Connecting Heterogeneous Agents under Incomplete Information[*]

Masaki Aoyagi[†]

June 5, 2020

Abstract

This paper studies the problem of a monopolistic platform which offers agents connection with one another. Agents have heterogeneous characteristics that are valued by some other agents and observed privately by the principal. The agents are privately informed about their heterogeneous preferences over the characteristics of the other agents. The platform solicits information from the agents about their preferences and then offers an allocation that consists of groups of connected agents and subscription fees. We study mechanisms which induce truthful reporting and acceptance of the proposed allocation as a unique equilibrium outcome. We identify asymptotically optimal mechanisms which fully extract the agents' informational rents in the limit as the market becomes large.

Key words: platform, network externalities, random graph, revenue maximization, strategy-proof.
JEL Codes: D42, D47, D62, D82, L12

1   Introduction

Platforms as intermediaries of economic activities are gaining importance in the modern economy. Credit cards facilitate trade between businesses and consumers, internet auction houses bring out demand and supply and realize gains from trade, and SNS connects people and disseminates information that helps decision making. Economic theory has by now devoted much attention to the functioning of platforms,

1

and offers a variety of theories. This paper presents yet another theory of platforms focusing on their role in connecting heterogeneous market participants. Specifically, we focus on the monopoly power exercised by a platform based on its proprietary access to information about market participants, and study its consequence on the design of a mechanism that offers connections among them. Specifically, we consider a market in which a platform has information about some qualitative characteristics of agents referred to here as their *salience types.*[1] An agent's characteristic for example is the type of good a seller supplies to the market, the type of skill a worker can provide to a prospective employer, or the type of project that an entrepreneur requires funding for. In short, a salience type of an agent is his value to various other market participants. The assumption on proprietary access to such information by a platform models the ability of large online platforms to privately collect big data on the subscribers' activities and use them to project their characteristics.[2]

We assume that the agents are heterogeneous in the sense that no two agents share the same characteristics. Although the platform has access to information about the agents' salience types, we assume that it does not directly observe their preferences over those types, and must collect information about them. We represent these preferences of each agent by their *preference types*, and assume for simplicity that each preference type regards each salience type as either preferable or not. The agents are heterogeneous also in their preferences so that no salience type is ex ante more popular than any other salience type. Agents value connection to any agent with their preferred salience type, but must use the service of a monopolistic platform to achieve it. The platform offers them connections by forming groups whose members are connected with each other. The value of a group to each member i equals $v > 0$ if i finds preferable the salience type of another member, and zero otherwise. An *allocation* is a pair of such a group assignment and a subscription fee profile. The platform's objective is to maximize the subscription revenue by creating an optimal group assignment. In contrast to the literature on platform design that assumes that agents are vertically differentiated by their quality types and that every agent uniformly prefers to be matched with agents of higher quality, we study the problem under the assumption of horizontal differentiation across agents.[3]

Formally, a mechanism used by the platform forms an allocation based on reported preference types. We require such a mechanism to be *strategy-proof* in the sense that truth-telling is weakly dominant for every agent when reporting their

---

[1]We use the term "salience type" following Gomes & Pavan (2016).

[2]A similar motivation is discussed in Fainmesser & Galeotti (2016). See Section 2.

[3]While we assume that the platform observes the agents' salience types but not their preference types, our argument would be essentially unchanged if it observes the preference types and not the salience types.

preference types. By definition, implementation of a strategy-proof mechanism does not require knowledge about the probability distribution of the agents' types. Unless the platform can force an allocation on the agents, however, strategy-proofness does not itself solve the multiplicity of equilibria under adoption externalities. That is, if the agents may choose whether or not to accept the platform's offer of an allocation, their optimal decision depends on the acceptance decisions of other agents: The agent may be willing to accept if and only if the other agents also accept. The multiplicity of equilibria in such an adoption game is a central concern in the analysis of network externalities, and our analysis is no exception.[4] In order to address this concern, we build on the idea of a revelation-suggestion mechanism proposed by Myerson (1982), and consider a mechanism consisting of two stages: The platform creates an allocation based on solicited preference types in the first stage, and lets the agents play an adoption game in the second stage. The mechanism is designed so that acceptance is a uniquely optimal action for each agent. As far as we are aware, this is the first attempt to introduce such a requirement in the formulation of a mechanism.

In our model, whether agent i finds the salience type of agent j preferable or not is independent of whether j finds i's salience type preferable, or whether i finds the salience type of yet another agent $k \neq j$ preferable. We denote by p the probability with which each agent finds the salience type of any other agent preferable. When the total number of agents equals n, the value $(n-1)p$ is the key quantity that represents the expected number of agents whose salience types each agent finds preferable, and hence can be interpreted as the density of the preference relationship among them.

Our objective is to characterize a mechanism that fully extracts the agents' informational rents and achieves the first-best allocation for the platform. First, we observe that in any market of fixed size n and for any probability $p \in (0, 1)$, no mechanism extracts full surplus if it is required to induce truth-telling in the first stage, and acceptance as a uniquely optimal action in the second stage. This observation prompts us to examine what happens when the market size n grows large and the probability p may change with n. In particular, we consider the limit as n grows large while focusing on the behavior of $(n-1)p$ as a function of n.

Our analysis identifies asymptotically optimal mechanisms in the following two cases. We call a market *dense* if $(n-1)p$ increases (slowly) as n grows large. The case where p is a constant is a special case of a dense market. On the other hand, we call a market *sparse* if $(n-1)p$ vanishes as n grows large. In a dense market,

---

[4]See Dybvig & Spatt (1983). The multiplicity problem is also known as "chicken and egg" in the two-sided market literature. See for example Caillaud & Jullien (2003).

we consider a *single group mechanism* that creates a single group consisting of a prefixed agent i_1 and all those agents who find the salience type of i_1 preferable either directly or indirectly in any number of steps: These are set J_1 of agents who find i_1 preferable, set J_2 of agents who find agents in J_1 preferable, set J_3 of agents who find agents in J_2 preferable, and so on. Agent i_1 is offered a small subsidy for inclusion into the group, but all other agents who are included in the group are charged a subscription fee close to their valuation of the group. The single group mechanism induces truth-telling and also induces acceptance as a uniquely optimal action through the iterative elimination of strictly dominated actions. When the condition for a dense market holds, this mechanism is asymptotically optimal in the sense that as n grows large, the associated revenue approaches the maximal level achievable under complete information about the agents' preferences. In a sparse market, on the other hand, we propose a *multiple group mechanism* which constructs multiple groups which again include two sets of agents. The first set consists of agents who themselves find (the salience type of) no other agent preferable but have salience types which are found preferable by others, or the set of agents in which every agent finds the salience type of another agent preferable either directly or indirectly. The second set consists of agents who find the salience type of some of the first set of agents preferable either directly or indirectly. Again, the first set of agents are offered a small subsidy for subscription into their group, but the second set of agents are charged a subscription fee close to their valuation of their group. The multiple group mechanism induces truth-telling, and also induces acceptance as a uniquely optimal action. Under the conditions for a sparse market, we show that this mechanism is asymptotically optimal in the limit as n grows large.

