more restrictive definition where other-types impose a negative externality, rather than a smaller positive externality to eliminate uninteresting cases.[29]

[29] All positively valued $h_{\ell m}$ implies no negative externality, making cultivation is a dominant strategy.

Definition 4. *In-group bias exists if* $h_{\ell\ell} > 0$ *for each* ℓ *and* $h_{\ell m} < 0$ *for all* $m \neq \ell$. *The in-group bias is weak if* $\sum_{m \neq \ell} |h_{\ell m}| > h_{\ell\ell}$ *for all* ℓ. *The in-group bias is strong if* $\sum_{m \neq \ell} |h_{\ell m}| < h_{\ell\ell}$ *for all* ℓ.

The strength of in-group bias is defined by the diagonal-dominance of H. If H is diagonal dominant, then the bias is strong and if H is diagonal non-dominant, then it is weak. Strong bias implies that the presence of a single own-type outweighs the presence of one of each of the $\overline{Y} - 1$ other types for each type. Weak bias implies the opposite. If the in-group bias is strong, then $\sum_{m \neq 0} |h_{0m}| > h_{00}$. An analogous relationship holds for weak in-group bias.

Proposition 7. *Suppose that* $\mathbf{q}^e = \mathbb{E}\mathbf{Q}^e$ *and that* θ *is sufficiently large. If, for some nonnegative* ε, $\sum_{m \neq \ell} |h_{\ell m}| - h_{\ell\ell} > \varepsilon$, *then there is a unique SPE in which* $\mathbf{c}^* = \mathbf{0}$. *If* $\sum_{m \neq \ell} |h_{\ell m}| - h_{\ell\ell} < -\varepsilon$, *then in every SPE,*

(i) $\mathbf{c}^* \neq \mathbf{0}$,

(ii) $\mathbf{c}^*_{jy_r} > 0$ *for at least one* j *and all* y_r.

If $\partial^2 v(q^e_j(\mathbf{c}), \tilde{q}^e_{j\ell})/\partial q^{e2}_{j\ell} = 0$, then $\varepsilon = 0$. If $\partial^2 v(q^e_j(\mathbf{c}), \tilde{q}^e_{j\ell})/\partial q^{e2}_{j\ell} > 0$, then $\varepsilon > 0$, and ε is increasing in the steepness of $\partial v(q^e_j(\mathbf{c}), \tilde{q}^e_{j\ell})/\partial q^e_{j\ell}$. If $\partial^2 v(q^e_j(\mathbf{c}), \tilde{q}^e_{j\ell})/\partial q^{e2}_{j\ell} < 0$, then weak in-group bias is not necessary for Proposition 7, though it is sufficient. When the in-group bias is sufficiently strong, the additional positive externality of increasing the presence of in-group outweighs the negative externality induced on existing out-groups. As a result, there is always the incentive to cultivate at least one trait. Moreover, increasing the presence of each type of user is also value-enhancing [vis-à-vis $v(q^e_j(\mathbf{c}), \cdot)$], so all traits are cultivated, though not necessarily by the same platform. When the in-group bias is sufficiently weak, the additional negative externality outweighs any positive externality generated by size effects and positive externalities from a larger in-group. In this environment, the ideal is a wholly in-group composition, though this composition is only obtainable under a small subset of conditions on $\mathbf{q}^e$. Under the assumption $\mathbf{q}^e = \mathbb{E}\mathbf{Q}^e$, $\tilde{q}^e_{j\ell} = [n/(2\overline{Y})] \sum_{m=1}^{\overline{Y}} h_{\ell m}$. From the perspective of each ℓ user, the platform is $H+$ if the in-group bias is weak, $H-$ if it is strong, and $H0$ if it is balanced.

Under asymmetric expectations, a wide variety of outcomes are possible. When the distribution of a specific type across platforms is significantly unbalanced and H exhibits weak in-group bias, an SPE without cultivation may no longer exist. The platform with the large share of type-ℓ users seeks to increase their presence. Although the in-group bias is weak, the number of individuals benefitting from the increase of the in-group outweighs the negative externality imposed on those individuals who suffer from the presence of an out-group member. Whether or not the SPE is effected depends on how weak the in-group bias is, i.e. the magnitude of $\sum_{m \neq \ell} |h_{\ell m}| - h_{\ell\ell}$.

Proposition 8. *Suppose θ is sufficiently large. For every asymmetric $\mathbf{q}^e$, there exists a positive ε' such that an SPE without cultivation exists if and only if $\sum_{m \neq \ell} |h_{\ell m}| - h_{\ell\ell} > \varepsilon'$.*

When $\sum_{m \neq \ell} |h_{\ell m}| - h_{\ell\ell} > \varepsilon'$, the aggregate negative externality outweighs the aggregate positive externality.

4.2.2 Out-Group Bias

Like in-group bias, I use a stronger than necessary definition of out-group bias: $h_{\ell\ell} < 0$ for each ℓ and $h_{\ell m} > 0$ for every $m \neq \ell$.

Definition 5. *The out-group bias is weak if $\sum_{m \neq \ell} h_{\ell m} > |h_{\ell\ell}|$ for all ℓ. The out-group bias is strong if $\sum_{m \neq \ell} h_{\ell m} < |h_{\ell\ell}|$ for all ℓ.*

When out-group bias is present, a tension between types emerges. Suppose there are two types of users, a minority type and a majority type. The minority is satisfied due to the presence of a large out group, while the majority is dissatisfied. Which party's concerns are more important to the platform, vis-à-vis profits, depends on the strength of the bias. If a type dominates others according to the magnitude of $h_{\ell\ell}$, then its satisfaction is weighted heavier. Under strong out-group bias, the goal is to shrink the presence of the most dominant types, saturating the platform with the various out-groups. Because this minority has the most intense preferences for the out-group, their satisfaction outweighs the dissatisfaction of the remaining types. Since the platform is not a monopoly, its competitor has a similar strategy in mind.

Proposition 9. *Suppose that θ is sufficiently large, $|h_{11}| > |h_{22}| > \cdots > |h_{\overline{Y}\overline{Y}}|$, and $h_{m\ell} = h'$ for all $m \neq \ell$ and some value $h' > 0$. If the out-group bias is sufficiently strong, then platform j's profits are maximized when, for a given $\mathbf{q}^e$, $q_{j1} < q_{j2} < \cdots < q_{j\overline{Y}}$.*

In the SPE, too much cultivation occurs, again resembling a prisoner's dilemma. Suppose that $q^e_{j\ell} \approx n_\ell/2$ for each ℓ. If a platform does not cultivate, then it's competitor has the incentive to cultivate, targeting $h_{\overline{YY}}$ the heaviest, moving downward sequentially, turning the most dominant types into minorities on the platform. Thus there is no SPE without cultivation. The above profile cannot be part of an SPE either. The non-cultivating platform can increase its value by targeting the same group, increasing their numbers, which increases value to the more dominant types by decreasing their relative share in the installed base. In the end $\mathbf{q}^e \approx \mathbf{q}^e(\mathbf{c}^*)$ with $\mathbf{c}^* \neq \mathbf{0}$: a prisoner's dilemma. An analogous argument holds with an asymmetric c^* when the initial distribution is asymmetric.

4.2.3 Grouping

Grouping is a generalized form of in-group bias. Akin to in-group bias, $h_{\ell\ell} \geq 0$ for each ℓ. Partition the types into K sets, $L_1, L_2, \ldots, L_K$. For each $k = 1, \ldots, K$, all $\ell, m \in L_k$ and all $m' \notin L_k$, $h_{\ell m} > 0$ and $h_{\ell m'} < 0$. H takes on a block-diagonal shape

$$\begin{pmatrix} + & + & - & - & - & - \\ + & + & - & - & - & - \\ - & - & + & + & - & - \\ - & - & + & + & - & - \\ - & - & - & - & + & + \\ - & - & - & - & + & + \end{pmatrix}.$$

Unlike in-group bias, the diagonals ($h_{\ell\ell}$) need not be positive. The positive and negative values can each have varying weights ($h_{\ell m}$ need not be equal). Such patterns are often found in online dating markets, where strong preferences over race, religion, education, and others are held.

The platform's ideal composition depends on the θ. For larger θ, each platform will target specific positive groupings, hoping that its competitor will target the alternative group. When $K = 2$, the platforms will cultivate traits corresponding to members of a group as best as possible (as there are overlaps when $y_r \in \ell$ and $y_r \in m$ for at least one y_r). If the two groups are somewhat even in size and value, then there is no tension. An SPE emerges in which each platform targets a different group. Which group is targeted by each platform depends on the initial state $\mathbf{q}^e$. If there is a large disparity between the size or value of each group, a prisoner's dilemma emerges and the platforms compete for the same group.