The paper is organized as follows. We discuss the related literature in Section 2 and formulate a model in Section 3. Section 4 provides the description of the mechanism and the properties required of it. The benchmark case of complete information is analyzed in Section 5. Section 6 presents the analysis of the models with incomplete information: Subsection 6.1 analyzes the single group mechanism for a dense market, and Subsection 6.2 analyzes the multiple group mechanism for a sparse market. We conclude with a discussion in Section 7.

2   Related Literature

The present paper belongs to the literature that studies the problem of a monopolistic platform which offers privately informed agents connection to each other. Damiano & Li (2007) study a two-sided market where each side of the market consists of agents with heterogeneous quality, and the value of a match between two

4

agents is a supermodular function of the two individual qualities. When a monopolist divides the market into groups within which matches can be formed, Damiano & Li (2007) derive the conditions under which the profit-maximizing fee schedules for the groups are efficient. Adopting the framework of Damiano & Li (2007), Hoppe et al. (2011) analyze the performance of the coarse matching scheme in which the market is divided into only two groups. Gomes & Pavan (2016) study efficient and profit-maximizing platforms for many-to-many matching in a two-sided market where agents have private types. Gomes & Pavan (2016) discuss the distinction between the salience and preference types of an agent, and assume in their main analysis that those two types are functions of a single private type of an agent. In Board (2009), a monopolist in a one-sided market creates and assigns groups to the population of agents with heterogeneous quality. Under various specifications of the functional relationship between the quality of a group and the qualities of its members, Board (2009) shows that the profit maximizing group structure is inefficient from a welfare point of view and attributes the inefficiencies to the creation of too many groups. On the other hand, Veiga (2013) shows that when the value of a larger group is always higher than that of a smaller group, the profit maximizing group structure always entails a single group.

There are some important differences between the present paper and the papers in the literature mentioned above: First, the ubiquitous assumption in the literature is that agents are vertically differentiated, and every agent prefers to be matched with agents of a higher type. In contrast, we suppose that agents are horizontally differentiated and their preferences over other agents are heterogeneous. In particular, an agent does not necessarily prefer to be matched with an agent who happens to be popular among other agents. Second, we assume that each agent's type is two-dimensional and one of the dimensions is privately observed by the platform. This assumption corresponds to the reality of online platforms that have proprietary access to data on their subscribers. Third, the literature assumes a screening mechanism in which an agent's allocation is determined by his type only and not by the entire type profile. In contrast, we consider a full mechanism where allocation is determined by the type profile of all agents.

A platform is a good with network externalities in the sense that its value to each agent depends positively on the adoption decisions of other agents. We may view the platform in our model trying to sell connection to agents. Monopoly sale of a network good under incomplete information is studied by Aoyagi (2013), where the value of the good to each agent is a function of the agent's private type and the number of subscriptions. Some models of network goods monopoly express local network externalities by agents located at the nodes of a network as in the present

paper. Among them, Fainmesser & Galeotti (2016) explore the possibilities that a platform has varying degrees of knowledge about the externalities relationship among consumers. Specifically, they study scenarios where the platform has access to information on how susceptible or influential each consumer is to the decisions of others. Although we also suppose that a platform has access to information about market participants, our formulation is different in that the realization of the externalities relationship is random and that the platform needs to collect information about it.[5]

Extensive literature on matching theory is on the design of a mechanism that matches agents with heterogeneous preferences as in the present model. Most closely related among them are the papers on coalition formation, which consider the problem of partitioning the set of agents into groups when the value of a group to any member is a function of the types of other members. For example, Cechlárová & Romero-Medina (2001) assume that the value of a group to any of its member is equal to the value of either his most-preferred member, or his least-preferred member, Alcalde & Revilla (2004) introduce a preference which evaluates groups based on the best subset of them, Dimitrov et al. (2006) suppose that each agent regards other agents as either friends or enemies and compares the numbers of friends and enemies when evaluating the quality of a group, and Rodríguez-Álvarez (2009) considers the preference domain that includes the friend-enemy preferences studied by Dimitrov et al. (2006). However, the primary focus of the literature on coalition formation is on the stability of matching, whereas our focus is on profit maximization through the extraction of informational rents. The agents' preferences over other agents assumed in this paper are binary so that each agent finds the salience type of another agent either preferable or not. This substantial simplification of the general preference specifications used in the matching literature enables us to characterize mechanisms that achieve asymptotic optimality under the unique enforceability requirement.

3 Model

The market consists of the set I of n agents. Each agent i has a *salience type* β_i and a *preference type* θ_i. Agents are *heterogeneous* in the sense that $\beta_i \neq \beta_j$ for

---

[5]Ohter models of monopoly under local network externalities include Candogan et al. (2012), who characterize the relationship between the location of a buyer in the network and the price he faces under imperfect and perfect price discrimination, and Bloch & Quérou (2013), who examine the optimality of price discrimination when each buyer is privately informed about the stand-alone valuation of the monopolist's good. Models of price competition between sellers of goods with local network externalities include Aoyagi (2018), Chen et al. (2018), and Chen et al. (2020).

any pair of agents i and j. Specifically, when B denotes the set of salience types in the population, there is a one-to-one correspondence $\phi : I \to B$ such that $\phi(i) \in B$ is the salience type of agent i. Let Φ be the set of all such correspondences. We assume that the platform privately observes the random realization of ϕ, but the agents don't. Our interpretation of this assumption is that it is prohibitively costly for participants in a large and physically spread market to find out if any other agent has the characteristics by randomly approaching them.

Agent i's preference type θ_i describes their preferences over the salience types of other agents. These preferences classify each salience type as either preferable or not. Formally, given ϕ, agent i's preference type θ_i is a subset of the set $B \setminus \{\phi(i)\}$ of salience types of other agents, and $\phi(j) \in \theta_i$ implies that i finds salience type $\phi(j)$ of agent j *preferable*. Let $\theta = (\theta_i)_{i \in I}$ denote a preference type profile, and Θ_ϕ be the set of preference type profiles under ϕ. A visual representation of (θ, ϕ) is a directed graph in which the agents are nodes and the type profile θ represents directed links among them: If i finds j's salience type $\phi(j)$ preferable (*i.e.*, $\phi(j) \in \theta_i$), then it is identified with a directed link from i to j and denoted $i \to j$. We call this directed graph a *preference network* and denote it by $D(\theta, \phi)$. In $D(\theta, \phi)$, agent i *indirectly prefers* agent j, denoted $i \rightsquigarrow j$, if there exists a sequence of agents $i_0, \ldots, i_K$ ($K \geq 1$) in I such that $i_0 = i$, $i_K = j$, and for every i_k and i_{k+1} ($k = 0, \ldots, K-1$), i_k finds $\phi(i_{k+1})$ preferable (*i.e.*, $\phi(i_{k+1}) \in \theta_{i_k}$). In terms of the preference network $D(\theta, \phi)$, i indirectly prefers j if there is a directed path from i to j: a directed link $i_k \to i_{k+1}$ exists for every such k. For each θ and $i \in I$, define

$$C_{\rightsquigarrow i}(\theta) = \{j \neq i : j \rightsquigarrow i\}$$

to be the set of agents who indirectly prefer $\phi(i)$.[6] A subset $H \subset I$ of two or more agents is *strongly preference-connected* if $i \rightsquigarrow j$ for every $i, j \in H$. $H \subset I$ is a *strong component* if it is strongly preference-connected and there exists no $H' \not\supset H$ that is strongly preference-connected. A *component* of the preference network $D(\theta, \phi)$ is the set of agents who are connected by the preference links when their direction is ignored.

A *group* is a set of two or more salience types, and an *assignment* is a collection g of disjoint (non-overlapping) groups. The interpretation is that agents i and j are connected to each other if and only if their salience types belong to the same group. Let $\mathcal{G}$ be the set of all possible assignments. Denoting by $S \equiv S(g)$ the set of indices of groups in assignment $g \in \mathcal{G}$, we express $g \in \mathcal{G}$ as $g = (G_s)_{s \in S}$, where each G_s is a group in g. Given $g = (G_s)_{s \in S}$, let $G = \cup_{s \in S} G_s$ be the set of salience types included

---

[6]For simplicity, ϕ is omitted in the description of a preference network when it is understood. We also often say that agent i prefers agent j when i finds j's salience type $\phi(j)$ preferable.

in some group in g. For example, if $B = \{1, \ldots, 8\}$, then $g = (\{1, 3\}, \{5, 7, 8\}) \in \mathcal{G}$ is an assignment that consists of two groups ($S = \{1, 2\}$) and excludes salience types 2, 4 and 6. Agent i's valuation $u_i(g, \phi, \theta_i)$ of assignment $g = (G_s)_{s \in S}$ is determined by his preference type θ_i as follows: For a constant $v > 0$,

$$u_i(g, \phi, \theta_i) = \begin{cases} v & \text{if } \phi(i) \in G_s \text{ and } \theta_i \cap G_s \neq \emptyset \text{ for some } s \in S, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

In other words, the value of an assignment to any agent i equals v if there is another member of the same group who has his preferred salience type, and zero if there is no such member, or if i does not belong to any group.

Upon observing the agents' salience types, a platform collects information about their preference types and determines an *allocation,* which consists of an assignment and transfers. Formally, an *allocation rule* is a pair (g, x), where $g : \Theta \times \Phi \to \mathcal{G}$ and $x = (x_i)_{i \in I} : \Theta \times \Phi \to \mathbf{R}^I$. In other words, $g(\theta, \phi)$ and $x(\theta, \phi) = (x_i(\theta, \phi))_{i \in I}$ are the assignment and subscription fee profile, respectively, under the reported preference type profile θ and the salience type profile ϕ. We also denote by $S(\theta, \phi)$ the set of indices of groups in $g(\theta, \phi)$, and by $G_s(\theta, \phi)$ the group $s \in S(\theta, \phi)$.

Each correspondence $\phi \in \Phi$ that determines the agents' salience types is equally likely. Conditional on ϕ, the probability distribution of the type θ_i is independent and identical across agents. We also suppose that for any agent i, his preferences over other agents j and k are independent (*i.e.*, for $j \neq k$, two events $j \in \theta_i$ and $k \in \theta_i$ are independent), and that the probability that i finds any agent $j \neq i$ preferable equals p. Let P_p denote the joint probability distribution of type profile θ given p, and E_p the corresponding expectation. For any i, the marginal distribution of θ_i is given by

$$\Pr(\theta_i \mid \phi) \equiv P_p(\theta_i) = p^{|\theta_i|}(1 - p)^{n - 1 - |\theta_i|},$$

where $|\theta_i|$ denotes the number of elements in θ_i. We will allow the probability p to depend on the number n of agents in the market.

The platform's cost of including any single agent into an assignment equals c, which is assumed to satisfy

$$0 < c < \frac{v}{2}. \tag{2}$$

It follows that social surplus generated when a group of two agents is formed is negative ($= -2c$) when neither agent finds the other agent preferable, but is positive ($\geq v - 2c$) when at least one agent finds the other agent preferable. Denote by $R(\theta \mid g, x)$ the platform's payoff from (g, x) when the type profile is θ:

$$R(\theta \mid g, x) = \sum_{i \in I} x_i(\theta) - |G(\theta)|c.$$

8

The platform's expected payoff $R(g, x)$ under (g, x) is defined accordingly:

$$R(g, x) = E_p\big[R(\theta \mid g, x)\big].$$

4   An Enforceable Mechanism

As mentioned in the Introduction, we address the multiplicity of equilibria in the adoption decisions by considering a version of a revelation-suggestion mechanism proposed by Myerson (1982). Specifically, we suppose that the platform offers the agents an allocation while giving them an option to reject it. The proposed mechanism works over two stages 1 and 2 as follows. In stage 1, the platform determines an allocation $(g(\hat{\theta}, \phi), x(\hat{\theta}, \phi))$ as a function of the reported preference type profile $\hat{\theta}$ and the observed salience type profile ϕ. In stage 2, the platform publicly announces the allocation $(g(\hat{\theta}, \phi), x(\hat{\theta}, \phi))$ as well as $(\hat{\theta}, \phi)$ and then lets the agents simultaneously choose whether or not to accept the allocation.

Formally, the mechanism Γ consists of the allocation rule (g, x), an action set $A_i = \{0, 1\}$ for every agent $(A = \prod_{i \in I} A_i)$, and a collection of outcome functions $(f^g)_{g \in \mathcal{G}}$. A_i is the set of actions available to agent $i \in I$ in the *subscription game* played in stage 2: $a_i = 0$ and $a_i = 1$ represent rejection and acceptance, respectively, of the proposed allocation. When allocation (g, x) with $g = (G_s)_{s \in S}$ is offered and the agents choose action profile $a \in A$ in the subscription game, $f^g : A \to \mathcal{G}$ modifies g by excluding those agents $i \in G$ who have chosen rejection.[7] Accordingly, if θ_i is agent i's true type, his payoff under the action profile a in the subscription game following (g, x) is written in terms of the payoff function defined in (1) by

$$U_i(a, \theta_i \mid g, x, \phi) = \begin{cases} u_i(f^g(a), \phi, \theta_i) - x_i & \text{if } a_i = 1 \text{ and } \phi(i) \in G, \\ -x_i & \text{if } a_i = 1 \text{ and } \phi(i) \notin G, \\ 0 & \text{if } a_i = 0. \end{cases} \quad (3)$$

The first line corresponds to an accepting agent (whose salience type is) included in the assignment, the second line to an accepting agent not included in the assignment, and the third line to a rejecting agent (who makes no transfer payment and gains no connection to other agents).[8] Importantly, the adoption game does not result in the inclusion of an excluded agent to the assignment, or reassignment of any agent to a different group.

---

[7]It is not crucial whether accepting agents are included or not. For example, the assignment $f^g(a)$ may include all accepting agents, or may exclude all agents (so that no group will be formed).

[8]Any agent excluded from the assignment chooses acceptance if and only if his transfer payment is non-positive.