Definition 6. *A group k is dominant if*

(i) $\min\{h_{m\ell} :\ m, \ell \in L_k\} - \max_{k' \neq k} \max\{h_{m\ell} :\ m, \ell \in L_{k'}\} \gg 0$

(ii) $|\max\{h_{m\ell} :\ m \in L_k,\ \ell \notin L_k\}| - |\min_{k' \neq k} \min\{h_{m\ell} :\ m \in L_{k'},\ \ell \in L_k\}| \gg 0.$

Two groups are similar if, for groups k and k',

(iii) $|h_{m\ell}| - |h_{m'\ell'}| \approx 0$ *for all* $m, \ell \in L_k,\ m', \ell' \in L_{k'}$

(iv) $|h_{m\ell'}| - |h_{m'\ell}| \approx 0$ *for all* $m, \ell \in L_k,\ m', \ell' \in L_{k'}$.

Statement (i) states that the type with the lowest within-group valuation in group k is sufficiently larger than the type with the highest within-group valuation in all other groups. Statement (ii) states that the type most tolerant of the out-group is significantly less tolerant than the least tolerant type in the all of the out-groups. Statements (iii) and (iv) are straightforward.

Proposition 10. *Suppose that $K = 2$, $y_r \in \ell \Rightarrow y_r \notin m\ \forall m \neq \ell$ (types and traits are one-to-one), $\mathbf{q}^e = \mathbb{E}\mathbf{Q}^e$ ($q^e_{j\ell} = n_l/2$ for all j, ℓ), and θ is sufficiently large.*

(i) *If the two groups are similar, then there exists a unique SPE $(\mathbf{c}^*, \mathbf{p}(\mathbf{q}^e(\mathbf{c})))$ in which*

  (a) $c^*_{jy_r} > 0$ *if and only if* $c^*_{-jy_r} = 0$,

  (b) $c^*_{jy_r} > 0$ *for at least one j and all y_r,*

  (c) *If $c^*_{jy_r}$ and $c^*_{jy_s} > 0$ for all $y_r \neq y_s$, then $h_{y_r y_s} > 0$.*

(ii) *If group 1 is dominant, then there exists a SPE $(\mathbf{c}^*, \mathbf{p}(\mathbf{q}^e(\mathbf{c})))$ in which*

  (a) $c^*_{jy_r} = c^*_{-jy_r}$ *for all y_r,*

  (b) $c^*_{jy_r} > 0$ *for all j and $y_r \in L_1$,*

  (c) $c^*_{jy_s} > 0$ *for all j and all y_s such that $h_{y_r y_s} > 0$ for all $y_r \in L_1$,*

  (d) $c^*_{jy_t} = 0$ *for all j and all y_t such that $h_{y_r y_t} < 0$ for all $y_r \in L_1$.*

Case (i) formalizes an equilibrium with delineation of the market by group. Each platform cultivates a different group and profits are strictly greater than under a zero-cultivation profile. Case (ii) formalizes the prisoner's dilemma.

When $K > 2$, the delineation of the user base by type is less precise. Without relaxing assumptions on ζ_A and ζ_B eliminating market coverage, cultivation is not as effective.[30] Each platform is forced to have both in-and-out groups.

So far, cultivation alters expectations by increasing the presence of types possessing cultivated traits. Often, other traits and types are affected. I refer to these effects as enhancing or refining.

4.2.4 Enhancing and Refining Effects

Cultivation often has a refining effect: targeting one set of traits negatively affects individuals with related traits, pushing them away from the platform. For example, cultivating a specific age, education, racial, or religious group in online dating. If a dating platform announces itself as a platform for Jewish singles to find other Jewish singles, both Jewish and non-Jewish singles will adjust their expectations accordingly. Not expecting to find a match, non-Jewish singles adjust their expectations downward. When traits are complementary, targeting one set of traits can enhance cultivation, attracting those with complementary traits. For example, a pet-friendly dating platform cultivating pet owners will also attract those without pets who like pets. This pattern represents an enhancing effect, where attracting one type attracts others.

Applying enhancing and refining effects to cultivation increases its efficacy. In the case of refinement, the size of the non-cultivated installed base shrinks, but this shrinkage increases the value of the composition effect. For large θ, this shrinkage is a net positive for the platform. A platform targeting a small coalition with strong in-group bias benefits from refinement. It shrinks its installed base so that in the limit only members of the coalition join the platform. This coalition is not valuable in the vertical sense; that is, they do not possess a higher income level or a higher willingness to pay for the platform in general. Rather, the value arises horizontally, through homogeneity. The platform's exclusivity is driving its value. It is not exclusive in the typical sense of niche or luxury goods, but exclusive to a specific group with preferences to remain isolated. Cultivation acts as a public good, where each platform prefers the other to cultivate and through refinement, receive a share of the benefits. After a war of attrition, one platform acquiesces and cultivates.

At face value, the market resembles the dominant firm-fringe firm paradigm; however the fringe platform is not "fringe" in the typical sense. Rather, the fringe (cultivating) platform

[30] Without market coverage, platforms will price and cultivate such that only one group joins each platform under case (i). Under case (ii), both of the platforms will cultivate and split the dominant group.

is strategically small by its own doing and the non-cultivating platform's "dominant" position by way of market share is only so as the rational response to its cultivating counterpart. It is driven not by the larger platform, but the smaller platform. Both platforms may be more profitable than they would be if they split the market evenly. The cultivating platform has fewer but more valuable adopters while the larger platform is compensated by the increasing returns generated by more users.

The dating website JDate, cultivating on the characteristic of religion (Judaism), faces such a trade-off. Jewish individuals make up approximately 2% of the adult population in the United States.[31]. Proposition 10(i) shows how it can be highly profitable to consolidate the network and leverage composition effects into prices matching its much larger competitors. JDate's single month price is \$39.99, while Match charges \$34.99, where JDate has only a fraction of the installed base of Match. Similarly, AttractiveWorld charges a significantly higher price than Match, \$76.81 vs. \$34.99.

5 Composition: Collusion and Competition

The model and analysis are also useful in understanding environments outside of the direct scope of the model. I emphasize two such environments. First, identifying collusion via price fixing can be difficult when the goods in question possess network externalities. This issue is exacerbated by preferences over composition. Ignoring composition effects leads to false positives by attributing higher-than-expected (or lower-than-expected) prices to factors outside of marketplace competition. Second, it is difficult to explain a multiproduct firm in a network industry when the products are similar. Without regards to composition, segmenting users into multiple platforms cannibalizes demand. By incorporating composition effects, such separation can be demand enhancing rather than cannibalizing.

5.1 Identifying Collusion

If composition is ignored and consumers are cultivated so that size remains constant but the composition is improved, then equilibrium prices will appear to be higher than what competition supports. To formally illustrate this point, suppose that: (i) $y_r \in \ell \Rightarrow y_r \notin m \ \forall m \neq \ell$, (ii) $n_\ell = n/\overline{Y}$ (equal number of users of each type), (iii) $q^e_{j\ell} = n_\ell/2$ for each ℓ and j (initial expectations divide each type equally among the platforms), (iv) $\theta \gg 0$

[31] http://www.pewresearch.org/fact-tank/2013/10/02/how-many-jews-are-there-in-the-united-states/, accessed March 25, 2018.

(composition matters). For simplicity, I also assume that $|h_{\ell m}|$ is constant for all ℓ, m and equal to some value h^*. H takes the block diagonal shape

$$H = \begin{pmatrix} \mathbf{h}^* & -\mathbf{h}^* \\ -\mathbf{h}^* & \mathbf{h}^* \end{pmatrix}, \quad \text{where } \mathbf{h}^* = \begin{pmatrix} h^* & \dots & h^* \\ \vdots & \ddots & \vdots \\ h^* & \dots & h^* \end{pmatrix}.$$

Using the analysis in Section 4, an SPE $(\mathbf{c}^*, \mathbf{p}(\mathbf{q}^e(\mathbf{c}))$ exists in which: (i) all types are cultivated, but no two are targeted by the same platform, (ii) there is symmetric cultivation, (ii) only in-group types are cultivated, and (iv) the platforms have symmetric sizes. In this SPE, $p_A(\mathbf{q}^e(\mathbf{c}^*)) = p_B(\mathbf{q}^e(\mathbf{c}^*))$. Setting $\theta = 0$ and $\mathbf{c} = \mathbf{0}$ replicates a model without composition effects. The outcome is prices $p'_A = p'_B$ such that $p'_j < p(\mathbf{q}^e(\mathbf{c}^*))$. Thus if prices $\mathbf{p}(\mathbf{q}^e(\mathbf{c}^*))$ are observed but a model without composition effects predicts prices $\mathbf{p}'$, then the platforms face unwarranted scrutiny for collusive behavior.

Prices can also be lower than what would be predicted by a model devoid of composition effects. If the platforms are $H-$, then each platform sets prices lower than if they were $H0$ or $H+$. Combining these lower-than-expected prices with high startup or fixed costs may lead to false positives in the other direction. Estimating prices without accounting for composition effects, but observing prices correspond to the non-cooperative equilibrium leads to predicting predatory pricing when no such pricing is occurring.

5.2 Multiplatform Monopolist

Christian Mingle, JDate, Silver Singles, Attractive world, and others are all owned by Spark Networks. Match, SingleParentMeet, SeniorPeopleMeet, and others are all owned by the Match Group Inc. A single entity owning and operating multiple platforms is not unique to the online dating industry.[32] Nonetheless, these examples illustrate the importance of accounting for composition effects when considering network goods and uncovers a phenomenon that is difficult without incorporating composition effects: multiproduct firms in network industries.[33]

Composition effects and strong preferences over types are sufficient to explain why users may prefer multiple independent platforms over a single unified platform. By separating users according to types, each smaller platform becomes increasingly valuable to its base. The gains from the increased positive composition effect outweighs loss from the decreased size

[32] For example, PayPal and Venmo are under the same ownership.

[33] Mergers with significant switching costs can explain multiproduct firms, though they cannot explain multiple products originating from the same owner.

effect. These exclusive platforms are then priced higher than the original larger platform, strictly increasing profits.

5.2.1 Formal Analysis

Instead of two platforms competing, suppose that there is a single owner that operates up to $J \geq 1$ platforms. Also suppose that ξ_A is constant across all users, i.e. $\phi(\xi_A)$ is a single-point-mass. This simplifies the analysis so that when operating a single platform, profit maximization entails capturing all n users when there are no composition effects ($\theta = 0$ or $H = \mathbf{0}_{\overline{Y}\times\overline{Y}}$). Cultivation is still a potentially profitable strategy, though any SPE from Section 4 in which a prisoner's dilemma emerges is immediately eliminated with a multiproduct monopolist. The remainder of the model is as described in Section 3.

Proposition 11. *For θ sufficiently small, the owner maximizes profits by operating a single platform.*

For simplicity, suppose $\theta = 0$. In this case, dividing a single large platform into smaller, independent platforms lowers $v(n, 0)$ to $v(q_j^e, 0)$ for all users. When there is a single platform, the profit maximizing price is given by $p^* = \xi_A + v(n, 0)$, with profits np^*. When there are $J > 1$ platforms, the price of each platform j is $p_j^*(0) = \xi_A + v(q_j^e, 0) < p^*$. The inequality follows from the fact that $n > q_j^e$ when $J > 1$. Profits are thus $\sum_j n_j p_j^* < np^*(0)$ as $\sum_j n_j = n$.

As θ grows larger, a more profitable opportunity arises when consumers hold specific preferences over the composition.

Proposition 12. *For θ sufficiently large and either in-group bias, out-group bias, or grouping, the owner maximizes profits by operating at least two platforms.*

More generally, any H that exhibits in-group bias, out-group bias or grouping on a subset of H satisfies the condition. The most common segmentation in online dating occurs via religious traits. More recently, segmentation has occurred via political beliefs. Under in-group bias, each $h_{\ell\ell} > 0$ and $h_{\ell m} < 0$ for ever $\ell \neq m$. The platform maximizes profits by setting $J = \overline{Y}$, where the price of each platform j is $p_j^*(\theta) = \xi_A + v(q_j^e, \theta\tilde{q}_{j\ell}^e)$. For θ sufficiently large, $p_j^*(\theta) > p^*$. Thus $\sum_j p_j^*(\theta) n_j > p^* n$.

The above analysis also uncovers why search filters in online dating sites do not represent a substitute solution to multiple platforms. These search algorithms decrease the cost of matching (increasing the value of the network effects); however, the valuation of the platform also depends on the presence of individuals for which there may not be a direct interaction. A man in his early thirties with children faces competition in the dating market from other men, both of the same and different types, e.g. a similar man without children. Similarly, women face competition from other women of the same and different types. These search algorithms cannot prevent an individual's desired type from interacting with the "competition." Separation of the platforms does mitigate the issue by eliminating the presence of other types. As religion is among the most dominant traits, companies have developed multiple platforms, segmenting their user base by religion. Similarly, political affiliations have recently grown in importance, more platforms are emerging that cultivate according to political leanings, including pro- and anti-Trump dating platforms and pro- and anti-Brexit dating platforms.[34] This idea extends well beyond online dating markets to platforms more generally.

6 Multisided Platforms and Dynamic Considerations

Sections 3 and 4 studied one-sided platforms in a static setting. Although the results (and proofs) are in this context, some of the insights can still be applied to more dynamic settings as well as multisided platforms (e.g. advertisers and end-users on a social network). In what follows, I provide some intuition on some of these extensions, with a focus on how composition and cultivation affects these environments. Complete analyses of these extensions are left for future work.

6.1 Multisided Platforms

Many modern platforms, including mobile-phone based dating markets, are multisided. End-users do not pay for the service. They act as loss-leaders while advertisers pay to reach users through the platform. It merits questioning if advertisers ever prefer, as defined by willingness to pay for access, a smaller platform.[35] Advertisers of niche, status, or luxury goods want to prevent over-saturation; however, these goods are often endowed with negative network externalities, where the value decreases as consumption permeates the economy. An

[34]https://www.forbes.com/sites/janetwburns/2016/10/07/8-political-dating-apps-to-help-you-escape-trump-or-bond-over-brexit/#7e11d2dd774a, accessed August 8, 2019.

[35]See Chandra & Collard-Wexler (2009) for a study of advertiser behavior under mergers.

argument exists for advertisers selling more goods without network externalities to prefer smaller platforms with a more appropriate composition.

Suppose a product targets a specific demographic. There are two platforms: one with 10,000 users, all of whom from the targeted demographic and one with 20,000 users, 5,000 of whom from the targeted demographic. In this case, the advertiser is willing to pay more for the smaller platform. Yet, if the second platform is augmented so that there are still 20,000 total users, with 11,000 from the targeted demographic, then the advertiser may still be willing to pay more for the smaller platform.

On the larger platform, advertisers compete for user attention with other advertisers. This competition includes advertisers targeting the specific demographic, advertisers targeting other demographics, and those targeting the general population. As a result, the likelihood of reaching the desired users is decreased, particularly by those advertisers targeting the general population.[36] Given a lower expected effectiveness, a lower willingness to pay follows.

In the context of social networks, advertisers face a similar issue, even without the presence of other advertisers. There is competition for views from other users. On larger platforms with many posts, the advertisements can become buried and suffer from similar clutter effects described in Ha & McCann (2008). The result is a higher willingness to pay for the smaller platform with the greater likelihood of exposure to the demographic of interest.

6.2 Dynamic Composition Effects

Section 4 illustrated how prices and profits are effected by composition effects, and the valuation of composition θ. Platforms are able to leverage composition through cultivation to increase prices and profits. As θ increases, so too do the benefits of cultivation. This relationship introduces a potential dynamic profitable opportunity for platforms: influencing the value of composition. If a platform can both cultivate a valuable installed base and increase how consumers value composition via θ, then over time, prices and profits rise.

Strategies similar to influencing θ have been utilized by identity groups, such as religious organizations. Carvalho & Sacks (2019) studies the dynamics of niche construction, where leaders of identity groups leverage the environment to increase payoffs. Platforms can employ similar strategies. Early investments in cultivation and increasing θ can yield significant long run dividends. Moreover, if there is significant lock-in, then the long run returns are even greater. Platforms are therefore willing to expend significant resources to increase θ.

[36]Evidence in support of this claim is found in Cho & Cheon (2004) and Ha & McCann (2008).

If θ can be increased by platform activities, then it is similarly plausible that a lack of investment can lead θ to decay. As θ shrinks, composition becomes less important and platforms will invest less in cultivating a composition. In the long run, the dynamics converge to those described in Cabral (2011).

Another margin through which platform developers can increase value is through ζ_j. Increasing the direct value of the platform relative to the outside option, all else equal, increases consumers' willingness to pay regardless of composition. Dating platforms such as Match employ this strategy. Match hosts offline meet-ups, where members meet face-to-face doing activities such as bowling and karaoke.[37]. These interactions alter the likelihood of interacting with other users off of the platform, which through social reinforcement can increase the value of the platform.[38]

7 Concluding Remarks

Prior to this article, composition effects have been largely absent from the study of network goods and platforms, and relatively absent from the general literature on markets with consumption externalities.[39] I have shown that many of the equilibrium and comparative static and dynamic properties of price competition in the presence of network externalities need not hold.

In particular, a platform's price is no longer necessarily monotonic in its market share, which implies that a platform's market share is no longer a sufficient statistic for its success in terms of profitability, as this correspondence is no longer one-to-one. The proper comparative static must incorporate the direct size effects and the indirect composition effects. I develop a model that analytically decomposes the comparative static to isolate these effects. The importance of composition effects opens up a new area of study within the realm of imperfect competition: the cultivation of heterogeneity and product differentiation over the externality. I have shown that platforms can use cultivation to coordinate expectations and endogenously influence the

[37]http://blog.match.com/category/stir-events/, accessed March 27, 2018

[38]For example, Nintendo and other video game developers hold tournaments, which influence social interactions.

[39]Exceptions include the study of local public goods and some work on club goods. In the provision of local public goods, heterogeneity among consumers, when coupled with preferences over heterogeneity, has substantial effects on outcomes. For instance, individuals' willingness to provide local public goods through taxation decreases as the degree of ethnic fragmentation (heterogeneity) in the local population increases (Easterly & Levine 1997, Alesina et al. 1999, Alesina & La Ferrara 2000). This finding indicates that individual attachment to goods with consumption externalities is directly affected by the composition of its users.

valuations in both the population of the platform's adopters and the population at large. The effects of cultivation can be similarly studied by decomposing the pricing effects into a direct effect through composition and indirect effect through the installed base. Although I linked many of the assumptions and results to online dating platforms, the results presented are more general and can be applied to many networks, particularly those found in the new economy.