Upon the public announcement of the report profile $\hat{\theta} = (\hat{\theta}_i)_{i \in I}$ by the platform in stage 2, the agents form beliefs about the other agents' preference types. Let $\mu_i : \Theta \to \Delta(\Theta_{-i})$ denote agent i's belief given the public announcement: $\mu_i(\hat{\theta})(\theta_{-i})$ is i's belief that the other agents' true preferences equal θ_{-i} when the report profile equals $\hat{\theta}$.[9] Define $\mu^* = (\mu_i^*)_{i \in I}$ to be the belief corresponding to truthful reporting:

$$\mu_i^*(\hat{\theta})(\hat{\theta}_{-i}) = 1 \text{ for every } \hat{\theta} \in \Theta \text{ and } i \in I.$$

Given that every agent is ex ante symmetric in terms of their salience types, we will in what follows focus on anonymous mechanisms that respect the agents' salience types but ignore their identities. Specifically, we focus on (g, x) such that for every $\phi, \hat{\phi} \in \Phi$ and every $\theta \in \Theta$,

$$g(\theta, \phi) = g(\theta, \hat{\phi}) \text{ and } x_i(\theta, \phi) = x_i(\theta, \hat{\phi}) \text{ for every } i.$$

That is, as long as the preference type profile θ is the same, the mechanism specifies the same assignment of salience types and transfers regardless of who has which salience type. We will then denote (g, x) as a function of θ only, and our analysis in the subsequent sections identifies the agents' salience types with their names: $\phi(i) = i$ for every i.

In stage 1, we suppose that Γ induces truthful reporting as a weakly dominant action *conditional on* the acceptance of the platform's offer $(g(\theta), x(\theta))$ by every agent for every θ in stage 2. Specifically, Γ is *strategy-proof* if

$$u_i(g(\theta_i, \theta_{-i}), \phi, \theta_i) - x_i(\theta_i, \theta_{-i}) \geq u_i(g(\theta_i', \theta_{-i}), \phi, \theta_i) - x_i(\theta_i', \theta_{-i})$$

for every θ_i, θ_i', θ_{-i}, and $i \in I$, and *individually rational* if (g, x) satisfies

$$u_i(g(\theta_i, \theta_{-i}), \phi, \theta_i) - x_i(\theta_i, \theta_{-i}) \geq 0$$

for every θ_i, θ_{-i}, and $i \in I$.

In stage 2, a mechanism Γ is *uniquely acceptable* if for any type profile $\theta \in \Theta$, $a^* = (1, \ldots, 1)$ is the unique Bayes Nash equilibrium of the subscription game when the agents' belief about the distribution of θ is given by μ^*, and agent i's payoff function is given by U_i in (3). In other words, acceptance of the suggested allocation is the uniquely optimal action for every agent i *conditional on* truthful reporting by every agent in stage 1. A mechanism is *uniquely enforceable* if it is strategy-proof, individually rational, and uniquely acceptable.[10]

---

[9]Since reporting is done privately and the types are independent, we ignore the dependence of μ_i on his own type.

[10]Note that unique enforceability does not imply the uniqueness of a PBE in the two-stage game.

10

By strategy-proofness, no unilateral deviation that involves misreporting in stage 1 and acceptance in stage 2 is profitable. No unilateral deviation that involves rejection in stage 2 is profitable either since any such deviation yields zero, whereas truthful reporting and acceptance yield at least zero by individual rationality. We hence have the following result.

Proposition 1. *If Γ is uniquely enforceable, then it implements (g, x) in PBE of the two-stage game.*

5 Complete Information

We begin our analysis with the benchmark case where the platform has complete information about the agents' preference type profile θ, or equivalently, the underlying preference network $D(\theta)$. In this case, we only require the mechanism Γ to be IR and uniquely acceptable. Define $R^*(\theta)$ to be the supremum of the platform's payoff from such mechanisms:

$$R^*(\theta) = \sup\{R(\theta \mid g, x) : \Gamma \text{ is IR and uniquely acceptable}\},$$

and $R^* = E_p[R^*(\theta)]$.

Consider the components of the preference network $D(\theta)$ consisting of two or more agents, and index them by s. Let $S(\theta)$ be the set of such indices. For each component $C_s(\theta) \subset I$ of $D(\theta)$, take any subset F of $C_s(\theta)$ and let

$$C_{\rightsquigarrow F}(\theta) = \cup_{i \in F} C_{\rightsquigarrow i}(\theta)$$

be the set of agents who are strongly preference-connected to some $i \in F$. Define $(F_s(\theta), Y_s(\theta))$ to be a solution to the maximization problem:

$$\max_{(F,Y)} |Y|(v - c) - |F|c \quad \text{subject to} \quad \begin{cases} F, Y \subset C_s(\theta), \\ Y = C_{\rightsquigarrow F}(\theta) \setminus F. \end{cases} \tag{4}$$

(4) has a solution since $C_s(\theta)$ is finite, and gives the maximal payoff that the platform can generate from this component. The intuition is as follows: The objective function corresponds to the platform's payoff when it forms a group $F \cup Y$ by offering free subscription to agents in F, but charging v to agents in Y. Since Y is chosen so that its members are strongly preference-connected to some agent in F, agents in Y are indeed willing to pay up to v for subscription into an assignment if it also includes F. The platform then maximizes its payoff by taking Y as large as possible (since $v - c > 0$), while taking F as small as possible (since $c > 0$). Note that the maximized payoff is strictly positive for every component of two or

11

more agents since $2c < v$. When the component C is as described in Figure 1, for example, (4) has the solution $F = \{3, 6\}$ and $Y = \{1, 2, 4, 7, 8, 9, 10\}$, and its value equals $7(v - c) - 2c = 7v - 9c$.[11]
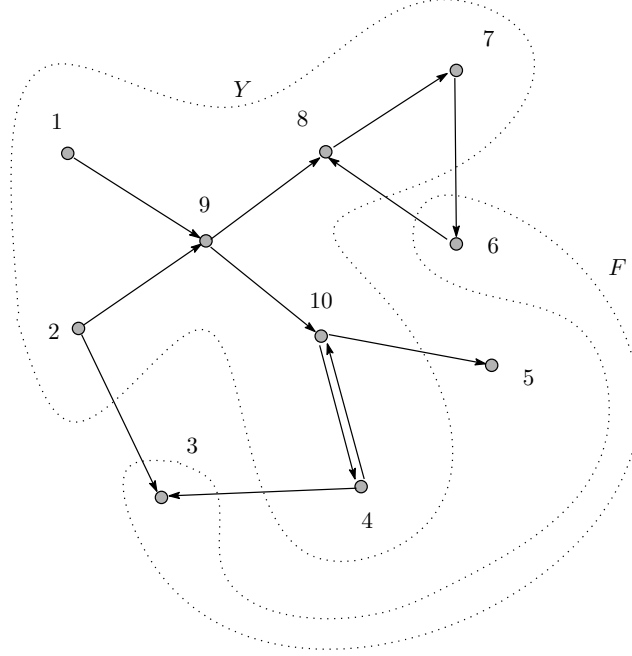