References

Albert, J. & Chib, S. (1993), 'Bayesian analysis of binary and polychotomous response data', *Journal of the American Statistical Association* **88**(422), 669–679.

Alesina, A., Baqir, R. & Easterly, W. (1999), 'Public goods and ethnic divisions', *The Quarterly Journal of Economics* **114**(4), 1243–1284.

Alesina, A. & La Ferrara, E. (2000), 'Participation in heterogeneous communities', *The Quarterly Journal of Economics* **115**(3), 847–904.

Algesheimer, R., Dholakia, U. & Herrmann, A. (2005), 'The social influence of brand community: Evidence from european car clubs', *Journal of Marketing* **69**(3), 19–34.

Ambrus, A. (2006), 'Coalitional rationalizability', *The Quarterly Journal of Economics* **121**(3), 903–929.

Ambrus, A. & Argenziano, R. (2009), 'Asymmetric networks in two-sided markets', *American Economic Journal: Microeconomics* **1**(1), 17–52.

An, M. & Kiefer, N. (1995), 'Local externalities and societal adoption of technologies', *Journal of Evolutionary Economics* **5**(2), 103–117.

Athey, S. & Ellison, G. (2011), 'Position auctions with consumer search', *The Quarterly Journal of Economics* **126**(3), 1213–1270.

Basu, K. (1989), 'A theory of association: Social status, prices and markets', *Oxford Economic Papers* **49**(4), 653–671.

Baylis, K. & Perloff, J. (2002), 'Price dispersion on the internet: Good firms and bad firms', *Review of Industrial Organization* **21**(3), 305–324.

Berman, E. (2000), 'Sect, subsidy, and sacrifice: An economist's view of ultra-orthodox jews', *The Quarterly Journal of Economics* **115**(3), 905–953.

Birke, D. (2009), 'The economics of networks: A survey of the empirical literature', *Journal of Economic Surveys* **23**(4), 762–793.

Brynjolfsson, E. & Kemerer, C. (1996), 'Network externalities in microcomputer software: An econometric analysis of the spreadsheet market', *Management Science* **42**(12), 1627–1647.

Burdett, K. & Judd, K. (1983), 'Equilibrium price dispersion', *Econometrica* **51**(4), 955–969.

Cabral, L. (2011), 'Price competition with network effects', *Review of Economic Studies* **78**(1), 83–111.

Carvalho, J.-P., Koyama, M. & Sacks, M. (2017), 'Education, identity, and community: lessons from jewish emancipation', *Public Choice* **171**(1-2), 119–143.

Carvalho, J.-P. & Sacks, M. (2019), 'Radicalization', *Working Paper* .

Chandra, A. & Collard-Wexler, A. (2009), 'Mergers in two-sided markets: An application to the Canadian newspaper industry', *Journal of Economics and Management Strategy* **18**(4), 1045–1070.

Chen, J., Doraszelski, U. & Harrington, J. (2009), 'Avoiding market dominance: product compatibility in markets with network effects', *RAND Journal of Economics* **40**(3), 455–485.

Chen, J. & Sacks, M. (2018), 'Reimbursing consumers' switching costs in network industries', *NET Institute Working Paper No. 16-13* .

Cho, C.-H. & Cheon, H. (2004), 'Why do people avoide advertising on the internet?', *Journal of Advertising* **33**(4), 89–97.

Damiano, E. & Li, H. (2007), 'Price discrimination and efficient matching', *Economic Theory* **30**(2), 243–263.

de Palma, A. & Leruth, L. (1996), 'Variable willingness to pay for network externalities with strategic standardization decisions', *European Journal of Political Economy* **12**(2), 235–251.

Doganoglu, T. (2003), 'Dynamic price competition with consumption externalities', *Netnomics* **5**(1), 43–69.

Easterly, W. & Levine, R. (1997), 'Africa's growth tragedy: Policies and ethnic divisions', *The Quarterly Journal of Economics* **112**(4), 1203–1250.

Farrell, J. & Klemperer, P. (2007), *Handbook of Industrial Organization*, Vol. 3, Elsevier B.V., chapter 31. Coordination and Lock-in: Competition with Switching Costs and Network Effects, pp. 1970–2056.

Farrell, J. & Saloner, G. (1985), 'Standardization, compatibility, and innovation', *RAND Journal of Economics* **16**(1), 70–83.

Farrell, J. & Saloner, G. (1986), 'Installed base and compatibility: Innovation, product preannouncements, and predation', *American Economic Review* **76**(5), 940–955.

Fisman, R. & Iyengar, S. (2008), 'Racial preferences in dating', *Review of Economic Studies* **75**(1), 117–132.

Fudenberg, D. & Tirole, J. (2000), 'Pricing a network good to deter entry', *Journal of Industrial Economics* **48**(4), 373–390.

Gandal, N. (1994), 'Hedonic price indexes for spreadsheets and an empirical test for network externalities', *RAND Journal of Economics* **25**(1), 160–170.

Gomes, R. & Pavan, A. (2011), 'Price discrimination in many-to-many matching markets', *Discussion Paper, Center for Mathematical Studies in Economics and Management Science, No. 1540* .

Goyal, S. (2009), *Connections: An Introduction to the Economics of Networks*, Princeton University Press, Princeton, New Jersey.

Ha, L. & McCann, K. (2008), 'An integrated model of advertising clutter in offline and online media', *International Journal of Advertising* **27**(4), 569–592.

Hagiu, A. & Spulber, D. (2013), 'First-party content and coordination in two-sided markets', *Management Science* **59**(4), 933–949.

Henkel, J. & Block, J. (2013), 'Peer influence in network markets: a theoretical and empirical analysis', *Journal of Evolutionary Economics* **23**(5), 925–953.

Hitch, G., Hortaçsu, A. & Ariely, D. (2010*a*), 'Matching and sorting in online dating', *American Economic Review* **100**(1), 130–163.

Hitch, G., Hortaçsu, A. & Ariely, D. (2010*b*), 'What makes you click? - mate preferences in online dating', *Quantitative Marketing and Economics* **8**(4), 393–427.

Iannaccone, L. R. (1992), 'Sacrifice and stigma: Reducing free-riding in cults, communities, and other collectives', *Journal of Political Economy* **100**(2), 271–291.

Iannaccone, L. R. (1994), 'Why strict churches are strong', *American Journal of Sociology* **99**(5), 1180–1211.

Jackson, M. (2010), *Social and Economic Networks*, Princeton University Press, Princeton, New Jersey.

Janssen, M. & Mendys-Kamphorst, E. (2007), 'Evolution of market shares with repeated purchases and heterogeneous network externalities', *Journal of Evolutionary Economics* **17**, 551–577.

Jeziorski, P. (2014), 'Effects of mergers in two-sided markets: The us radio industry', *American Economic Journal: Microeconomics* **6**(4), 35–73.

Katz, M. & Shapiro, C. (1985), 'Network externalities, competition, and compatibility', *American Economic Review* **75**(3), 424–440.

Klemens, B. (2006), *Math you can't use: patents, copyright, and software*, The Brookings Institution, Washington, D.C.

Laffont, J.-J., Rey, P. & Tirole, J. (1998*a*), 'Network competition: I. overview and nondiscriminatory pricing', *RAND Journal of Economics* **29**(1), 1–37.

Laffont, J.-J., Rey, P. & Tirole, J. (1998*b*), 'Network competition: Ii. price discrimination', *RAND Journal of Economics* **29**(1), 38–56.

Markovich, S. (2008), 'Snowball: A dynamic oligopoly model with indirect network effects', *Journal of Economic Dynamics and Control* **32**(3), 909–938.

Markovich, S. & Moenius, J. (2009), 'Winning while losing: competition dynamics in the presence of indirect network effects', *International Journal of Industrial Organization* **27**(3), 346–357.

Marx, P. & Schummer, J. (2019), 'Revenue from matching platforms', *Working Paper* .

McBride, M. (2008), 'Religious pluralism and religious participation: A game theoretic analysis', *American Journal of Sociology* **114**(1), 77–108.

McBride, M. (2015), 'Why churches need free-riders: religious capital formation and religious group survival', *Journal of Behavioral and Experimental Economics* **58**, 77–87.

Mitchell, M. & Skrzypacz, A. (2006), 'Network externalities and long-run market shares', *Economic Theory* **29**(3), 621–648.

Salop, S. & Stiglitz, J. (1977), 'Bargains and ripoffs: A model of monopolistically competitive price dispersion', *Review of Economic Studies* **44**(3), 493–510.

Shankar, V. & Bayus, B. (2003), 'Network effects and competition: an empirical analysis of the home video game industry', *Strategic Management Journal* **24**(4), 375–384.

Shy, O. (2001), *The Economics of Network Industries*, Cambridge: Cambridge University Press, New York, New York.