Figure 1: Optimal Mechanism under Complete Information

We next show that for any $\varepsilon > 0$, there exists a mechanism Γ such that $R(\theta \mid g, x) = \sum_{s \in S(\theta)} \{|Y_s(\theta)|(v - c) - |F_s(\theta)|c\} - \varepsilon$ for the solution $(F_s(\theta), Y_s(\theta))$ to (4) $(s \in S(\theta))$. First, for each θ, let group $G_s(\theta)$ be defined by $G_s(\theta) = F_s(\theta) \cup Y_s(\theta)$, and assignment $g(\theta)$ by $g(\theta) = (G_s(\theta))_{s \in S(\theta)}$. In line with the intuition provided above, the transfer is specified so that agents in $F_s(\theta)$ are offered a small subsidy whereas agents in $Y_s(\theta)$ are charged a fee close to their full valuation v. Formally, let

$$F(\theta) = \cup_{s \in S(\theta)} F_s(\theta), \quad \text{and} \quad Y(\theta) = \cup_{s \in S(\theta)} Y_s(\theta),$$

and for $\varepsilon \in (0, v)$, define

$$x_i(\theta) = \begin{cases} -\frac{\varepsilon}{n} & \text{if } i \notin Y(\theta),^{12} \\ v - \frac{\varepsilon}{n} & \text{if } i \in Y(\theta). \end{cases} \tag{5}$$

---

[11] Recall that an agent's salience type is identified with his index. Note that this is not the unique solution to the maximization problem. The same value can be achieved when, for example, $F = \{3, 8\}$ and $Y = \{1, 2, 4, 7, 8, 9, 10\}$.

Proposition 2. *Suppose that the platform has complete information about θ. Then*

$$R^*(\theta) = \sum_{s \in S(\theta)} \{|Y_s(\theta)|(v-c) - |F_s(\theta)|c\}.$$

Furthermore, Γ with (g,x) defined in (5) is IR and uniquely acceptable, and satisfies $R(\theta \mid g,x) = R^*(\theta) - \varepsilon$ for every θ.

In the subscription stage, acceptance is a strictly dominant action for every agent in F, and is an iteratively strictly dominant action for agents in Y. The number of iteration required for each agent in Y to find out that acceptance is an optimal action equals the length of the shortest directed path that connects him with agents in F.

6   Incomplete Information

We now return to the incomplete information environment in which the platform cannot directly observe the realization of θ. With incomplete information, the first-best payoff R^* is not achievable as shown in the following proposition.

Proposition 3. *For any market size n and probability p, there exists $\kappa \equiv \kappa(n,p) > 0$ such that if Γ is a uniquely enforceable mechanism, then the platform's expected payoff under (g,x) is bounded away from the optimal level by κ: $R(g,x) \le R^* - \kappa$.*

The proposition is based on the observation that a mechanism is uniquely acceptable only if it offers a negative subscription fee to at least one agent. We should emphasize that κ, which is interpreted as a lower bound for the agents' informational rents, depends on both p and n. For a fixed market size n, the extreme values of the probability p eliminate those information rents: On the one hand, as p approaches zero, all agents become isolated with probability close to one, and hence the platform's payoff from any IR mechanism approaches zero. On the other hand, as p approaches one, the underlying preference network becomes complete (*i.e.,* there is a two-way link between every pair of agents) with probability close to one, and the mechanism described in the following subsection that forms a single group will become optimal. Our primary interest hence is on what happens for the less extreme values of p as the market size n grows large. As seen below, different mechanisms entail different allocations of the informational rents among the agents.

---

[12]In this specification, note that agents not included in the assignment are also offered a small subsidy.

6.1 Single Group Mechanism

We first consider the case where the expected number of agents $(n-1)p$ in any agent's preference set θ_i (slowly) increases with the market size n. The expected total number of links $n(n-1)p$ in the underlying network $D(\theta)$ hence grows (a little) faster than linearly with n.

A single group mechanism forms at most one group regardless of the underlying preference network. Specifically, for some predetermined agent i_1, the single platform is formed if the set of agents who are strongly preference-connected to i_1 is non-empty, and the group includes i_1 and all those agents who are strongly preference-connected to him. No group is formed if no agent is strongly preference-connected to i_1. Define $Y(\theta) \equiv C_{\rightsquigarrow i_1}(\theta)$ to be the set of agents strongly preference-connected to i_1. Formally, Γ is a *single group mechanism based on i_1* if there exists $\varepsilon \in (0, v)$ such that for every θ: $S(\theta)$ is a singleton, and

1) if $Y(\theta) = \emptyset$, then $G(\theta) = \emptyset$ and $x_i(\theta) = -\frac{\varepsilon}{n}$ for every i;

2) if $Y(\theta) \neq \emptyset$, then $G(\theta) = \{i_1\} \cup Y(\theta)$, and

$$x_i(\theta) = \begin{cases} v - \frac{\varepsilon}{n} & \text{if } i \in Y(\theta), \\ -\frac{\varepsilon}{n} & \text{if } i \notin Y(\theta). \end{cases}$$

Figure 2 illustrates the construction of a single group mechanism based on $i_1 = 6$. It is important that the specification of i_1 cannot be contingent on the realization of θ. Note also that only agent i_1 may possibly enjoy informational rents close to v under this mechanism.

Proposition 4. *A single group mechanism Γ is uniquely enforceable.*

For a fixed market size n, the single group mechanism is inefficient unless p is very large since with non-negligible probability, only a few agents are strongly preference-connected to i_1, and hence the platform formed is small compared with the total market size. In Figure 2, for example, a single group mechanism based on $i_1 = 1$ or 2 will create no group. As seen below, however, this problem becomes less serious as the market size grows as far as the probability p does not vanish "very fast" with n. Formally, let Γ be the single group mechanism specified above, and denote by $\Theta^\varepsilon(n)$ the set of type profiles θ at which the expected revenue under the single group mechanism Γ described above is within ε of the optimal level:

$$\Theta^\varepsilon(n) = \{\theta : R(\theta \mid g, x) > R^*(\theta) - \varepsilon\}.$$
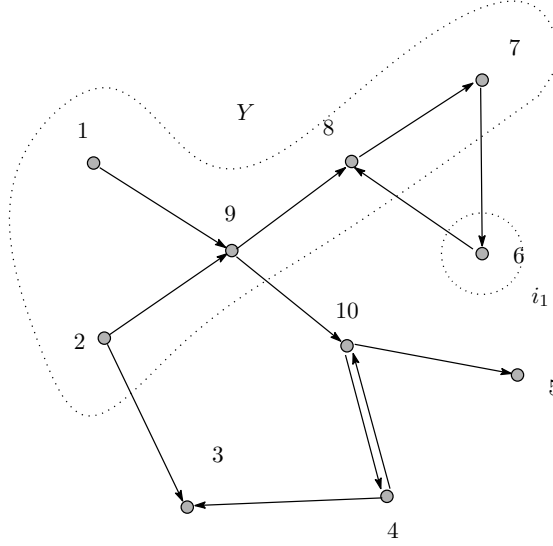
14

Figure 2: Construction of a Single Network Mechanism when $i_1 = 6$

Proposition 5. *Let $\gamma > 0$ and suppose that $p(n) \geq \frac{(1+\gamma)\log n}{n}$ for every large n. Then the single group mechanism Γ is asymptotically optimal in the sense that for every $\varepsilon > 0$,*

$$\lim_{n \to \infty} P_p(\theta \in \Theta^\varepsilon(n)) = 1.$$

The proposition builds on a theory of random graphs. A directed graph D is *Hamiltonian* if there exists a directed closed trail that goes through every node of D exactly once as in Figure 3. If $D(\theta)$ is Hamiltonian, then every agent is strongly preference-connected to every other agent, and hence to i_1. It follows that if i_1 subscribes, then all other agents are willing to pay up to v for subscription. It is known that when the probability $p(n)$ satisfies the condition stated in Proposition 5, the probability that $D(\theta)$ is Hamiltonian approaches one as the market becomes large.

Proposition 5 includes as a special case where the probability p of each link is independent of the market size.

Corollary 1. *Suppose $p > 0$ is a constant. Then the single group mechanism Γ is asymptotically optimal as $n \to \infty$.*
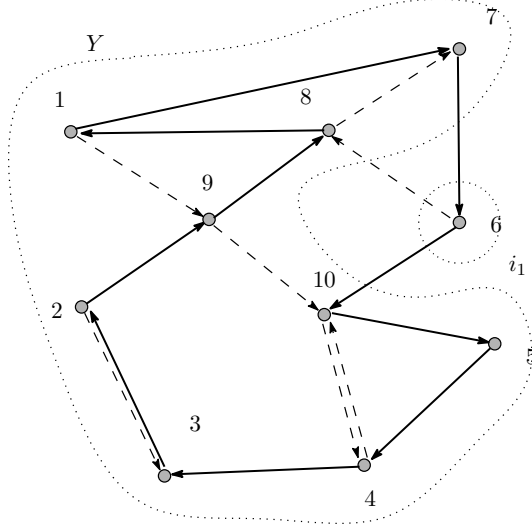
15

Figure 3: Hamiltonian graph
The directed path indicated by the solid line visits every agent just once and
returns to the original agent.

6.2   Multiple Group Mechanism

We next consider the case where the expected number of agents $(n-1)p$ in any
agent's preference set θ_i vanishes as the market size $n \to \infty$. However, $n(n-1)p$,
which is the expected total number of links in the underlying network $D(\theta)$, may
still increase with n.

   Given $i \in I$ and $Y \subset I$, we say that j is *strongly preference-connected to i within*
Y, and denote $j \rightsquigarrow_Y i$, if there is a directed path $j \to i_1 \to \cdots \to i_k \to i$ from j to i
such that all intermediate agents $i_1, \ldots, i_k$ on the path belong to Y. Given any F,
$Y \subset I$ with $F \cap Y = \emptyset$, define also

$$C_{\rightsquigarrow F}(Y, \theta) = \{j \notin F : j \rightsquigarrow_Y i \text{ for some } i \in F\}$$

to be the set of agents outside F who are strongly preference-connected to some i
in F within Y in the network $D(\theta)$. If Y satisfies $Y \subset C_{\rightsquigarrow F}(Y, \theta)$, then conditional
on subscription by agents in F, every agent $i \in Y$ has acceptance as an iteratively
dominant action when a subscription fee is less than v.

   For any θ, consider a collection $\mathcal{F}(\theta)$ of subsets F of I such that either

1) $F = \{i\}$ for some i such that $\theta_i = \emptyset$ and $C_{\rightsquigarrow i}(\theta) \neq \emptyset$, or

2) $|F| \geq 2$ and F is a strong component.

16

F is a set of candidates who are offered free subscription. The only complication is that there may be an overlap of agents who find preferable two different free agents i and j. In such a case, agent j is redundant in attracting those agents in the overlap. In this sense, the optimal assignment must partition the set of agents who find free agents preferable, and eliminate any redundant free agents. Formally, let $\bar{S}(\theta)$ be set of indices of elements of $\mathcal{F}(\theta)$, and denote its generic element by F_s. Let

$$Y = \left(\cup_{s \in \bar{S}(\theta)} C_{\leadsto F_s}(\theta) \right) \setminus \left(\cup_{s \in \bar{S}(\theta)} F_s \right).$$

Y is the set of agents who are outside $\cup_{s \in \bar{S}(\theta)} F_s$ but are strongly preference-connected to some F_s. If $Y \neq \emptyset$, partition Y into disjoint subsets $(Y_s)_s$ (allowing $Y_s = \emptyset$ for some s) such that

$$Y_s \subset C_{\leadsto F_s}(Y_s, \theta).$$

That is, each $j \in Y_s$ is strongly preference-connected within Y_s to F_s so that they are willing to pay up to v if agents in F_s also subscribe. If there are multiple such partitions, take the one that minimizes the number of free agents included, or equivalently, the number of elements in the set

$$\{s : F_s = \{i\} \text{ and } Y_s \neq \emptyset\}.$$

Let now $S(\theta) \subset \bar{S}(\theta)$ be the subset of indices such that

$$S(\theta) = \{s \in \bar{S}(\theta) : F_s = \{i\} \text{ and } Y_s \neq \emptyset\} \cup \{s \in \bar{S}(\theta) : |F_s| \geq 2\}.$$

Let $G_s(\theta) = F_s(\theta) \cup Y_s(\theta)$ for each $s \in S(\theta)$, and

$$F(\theta) = \cup_{s \in S(\theta)} F_s(\theta), \quad \text{and} \quad Y(\theta) = \cup_{s \in S(\theta)} Y_s(\theta).$$

Consider a mechanism such that for each θ, its groups are given by $G_s(\theta)$ for $s \in S(\theta)$, and agents in $Y(\theta)$ are charged subscription fees close to v, while other agents are offered small subsidies. Formally, for $\varepsilon \in (0, v)$, define a multiple group mechanism Γ by

$$g(\theta) = (G_s(\theta))_{s \in S(\theta)}, \quad \text{and} \quad x_i(\theta) = \begin{cases} v - \frac{\varepsilon}{n} & \text{if } i \in Y(\theta), \\ -\frac{\varepsilon}{n} & \text{otherwise.} \end{cases}$$

Figure 4 illustrates the construction of a multiple group mechanism, where $F = F_1 \cup F_2$.[13]

---

[13]Note that agents 3 and 4 are not included in the assignment since they are redundant in terms of the agents they attract.
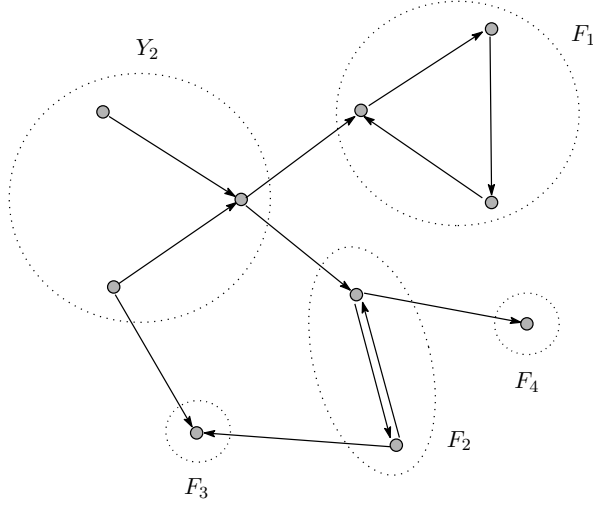
Figure 4: Construction of a Multiple Group Mechanism

Proposition 6. *A multiple group mechanism* Γ *is uniquely enforceable.*

As seen, the multiple group mechanism leaves the informational rents to a strong component of agents, and hence can be inefficient when the probability p is high. On the other hand, it is almost efficient when there is no strong component as seen in the following lemma.