Shy, O. (2011), 'A short survey of network economics', *Review of Industrial Organization* **38**(2), 119–149.

Sorensen, A. (2000), 'Equilibrium price dispersion in retail markets for prescription drugs', *Journal of Political Economy* **108**(4), 833–850.

Stahl, D. (1989), 'Oligopolistic pricing with sequential consumer search', *American Economic Review* **79**(4), 700–712.

Sweeting, A. (2010), 'The effects of mergers on product positioning: evidence from the music radio industry', *RAND Journal of Economics* **41**(2), 372–397.

Varian, H. (1980), 'A model of sales', *American Economic Review* **70**(4), 651–659.

Veiga, A., Weyl, E. & White, A. (2017), 'Multidimensional platform design', *American Economic Review: Papers and Proceedings* **107**(5), 191–195.

Vossmeyer, A. (2014), 'Determining the proper specification for endogenous covariates in discrete data settings', *Advances in Econometrics* **34**, 223–247.

Waldfogel, J. (2003), 'Preference externalities: an empirical study of who benefits whom in differentiated-product markets', *RAND Journal of Economics* **34**(3), 557–568.

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

White, A. & Weyl, E. (2016), Insulated platform competition. http://ssrn.com/abstract=1694317.

Appendix: Proofs

Lemma 1

Proof. For a given cultivation profile, the platforms' objective functions are

$$\max_{p_A}\ p_A\left[n-\sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell)\right]$$
$$\max_{p_B}\ p_B\sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell).$$

The corresponding first-order necessary conditions are

$$p_A=\frac{n-\sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell)}{\sum_{\ell=1}^{\overline{Y}} n_\ell \phi\left(\omega_\ell\right)} \tag{16}$$
$$p_B=\frac{\sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell)}{\sum_{\ell=1}^{\overline{Y}} n_\ell \phi\left(\omega_\ell\right)} \tag{17}$$

multiplying by one judiciously (n^{-1}/n^{-1}), (16) and (17) are rewritten as

$$p_A=\frac{1-\sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n} \Phi(\omega_\ell)}{\sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n} \phi\left(\omega_\ell\right)}$$
$$p_B=\frac{\sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n} \Phi(\omega_\ell)}{\sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n} \phi\left(\omega_\ell\right)}.$$

Invoking Lemma 2, I substitute the for the representative consumer:

$$p_A=\frac{1-\Phi(\omega_0)}{\phi\left(\omega_0\right)} \tag{18}$$
$$p_B=\frac{\Phi(\omega_0)}{\phi\left(\omega_0\right)}. \tag{19}$$

Subtracting (19) from (18) yields

$$p_A-p_B=\frac{1-2\Phi(\omega_0)}{\phi\left(\omega_0\right)} \tag{20}$$

By Assumption 1(iv), the right-hand side of (20) is strictly decreasing in p_A while the left-hand side is unbounded and strictly increasing in p_A. A standard application of the intermediate value theorem shows that for any $\mathbf{q}^e$, there is a unique p_A-p_B satisfying the equilibrium conditions. Given p_A-p_B and (18) and (19), a unique p_A and p_B can be backed out for each $\mathbf{q}^e$ using (18) and (19). □

Lemma 2

Proof. The existence of such an individual ω_0 follows from two points. First note that $\Phi(\omega_\ell) \leq 1$ for all ℓ and $\sum_{\ell=1}^{\overline{Y}} n_\ell = n$, which implies that

$$\sum_{\ell=1}^{\overline{Y}} n_\ell \Phi(\omega_\ell) \leq n \Longrightarrow \sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n} \Phi(\omega_\ell) \leq 1.$$

By the continuity of $\Phi(\cdot)$, for every $\sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n} \Phi(\omega_\ell) = \Omega \in [0, 1]$, there must exist a value ω_0 such that

$$\Omega = \sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n} \Phi(\omega_\ell) = \Phi(\omega_0).$$

□

Propositions 1 and 2

Proof. From (18) and (19),

$$0 = -p_A(\mathbf{q}^e) + \frac{1 - \Phi(\omega_0^*)}{\phi\left(\omega_0^*\right)}.$$
$$0 = -p_B(\mathbf{q}^e) + \frac{\Phi(\omega_0^*)}{\phi\left(\omega_0^*\right)}.$$

Totally differentiating each (and setting those differentials not of interest to zero) yields

$$0 = -dp_A(\mathbf{q}^e) - \frac{1}{\phi(\omega_0^*)^2}\left(\phi(\omega_0^*)^2 + \phi'(\omega_0^*)[1 - \Phi(\omega_0^*)]\right)\Bigg(dp_A(\mathbf{q}^e) - dp_B(\mathbf{q}^e) - \Bigg[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \theta\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e}\frac{\partial \tilde{q}_{A0}^e}{\partial q_{A\ell}^e}\Bigg] dq_{A\ell}^e + \Bigg[\frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e}\frac{\partial q_{B\ell}^e}{\partial q_{A\ell}^e} + \theta\frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e}\frac{\partial \tilde{q}_{B0}^e}{\partial q_{A\ell}^e}\Bigg] dq_{A\ell}^e\Bigg)$$

$$0 = -dp_B(\mathbf{q}^e) + \frac{1}{\phi(\omega_0^*)^2}\left(\phi(\omega_0^*)^2 - \phi'(\omega_0^*)\Phi(\omega_0^*)\right)\Bigg(dp_A(\mathbf{q}^e) - dp_B(\mathbf{q}^e) - \Bigg[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \theta\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e}\frac{\partial \tilde{q}_{A0}^e}{\partial q_{A\ell}^e}\Bigg] dq_{A\ell}^e + \Bigg[\frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e}\frac{\partial q_{B\ell}^e}{\partial q_{A\ell}^e} + \theta\frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e}\frac{\partial \tilde{q}_{B0}^e}{\partial q_{A\ell}^e}\Bigg] dq_{A\ell}^e\Bigg).$$

As $q_{B\ell}^e = n_\ell - q_{A\ell}$ for all ℓ, $\frac{\partial q_B^e}{\partial q_{A\ell}^e} = -1$ and $\frac{\partial \tilde{q}_{B0}^e}{\partial q_{A\ell}^e} = -h_{0\ell}$.

Making these substitutions, multiplying both sides of each total differential by $1/dq_{A\ell}^e$, and grouping like terms yields

$$0 = -\left(2 + \frac{\phi'(\omega_0^*)[1 - \Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}\right)\frac{dp_A(\mathbf{q}^e)}{dq_{A\ell}^e} + \left(1 + \frac{\phi'(\omega_0^*)[1 - \Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}\right)\frac{dp_B(\mathbf{q}^e)}{dq_{A\ell}^e} + \left(1 + \frac{\phi'(\omega_0^*)[1 - \Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}\right)\left[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e} + \theta\left(\frac{\partial v(q_B^e, \theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e}\right)h_{0\ell}\right]$$

$$0=\left(1-\frac{\phi'(\omega_0^*)\Phi(\omega_0^*)}{\phi(\omega_0^*)^2}\right)\frac{dp_A(\mathbf{q}^e)}{dq^e_{A\ell}}-\left(2-\frac{\phi'(\omega_0^*)\Phi(\omega_0^*)}{\phi(\omega_0^*)^2}\right)\frac{dp_B(\mathbf{q}^e)}{dq^e_{A\ell}}$$
$$-\left(1-\frac{\phi'(\omega_0^*)\Phi(\omega_0^*)}{\phi(\omega_0^*)^2}\right)\underbrace{\left[\frac{\partial v(q_A^e,\theta\tilde{q}^e_{A0})}{\partial q^e_{A\ell}}+\frac{\partial v(q_B^e,\theta\tilde{q}^e_{B0})}{\partial q^e_{B\ell}}+\theta\left(\frac{\partial v(q_B^e,\theta\tilde{q}^e_{A0})}{\partial \tilde{q}^e_{A0}}+\frac{\partial v(q_B^e,\theta\tilde{q}^e_{B0})}{\partial \tilde{q}^e_{B0}}\right)h_{0\ell}\right]}_{\equiv\mu}.$$

For expositional convenience, I henceforth use μ to refer to the underbraced statement. The system can be rewritten in matrix form as

$$\underbrace{\begin{pmatrix}-2-\frac{\phi'(\omega_0^*)[1-\Phi(\omega_0^*)]}{\phi(\omega_0^*)^2} & 1+\frac{\phi'(\omega_0^*)[1-\Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}\\ 1-\frac{\phi'(\omega_0^*)\Phi(\omega_0^*)}{\phi(\omega_0^*)^2} & -2+\frac{\phi'(\omega_0^*)\Phi(\omega_0^*)}{\phi(\omega_0^*)^2}\end{pmatrix}}_{\equiv M}\begin{pmatrix}\frac{dp_A(\mathbf{q}^e)}{dq^e_{A\ell}}\\ \frac{dp_B(\mathbf{q}^e)}{dq^e_{B\ell}}\end{pmatrix}=\begin{pmatrix}-1-\frac{\phi'(\omega_0^*)[1-\Phi(\omega_0^*)]}{\phi(\omega_0^*)^2}\\ 1-\frac{\phi'(\omega_0^*)\Phi(\omega_0^*)}{\phi(\omega_0^*)^2}\end{pmatrix}\mu.$$

Note that

$$det(M)=3+\frac{\phi'(\omega_0^*)[1-2\Phi(\omega_0^*)]}{\phi(\omega_0^*)}\neq 0$$

so the system is solvable. Moreover, $\phi'(\omega_0^*)<0$ if and only if $\Phi(\omega_0^*)>1/2$, so $det(M)>0$. The comparative statics in Proposition 2 follow from a standard application of Cramer's rule. For Proposition 1, the statics are found by the same procedure but setting changes in platform B's installed base to zero.