Lemma 1. *Let* Γ *be a multiple group mechanism. If* $D(\theta)$ *contains no strong component, then* $R(\theta \mid g, x) = R^*(\theta) - \varepsilon$.

When the expected cardinality $(n-1)p$ of the type set θ_i of each agent becomes small as n grows large, the probability that there exists a strong component becomes small. It then follows from Lemma 1 that in such a market, the multiple group mechanism becomes almost efficient. As before, let $\Theta^\varepsilon(n)$ be the set of type profiles θ at which the expected revenue under the multiple group mechanism Γ described above is within ε of the optimal level:

$$\Theta^\varepsilon(n) = \{\theta : \; R(\theta \mid g, x) > R^*(\theta) - \varepsilon\}.$$

Proposition 7. *Suppose that* $p(n) = \frac{\tau(n)}{n}$ *for every n for some function* $\tau : \mathbf{N} \to \mathbf{R}_{++}$ *satisfying* $\lim_{n\to\infty} \tau(n) = 0$. *Then the multiple group mechanism is asymptotically optimal in the sense that for every* $\varepsilon > 0$,

$$\lim_{n\to\infty} P_p(\theta \in \Theta^\varepsilon(n)) = 1.$$

18

The key insight behind this result also comes from a result on random graphs. Let $U(\theta)$ be the undirected graph obtained from $D(\theta)$ by ignoring the direction of its links. When the probability $p(n)$ satisfies the condition of Proposition 7, it can be shown that the probability that there is a cycle of three or more agents in $U(\theta)$ approaches zero as $n \to \infty$. If there is no such cycle in $U(\theta)$, then there is no directed cycle of three or more agents in $D(\theta)$. Furthermore, it can be readily shown that the probability that $D(\theta)$ contains directed cycles of just two agents also approaches zero as $n \to \infty$. It follows that the probability that there is a strong component in $D(\theta)$ approaches zero as $n \to \infty$, and hence a small subsidy needs to be offered only to those agents with $\theta_i = \emptyset$. These agents, however, must also be offered small subsidies under complete information, and hence there is no loss in the platform's payoff compared with the complete information benchmark.

7   Conclusion

We study platform design when a platform offers connections to agents when it exercises a monopoly power based on its proprietary access to information about the agents' heterogeneous salience types. The platform collects information about the agents' preferences and proposes an allocation that consists of groups of connected agents and transfer payments. Our focus is on uniquely enforceable mechanisms that work over two stages and make truthful reporting of private preference types weakly dominant in the first stage, and acceptance of the proposed allocation a uniquely optimal action for every agent in the second stage. The latter requirement addresses the multiplicity of equilibria in the adoption decisions in the presence of network externalities. We show that in the limit as the number of market participants grows large, there exist uniquely enforceable mechanisms that extract full informational surplus from the agents. These mechanisms are classified by the number of groups of connected agents they create. Specifically, the classification is based on how the expected number $(n-1)p$ of salience types that each agent finds preferable varies with the market size n. In markets where each agent finds a relatively large number of salience types preferable, an optimal mechanism forms one group of connected agents and offers a small subsidy just to one agent. In contrast, in markets where each agent finds a relatively small number of salience types preferable, an optimal mechanism forms many small groups of connected agents.

The following are some open questions. First, one important case not covered by our analysis is when the expected size of the preference set θ_i of each agent remains constant regardless of n, i.e., $p(n) = \frac{k}{n}$ so that $(n-1)p \to k$ for some constant k. This case, however, corresponds to the threshold in the existence of cycles in the large

19

random graphs, and the asymptotic behavior of the preference network is difficult to analyze. Second, our market is one-sided in the sense that every agent is potentially interested in a match with any other agent. On the other hand, platform design is often studied in a two-sided market, where agents in one side are only interested to be matched with some of the agents on the other side as in the case of sellers and buyers of some good. While the present framework readily extends to accommodate models of two-sided markets, extending the analysis requires a different technique. Third, we have assumed for simplicity that the value of a platform can take only two values 0 and $v > 0$ for each agent. As mentioned earlier, models in the matching literature assume more general value specifications, and it would be interesting to study their implications in our context. For example, one natural specification is one where the value of a group increases monotonically with the number of agents in it with preferable salience types.

Appendix

Proof of Proposition 2. We first show that

$$R^*(\theta) \leq \sum_{s \in S(\theta)} \{|Y_s(\theta)|(v - c) - |F_s(\theta)|c\}.$$

Take any component $C_s(\theta)$ ($s \in S(\theta)$). Suppose to the contrary that there exists an IR and uniquely acceptable mechanism with the allocation rule $(\hat{g}, \hat{x})$ that yields the payoff $R_s(\theta \mid \hat{g}, \hat{x}) > |Y_s(\theta)|(v - c) - |F_s(\theta)|c$. Let $\hat{Y} = \{i \in C_s(\theta) : \hat{x}_i(\theta) > 0\}$ be the set of agents in $C_s(\theta)$ who are charged strictly positive subscription fees, and $\hat{F} = \{i \in C_s(\theta) : \hat{x}_i(\theta) \leq 0\}$ be the set of agents in $C_s(\theta)$ who are offered subsidies for subscription. By assumption, there exists $i_1 \in \hat{Y}$ who is not strongly preference-connected to $\hat{F}$. Let $J_1 = \{i_1\}$, and let $J_2 \subset \hat{Y}$ be the set of agents to whom i_1 is strongly preference-connected: $J_2 = \{j \in \hat{Y} : i_1 \rightsquigarrow j\}$. Since i_1 is charged a positive subscription fee, IR implies that $J_2 \neq \emptyset$, and furthermore, since i_1 is not strongly preference-connected to $\hat{F}$, no $j \in J_2$ is strongly preference-connected to $\hat{F}$. Let then $J_3 \subset \hat{Y}$ be the set of agents j to whom agents in J_2 are strongly preference-connected. In the same way, we can iteratively construct a sequence $J_4, J_5, \ldots$ of subsets of $\hat{Y}$ so that no agent in those subsets is strongly preference-connected to $\hat{F}$. Since $\hat{Y}$ is finite, however, we will have $J_{k+1} \subset \cup_{\ell=1}^{k} J_\ell$ for some k. We then have a contradiction to uniquely acceptable since then for agents i in the set $\cup_{\ell=1}^{k} J_\ell$, no subscription $a_i = 0$ is a Nash equilibrium action since they are all charged positive subscription fees.