I prove the monotonicity and non-monotonicity of the comparative statics without assuming differentiability for a more general result. First I illustrate monotonicity. Without loss of generality, consider expectations such that $q_A^e\le q_B^e$. Then, $\omega_0^*\ge\Phi^{-1}\left(\frac{1}{2}\right)$.

By items (ii) and (iii) of Assumption 1, $\Phi^{-1}(1/2)=0$. Therefore,

$$p_A(\mathbf{q}^e)-p_B(\mathbf{q}^e)\ge v\left(q_A^e,\theta\tilde{q}^e_{A0}\right)-v\left(q_B^e,\theta\tilde{q}^e_{B0}\right)\tag{21}$$

Now suppose that expectations are such that $\tilde{q}^e_{A0}\ge\tilde{q}^e_{B0}$. By Assumption 3(iii), there exists a θ' such that for all $\theta\ge\theta'$, the right-hand side of (21) is nonnegative, which implies that $p_A(\mathbf{q}^e)-p_B(\mathbf{q}^e)\ge 0$. This θ' is a function of the expectations $\mathbf{q}^e$.

To prove non-monotonicity, define $\overline{q}=\max\{\tilde{q}^e_{A0},\tilde{q}^e_{B0}\}$. Now suppose that q_j^e increases by one. Let $\eta_j<0$ and $\eta_{-j}>0$ denote the respective changes to $\tilde{q}^e_{j0}$ and $\tilde{q}^e_{-j0}$, where by construction $\eta_j<0$ and $\eta_{-j}>0$ iff $h_{0\ell}<0$ (when $h_{0\ell}=0$, $\eta_j=\eta_{-j}=0$). It is sufficient that the values of the networks move non-monotonically with market share:

$$v\left(q_j^e+1,\theta(\tilde{q}^e_{j0}+\eta_j)\right)\le v\left(q_j^e,\theta\tilde{q}^e_{j0}\right)$$
$$v\left(q_{-j}^e-1,\theta(\tilde{q}^e_{-j0}+\eta_{-j})\right)\ge v\left(q_{-j}^e,\theta\tilde{q}^e_{-j0}\right).$$

Define θ'' as the minimum θ that satisfies both inequalities. It follows that the inequalities are also satisfied for all $\theta\ge\theta''$. Setting $\theta^*(\mathbf{q}^e)=\max\{\theta',\theta''\}$ in the case of Proposition 1 and $\theta^{**}(\mathbf{q}^e)=\max\{\theta',\theta''\}$ in the case of Proposition 2 completes the proof. □

Proposition 3

Proof. For any fixed set of expectations ($\mathbf{q}^e$, a change in θ requires that the relative hedonic price of good A for the representative consumer,

$$\left[p_A(\mathbf{q}^e)-v\left(q_A^e,\theta\tilde{q}^e_{A0}\right)\right]-\left[p_B(\mathbf{q}^e)-v\left(q_B^e,\theta\tilde{q}^e_{B0}\right)\right]\equiv\omega_0^*,$$

must remain constant. Otherwise, the marginal consumer moves and expectations are no longer fulfilled. Moreover, platform j's expected hedonic price (as valued by the representative consumer), $p_j(\mathbf{q}^e) - v\left(q^e_j, \theta \tilde{q}^e_{j0}\right)$, cannot decrease given a change in θ. Otherwise, *ceteris paribus*, $p_j(\mathbf{q}^e)$ is not a maximizer, as it would imply the existence of a higher price p'_j that satisfies the equilibrium conditions.

Now, suppose that both platform A's network and platform B's network are $H+$. If θ increases, then both $v\left(q^e_A, \theta \tilde{q}^e_{A0}\right)$ and $v\left(q^e_B, \theta \tilde{q}^e_{B0}\right))$ increase, so to keep the hedonic prices and relative hedonic prices constant, $p_A(\mathbf{q}^e)$ and $p_B(\mathbf{q}^e)$ must both increase. By symmetry, if θ increases but both platforms' networks are $H-$, the equilibrium prices must both decrease. Next, suppose that platform A's network is $H+$ while platform B's network is $H-$. Then it follows that $p_A(\mathbf{q}^e)$ is increasing while $p_B(\mathbf{q}^e)$ is decreasing. By symmetry, the reverse holds when firm A's network is $H-$ and firm B's network is $H+$. If a platform is heterogeneity-neutral, then it follows from the above argument that its price is unaffected by θ. □

Corollary 1

Proof. Corollary 1 follows from Propositions 1-3. Consider expectations $q^e_A < q^e_B$ with $\tilde{q}^e_{A0} > \tilde{q}^e_{B0}$. By Lemma 2 and Assumption 3(iii), $p_A - p_B$ is strictly increasing in θ. Therefore, there exists a $\overline{\theta}(\mathbf{q}^e)$ such that if $\theta > \overline{\theta}(\mathbf{q}^e)$, then in the equilibrium of the subgame, $p_A(\mathbf{q}^e)q^e_A > p_{B(\mathbf{q}^e)}q^e_B$. □

Proposition 4

Proof. Proposition 4 follows from an envelop theorem argument. Fix expectations at $\mathbf{q}^e$. The platforms' objectives are given by

$$np_A(1 - \Phi(\omega_0)]$$
$$np_B\Phi(\omega_0).$$

Differentiating each with respect to q^e_{A0} and evaluating at $p(\mathbf{q}^e)$ yields $np_A\phi(\omega^*_0)\mu$ and $-np_B\phi(\omega^*_0)\mu$, respectively, where μ is defined in the Proof of Propositions 1 and 2. The statement regarding the signs of the comparative statics follows. □

Proposition 5

Proof. Without loss of generality, consider platform B. Recall that

$$\frac{1}{n}D_B(p_B, p_A, \mathbf{q}^e) = \sum_{\ell=1}^{\overline{Y}} \frac{n_\ell}{n}\Phi(\omega_\ell),$$

which by Lemma 2, can be rewritten as $n^{-1}D_B(p_B, p_A, \mathbf{q}^e) = \Phi(\omega^*)$. Therefore,

$$\frac{\partial D_B(p_B, p_A, \mathbf{q})}{\partial p_B} = -n\phi(\omega_0),$$

so platform B's elasticity is given by

$$E_B = -np_B\left(\frac{\Phi(\omega_0)}{\phi(\omega_0)}\right)^{-1}.$$

By Assumption 1(iv), $(\Phi(\omega_0)/\phi(\omega_0))^{-1}$ is strictly decreasing in ω_0, where

$$\begin{aligned}\omega_0 &= [p_A - v(q_A^e, \theta\tilde{q}_{A0}^e] - [p_B - v(q_B^e, \theta\tilde{q}_{B0})] \\ &= [p_A - p_B] - [v(q_A^e, \theta\tilde{q}_{A0}^e) - v(q_B^e, \theta\tilde{q}_{B0}^e)].\end{aligned}$$

As $v(q_B^e, \theta\tilde{q}_{B0}^e)$ is strictly increasing in $\tilde{q}_{B0}^e$ and $v(q_A^e, \theta\tilde{q}_{A0}^e)$ is strictly decreasing in $\tilde{q}_{A0}^e$, ω_0 is increasing in $\tilde{q}_{B0}^e$ and $\tilde{q}_{B0}^e - \tilde{q}_{A0}^e$, completing the proof. □

Lemma 3

Proof. Recall the objectives (14) and (15). At $p_A = p_A^*$ and $p_B = p_B^*$, (14) and (15) are upper semicontinuous. Given that $\mathbf{c}$ is defined over a compact space, a solution to the system exists. □

Proposition 6

Proof. Without loss of generality, consider platform A. The FOC with respect to c_{Ay_r} is given by

$$p_A(\mathbf{q}^e(\mathbf{c}^*))\phi(\omega_0^*) \sum_{\ell:y_r\in\ell} \left[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e} + \theta \left(\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e} \right) h_{0\ell} \right] \frac{\partial f(q_{A\ell}, \mathbf{c}^*)}{\partial c_{Ay_r}} = \frac{\partial \kappa(\mathbf{c}_A^*)}{\partial c_{Ay_r}},$$

where

$$\frac{\partial f(q_{j\ell}, \mathbf{c}^*)}{\partial c_{jy_r}} = \begin{cases} \frac{\alpha_r}{\sum_{y_s\in\ell}\alpha_s} & \text{if } q_{A\ell}^e(\mathbf{c}^*) = q_{j\ell}^e + \frac{\sum_{y_s\in\ell}\alpha_s(c_{jy_s} - c_{-jy_s})}{\sum_{y_s\in\ell}\alpha_s} \\ 0 & \text{otherwise.} \end{cases}$$

For statement (i), as $\theta \to 0$, the FOC becomes

$$p_A(\mathbf{q}^e(\mathbf{c}^*))\phi(\omega_0^*) \sum_{\ell:y_r\in\ell} \left[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e} \right] \frac{\partial f(q_{A\ell}, \mathbf{c}^*)}{\partial c_{Ay_r}} = \frac{\partial \kappa(\mathbf{c}_A^*)}{\partial c_{Ay_r}}.$$

At $\mathbf{c} = \mathbf{0}$,

$$p_A(\mathbf{q}^e(\mathbf{c}^*))\phi(\omega_0^*) \sum_{\ell:y_r\in\ell} \left[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e} \right] \frac{\partial f(q_{A\ell}, \mathbf{0})}{\partial c_{Ay_r}} > 0$$

for at least one y_r. By Assumption 2(iv), $\kappa(\mathbf{0}) \to 0$. By continuity, this holds for θ sufficiently small. Define $\underline{\theta}(\mathbf{q}^e)$ as the largest θ such that the inequality holds.

Statement (ii) follows from $p_A(\mathbf{q}^e(\mathbf{c}^*))\phi(\omega_0^*) > 0$.

For statement (iii), consider all ℓ such that $\{\ell : y_r \in \ell\} = \{\ell : y_s \in \ell\}$. The corresponding FOCs are

$$p_A(\mathbf{q}^e(\mathbf{c}^*))\phi(\omega_0^*) \sum_{\ell:y_r\in\ell} \left[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e} + \theta \left(\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e} \right) h_{0\ell} \right] \frac{\partial f(q_{A\ell}, \mathbf{c}^*)}{\partial c_{Ay_r}} = \frac{\partial \kappa(\mathbf{c}_A^*)}{\partial c_{Ay_r}}$$

$$p_A(\mathbf{q}^e(\mathbf{c}^*))\phi(\omega_0^*) \sum_{\ell:y_r\in\ell} \left[\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial q_{A\ell}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial q_{B\ell}^e} + \theta \left(\frac{\partial v(q_A^e, \theta\tilde{q}_{A0}^e)}{\partial \tilde{q}_{A0}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{B0}^e)}{\partial \tilde{q}_{B0}^e} \right) h_{0\ell} \right] \frac{\partial f(q_{A\ell}, \mathbf{c}^*)}{\partial c_{Ay_r}} = \frac{\partial \kappa(\mathbf{c}_A^*)}{\partial c_{Ay_s}}.$$

The left-hand side (LHS) of each FOC is identical except for $\partial f(q_{A\ell}, \mathbf{c}^*)/\partial c_{Ay_r}$. Thus the LHS of the FOC for c_{Ay_r} is greater (less than) [equal to] the LHS of the FOC for c_{Ay_s} if and only if

$$\frac{\alpha_r}{\sum_{y_t\in\ell}\alpha_t} > \ (<)\ [=] \frac{\alpha_s}{\sum_{y_t\in\ell}\alpha_t}$$
$$\alpha_r > \ (<)\ [=]\ \alpha_s.$$

An analogous argument holds for platform B. □

Proposition 7

Proof. Without utilizing Lemma 2, the FOC with respect to c_{Ay_r} is given by

$$p_A(\mathbf{q}^e(\mathbf{c}^*)) \sum_{m=1}^{\overline{Y}} \phi(\omega_m^*) \sum_{\ell:y_s\in\ell} \left[\frac{\partial v(q_A^e, \theta\tilde{q}_{Am}^e)}{\partial q_{Am}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{Bm}^e)}{\partial q_{Bm}^e} + \theta \left(\frac{\partial v(q_A^e, \theta\tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{Bm}^e)}{\partial \tilde{q}_{Bm}^e} \right) h_{m\ell} \right] \frac{\alpha_s}{\sum_{y_t\in\ell}\alpha_t} \leq \frac{\partial \kappa(\mathbf{c}_A^*)}{\partial c_{Ay_r}}.$$

At $\mathbf{c} = \mathbf{0}$, $\omega_m^* = \omega_\ell^*$ for all m, ℓ. It is sufficient to show that the LHS of the above is nonpositive at $\mathbf{c} = 0$. Utilizing the relationship between ω_m^* and ω_ℓ^*, $c_{Ay_r}^* = 0$ if

$$\sum_{m=1}^{\overline{Y}} \sum_{\ell:y_s\in\ell} \left[\frac{\partial v(q_A^e, \theta\tilde{q}_{Am}^e)}{\partial q_{Am}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{Bm}^e)}{\partial q_{Bm}^e} + \theta \left(\frac{\partial v(q_A^e, \theta\tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} + \frac{\partial v(q_B^e, \theta\tilde{q}_{Bm}^e)}{\partial \tilde{q}_{Bm}^e} \right) h_{m\ell} \right] \frac{\alpha_s}{\sum_{y_t\in\ell}\alpha_t} \leq 0.$$

For θ sufficiently large,

$$\sum_{m=1}^{\overline{Y}} \sum_{\ell:y_s\in\ell} \left[\left(\frac{\partial v(q_A^e, \theta\tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} \right) h_{m\ell} \right] \leq 0$$

is sufficient. Given the finite summations, the above can be rewritten as

$$\sum_{\ell:y_s\in\ell} \sum_{m=1}^{\overline{Y}} \left[\left(\frac{\partial v(q_A^e, \theta\tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} \right) h_{m\ell} \right] \leq 0.$$

There exists a positive ε such that if $\sum_{m\neq\ell} |h_{m\ell}| - h_{\ell\ell} > \varepsilon$, then for each ℓ,

$$\sum_{m=1}^{\overline{Y}} \left[\left(\frac{\partial v(q_A^e, \theta\tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} \right) h_{m\ell} \right] \leq 0,$$

If true for each ℓ, then the inequality also holds for all $\ell : y \in \ell$, proving the first statement.

To prove statements (i) and (ii), assume sufficiently strong in-group bias. Reversing the inequalities, an identical argument as above follows, except if $\sum_{m\neq\ell}|h_{m\ell}| - h_{\ell\ell} < -\varepsilon$ for a positive ε, then it follows that

$$\sum_{m=1}^{\overline{Y}} \left[\left(\frac{\partial v(q_A^e, \theta \tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} \right) h_{m\ell} \right] > 0.$$

□

Proof of Proposition 8

Proof. This proof follows a nearly identical argument to that of Proposition 7. Without utilizing Lemma 2, the FOC with respect to c_{Ay_r} is given by

$$p_A(\mathbf{q}^e(\mathbf{c}^*)) \sum_{m=1}^{\overline{Y}} \phi(\omega_m^*) \sum_{\ell:s\in\ell} \left[\frac{\partial v(q_A^e, \theta \tilde{q}_{Am}^e)}{\partial q_{Am}^e} + \frac{\partial v(q_B^e, \theta \tilde{q}_{Bm}^e)}{\partial q_{Bm}^e} \right.$$
$$\left. + \theta \left(\frac{\partial v(q_A^e, \theta \tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} + \frac{\partial v(q_B^e, \theta \tilde{q}_{Bm}^e)}{\partial \tilde{q}_{Bm}^e} \right) h_{m\ell} \right] \frac{\alpha_r}{\sum_{y_t\in\ell} \alpha_t} \leq \frac{\partial \kappa(\mathbf{c}_A^*)}{\partial c_{Ay_r}}.$$

at $\mathbf{c} = 0$, we require the LHS to be nonpositive. Assuming θ sufficiently large, The above then simplifies to

$$\sum_{m=1}^{\overline{Y}} \phi(\omega_m^*) \sum_{\ell:y_s\in\ell} \left[\left(\frac{\partial v(q_A^e, \theta \tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} + \frac{\partial v(q_B^e, \theta \tilde{q}_{Bm}^e)}{\partial \tilde{q}_{Bm}^e} \right) h_{m\ell} \right] \leq 0.$$

Define

$$A_m = \phi(\omega_m^*) \left(\frac{\partial v(q_A^e, \theta \tilde{q}_{Am}^e)}{\partial \tilde{q}_{Am}^e} + \frac{\partial v(q_B^e, \theta \tilde{q}_{Bm}^e)}{\partial \tilde{q}_{Bm}^e} \right) > 0,$$

so the above can be rewritten as $\sum_{m=1}^{\overline{Y}} \sum_{\ell:y_s\in\ell} A_m h_{m\ell} \leq 0$. The argument then proceeds identically to that of Proposition 7. □

Proposition 9

Proof. Without loss of generality, suppose that $|h_{11}| > |h_{22}| > \cdots > |h_{\overline{YY}}|$ and $h_{m\ell} = h'$ for all $m \neq \ell$. Recall from Proposition 4 that

$$\frac{\partial \pi_A(\mathbf{p}(\mathbf{q}^e))}{\partial q_{A\ell}^e} = np_A(\mathbf{q}^e) \sum_{m=1}^{\overline{Y}} \phi(\omega_m^*) \left[\frac{\partial v(q_A^e, \theta \tilde{q}_{Am}^e)}{\partial q_{A\ell}^e} + \frac{\partial v(q_B^e, \theta \tilde{q}_{Bm}^e)}{\partial q_{B\ell}^e} + \theta \left(\frac{v(q_A^e, \theta \tilde{q}_{Am})}{\partial \tilde{q}_{Am}^e} + \frac{v(q_B^e, \theta \tilde{q}_{Bm})}{\partial \tilde{q}_{Bm}^e} \right) h_{m\ell} \right].$$

For θ sufficiently large and sufficiently strong out-group bias: $\sum_{m\neq\ell} h_{\ell m} - |h_{\ell\ell}| < -\varepsilon$ for a positive ε, profits are decreasing in $q_{A\ell}^e$ if

$$\sum_{m=1}^{\overline{Y}} h_{m\ell} < 0.$$

By Definition 5, $h_{m\ell} > 0$ for all $m \neq \ell$ and $h_{\ell\ell} < 0$ for all ℓ. Aggregating over all ℓ, profits are maximized by minimizing $\sum_\ell q_{A\ell} h_{\ell\ell}$. With the $h_{\ell\ell}$ fixed, for every total size q_A^e, profits are maximized when $q_{A1} < q_{A2} < \cdots < q_{A\overline{Y}}$. An analogous argument holds for platform B. □

Proposition 10

Proof. Suppose that $K = 2$, $y_r \in \ell \Rightarrow y_r \notin m \; \forall m \neq \ell$, $\mathbf{q}^e = \mathbf{q}^e = \mathbb{E}\mathbf{Q}^e$ and that θ is sufficiently large. As $y_r \in \ell \Rightarrow y_r \notin m \; \forall m \neq \ell$, I use traits and types interchangeably. As θ sufficiently large, size effects can be ignored and composition effects dominate. By Lemma 3, at least one equilibrium exists and is given by the solution to the system of $2\overline{Y}$ FOCs.

For (i), suppose that $\overline{h} - \underline{h} < \varepsilon$ for a small positive ε. Consider the cultivation profile $\mathbf{c}^*$ such that $c^*_{A\ell} > 0$ and $c^*_{B\ell} = 0$ for all $\ell \in L_1$, while $c^*_{A\ell} = 0$ and $c^*_{B\ell} > 0$ for all $\ell \in L_2$. Under $\mathbf{c}^*$, $q^e_A = n/2$ and $q^e_{A\ell} > n_\ell/2$ for all $\ell \in L_1$ and $q^e_{A\ell} < n_\ell/2$ for all $\ell \in L_2$. Now, consider a unilateral deviation by platform A to a profile $\mathbf{c}'$, where all values are identical to $\mathbf{c}^*$ except that $c_{A\ell} = \varepsilon'$ for some $m\ell \in L_2$ and arbitrarily small positive ε'. By Assumption 2, if this deviation does not exist, there can be no deviation to a value greater than ε'. Introduction of a user from L_2 imposes a negative externality on all users from L_1. By proposition 4,

$$\frac{\partial \pi_A(\mathbf{p}(\mathbf{q}^e))}{\partial q^e_{A\ell}} = np_A(\mathbf{q}^e) \sum_{m=1}^{\overline{Y}} \phi(\omega^*_m) \left[\left(\frac{v(q^e_A, \theta \tilde{q}_{Am})}{\partial \tilde{q}^e_{Am}} + \frac{v(q^e_B, \theta \tilde{q}_{Bm})}{\partial \tilde{q}^e_{Bm}} \right) h_{m\ell} \right] dq^e_{A\ell}$$

as $h_{m\ell} < 0$ for all $m \in L_1$ and $q^e_{Am} > q^e_{A\ell}$ for each $m \in L_1$ and $\ell \in L_2$. The composition worsens for a strict majority of users, and improves for a small minority. As $\overline{h} - \underline{h} < \varepsilon$, the net effect is negative. Thus there is no profitable deviation.

Now consider a unilateral deviation to a profile $\mathbf{c}''$, where all is identical except that $c_{A\ell} = 0$ for some $\ell \in L_1$. Then $q^e_{A\ell}(\mathbf{c}'') = n_\ell/2 < \mathbf{q}^e(\mathbf{c}^*)$. Moreover, $\tilde{q}^e_{A\ell}(\mathbf{c}'') < \tilde{q}^e_{A\ell}(\mathbf{c}^*)$ for all $\ell \in L_1$. Given Assumption 2(iv), $\mathbf{c}''$ yields a strict decrease in platform A's profits via a lower price due to poorer composition and fewer users due to the loss of a cultivated trait. By symmetry, a similar argument holds for platform B.

To prove uniqueness, consider any profile $\mathbf{c}' \neq \mathbf{c}^*$. Cases in which a platform targets a subset of a partition are ruled out by the above argument. The only remaining potential candidate $\mathbf{c}'$ is both platforms targeting the same partition. However, if one platform deviates by targeting the alternate partition, it receives a strict increase in profits through increased composition. The mathematical argument is analogous to the one presented above.

For (ii), suppose that, without loss of generality, $\hat{h} \in L_1$. Consider the profile $\mathbf{c}^*$ with $c^*_{j\ell} > 0$ for all j and all $\ell \in L_1$. By symmetry, $c^*_{A\ell} = c^*_{B\ell}$ for all ℓ, so $q^e_A = q^e_B = n/2$ and $q^e_{A\ell} = q^e_{B\ell} = n_\ell/2$ for all ℓ. A deviation by platform A to $\mathbf{c}'$, in which all is identical except $c_{A\ell} < c_{B\ell}$ for some $\ell \in L_1$ yields a change in profits of

$$p_A(\mathbf{q}^e) \sum_{m=1}^{L} n_m \phi(\omega^*_m) \left[\frac{\partial v(q^e_A, \theta \tilde{q}^e_{Am})}{\partial q^e_{A\ell}} + \frac{\partial v(q^e_B, \theta \tilde{q}^e_{Bm})}{\partial q^e_{B\ell}} + \theta \left(\frac{v(q^e_A, \theta \tilde{q}_{Am})}{\partial \tilde{q}^e_{Am}} + \frac{v(q^e_B, \theta \tilde{q}_{Bm})}{\partial \tilde{q}^e_{Bm}} \right) h_{m\ell} \right] dq^e_{A\ell}.$$

As $q^e_A(\mathbf{c}') < q^e(\mathbf{c}^*)$, there is a negative size effect. For all type m users, $m \in L_1$, $\tilde{q}^e_{Am}(\mathbf{c}') < \tilde{q}^e_{Am}(\mathbf{c}^*)$. Moreover, for each $m \notin L_1$, the net increase $\tilde{q}^e_{Am}(\mathbf{c}') - \tilde{q}^e_{Am}(\mathbf{c}^*) < \tilde{q}^e_{Am}(\mathbf{c}^*) - \tilde{q}^e_{Am}(\mathbf{c}')$. Thus $p_A(\mathbf{q}(\mathbf{c}')) < p_A(\mathbf{q}(\mathbf{c}^*))$. By group 1's dominance in Definition 6, this implies a strict decrease in profits.

The only alternative deviation necessary to consider is platform A cultivating a type m individual for $m \in L_2$. Again, by the definition of dominance, the negative externality imposed on group 1 dominates the positive externality on group 2, implying a strict decrease in profits. Thus there is no unilateral deviation. □

Proposition 11

Proof. As in the text, if there is a single platform A, then a constant ξ_A and $\theta = 0$ implies that profits are maximized at $p^*(0) = \xi_A + v(q^e_A, 0)$, where $p^*(0)$ is $p^*(\theta)$ evaluated at $\theta = 0$. In equilibrium, $q^e_A = n$. Profits are np^*. Under J, platforms, each platform j maximizes profits at price $p^*_j(0) = \xi_A + v(q^e_j, 0)$. If $q^e_j < n$, then $p^*(0) > p^*_j(0)$ for all j. Total profits across the J platforms is $\sum_j q^e_j p^*_j(0) < \sum_j q^e_j p^*(0) = np^*(0)$. By continuity, the inequality holds for small but positive θ, completing the proof. □

Proposition 12

Proof. Suppose that θ is sufficiently large and H consists of either in-group bias, out-group bias, or grouping. Define $\underline{\ell}$ as the type ℓ user with the lowest heterogeneity-weighted installed base value in platform A.

A single platform A maximizes prices by setting

$$p^*(\theta) = \begin{cases} \xi_A + v(n, \theta \tilde{q}^e_{A\underline{\ell}}) & \text{if } q^e_A = n \\ \xi_A + v(q^e, \theta \tilde{q}^e_{A\underline{\ell}}) & \text{if } q^e_A < n. \end{cases}$$

If $q^e < n$, then creating a second platform with the remaining $n - q^e_A$ users completes the proof. If $q^e_A = n$, then by the construction of H, there must exist at least one type ℓ such that by removing type ℓ, $p^*(\theta)$ increases. Then by the above case, creating at least one new platform for the removed type(s) completes the proof. □