The mechanism with allocation rule (g, x) defined in (5) is clearly IR and also uniquely acceptable: For $i \in F(\theta)$, $x_i(\theta) < 0$ so that subscription $a_i = 1$ is a

dominant action. For $j \in Y_s(\theta)$, $a_j = 1$ is an iteratively dominant action since $Y_s(\theta)$ consists of all agents who are strongly preference-connected to some $i \in F_s(\theta)$: For any $j \in Y_s(\theta)$, either $i \in \theta_j$ for some $i \in F_s(\theta)$ or there exists $k \in Y_s(\theta)$ with $k \in C_{\rightsquigarrow F}(\theta)$. Rejection is dominated in the second round of the iterative elimination procedure in the first case, whereas it is iteratively dominated for j in one round after it is dominated for k in the second case. Finally, the platform's payoff from $C_s(\theta)$ under (g, x) equals

$$R(\theta \mid g, x) = \sum_{s \in S(\theta)} \{|Y_s(\theta)|(v - c) - |F_s(\theta)|c\} - \varepsilon.$$

Since ε is arbitrary, $R^*(\theta) \geq \sum_{s \in S(\theta)} \{|Y_s(\theta)|(v - c) - |F_s(\theta)|c\}$. We hence obtain the stated conclusion. $\qquad \square$

Proof of Proposition 3. Let $\delta = \left(\min\{p, 1 - p\}\right)^{\binom{n}{2}} > 0$ so that $P(\theta) \geq \delta$ for every θ. Consider the set Θ^{sc} of type profiles θ such that $D(\theta)$ is strongly preference-connected. Define

$$\kappa = \delta \left|\Theta^{\text{sc}}\right| (n - 1)(v - c) > 0.$$

Fix any uniquely enforceable mechanism with allocation rule (g, x), and write $R(\theta) = R(\theta \mid g, x)$ for simplicity. For any $\theta \in \Theta^{\text{sc}}$, it is optimal to include every agent in the assignment. Furthermore, one agent must be offered free subscription since otherwise, there would be a NE of the adoption game in which no agent accepts. Hence, the optimal revenue under complete information is given by

$$R^*(\theta) = (n - 1)v - nc.$$

On the other hand, for any θ, let $J(\theta) = \{j \in I : x_j(\theta) > 0\}$ be the set of agents who are charged positive subscription fees under (g, x). Since Γ is IR, $J(\theta) \subset G(\theta)$, the set of agents included in the assignment. Since Γ is IR and uniquely acceptable, we must again have one agent included in the assignment with a non-positive subscription fee so that the platform's revenue under θ satisfies

$$R(\theta) \leq |J(\theta)|v - (|J(\theta)| + 1)c.$$

It follows that for any $\theta \in \Theta^{\text{sc}}$,

$$R^*(\theta) - R(\theta) \geq (n - |J(\theta)| - 1)(v - c). \tag{6}$$

Take any $\theta \in \Theta^{\text{sc}}$ and $j \in J(\theta)$, and consider now a profile $(\emptyset, \theta_{-j})$ where agent j's type is replaced by $\theta'_j = \emptyset$. Clearly, $(\emptyset, \theta_{-j}) \notin \Theta^{\text{sc}}$ and $R^*(\emptyset, \theta_{-j}) = (n - 1)v - nc$. We must have either $j \notin G(\emptyset, \theta_{-j})$ or $\theta_j \cap G(\emptyset, \theta_{-j}) = \emptyset$ since otherwise, agent j

would report $\theta'_j = \emptyset$ when his true type is θ_j, and since $x_j(\emptyset, \theta_{-j}) \leq 0$ by IR, his utility from misreporting would be $v - x_j(\emptyset, \theta_{-j}) \geq v > v - x_j(\theta)$, which is his payoff when he reports his type truthfully. It follows that $G(\emptyset, \theta_{-j})$ consists of at most $n - 1$ agents, and hence the fact that Γ is IR and uniquely acceptable implies that

$$R(\emptyset, \theta_{-j}) \leq (n-2)v - (n-1)c.$$

This further implies that

$$R^*(\emptyset, \theta_{-j}) - R(\emptyset, \theta_{-j}) \geq v - c. \tag{7}$$

We then have from (6) and (7),

$$
\begin{aligned}
R^* - R &= \sum_{\theta} P(\theta) \left\{ R^*(\theta) - R(\theta) \right\} \\
&\geq \sum_{\theta \in \Theta^{\mathrm{sc}}} \left[P(\theta) \, (n - |J(\theta)| - 1)(v - c) + \sum_{j \in J(\theta)} P(\emptyset, \theta_{-j}) \, (v - c) \right] \\
&= \delta \sum_{\theta \in \Theta^{\mathrm{sc}}} \left[(n - |J(\theta)| - 1)(v - c) + |J(\theta)| \, (v - c) \right] \\
&= \delta \, |\Theta^{\mathrm{sc}}| \, (n - 1)(v - c).
\end{aligned}
$$

$\square$

Proof of Proposition 4. Γ is clearly individually rational. To see that it is strategy-proof, note first that agent i_1 has no incentive to misreport his type since the group formed or his transfer is independent of his report. Take any $i \neq i_1$.

1) If i is strongly preference-connected to i_1 and hence belongs to the group ($i \in Y(\theta)$) under θ, then his payoff is given by

$$u_i(g(\theta), \theta_i) - x_i(\theta) = v - (v - \frac{\varepsilon}{n}) = \frac{\varepsilon}{n}.$$

If i reports θ'_i such that i still belongs to the group under (θ'_i, θ_{-i}), then his payoff is unchanged. If i reports θ'_i such that i is not strongly preference-connected to i_1, then he is not offered an assignment and his payoff equals $\frac{\varepsilon}{n}$ which is the transfer made to him. It follows that i has no incentive to misreport his type.

2) If i is not strongly preference-connected to i_1 and hence does not belong to the group under θ, then his payoff equals the transfer made to him: $\frac{\varepsilon}{n}$. If he reports θ'_i that makes him still not strongly preference-connected to i_1 under (θ'_i, θ_{-i}), then his payoff is unchanged. On the other hand, if he reports θ'_i

22

which makes him strongly preference-connected to i_1, then his payoff equals $v - (v - \frac{\varepsilon}{n}) = \frac{\varepsilon}{n}$ if the assignment $g(\theta_i', \theta_{-i})$ also includes some agent in θ_i, and $-(v - \frac{\varepsilon}{n})$ otherwise. In every case, i has no incentive to misreport his type.

We have hence shown that Γ is strategy-proof. To see that Γ also is uniquely acceptable, note that $a_i = 1$ is a strictly dominant action for $i = i_1$, and $a_i = 1$ is iteratively strictly dominant for any $i \in Y(\theta)$. $\qquad \square$

Proof of Proposition 5. Let $\Theta^H(n)$ be the set of type profiles $\theta = (\theta_1, \dots, \theta_n)$ such that the directed graph $D(\theta)$ has a Hamiltonian cycle: There exists a directed closed path that goes through every agent. For any $\theta \in \Theta^H(n)$, hence, every agent $j \neq i$ is strongly preference-connected to i, and hence the group formed by the single group mechanism covers all the agents. This further implies that $\Theta^H(n) \subset \Theta^\varepsilon(n)$. On the other hand, when the probability $p(n)$ satisfies the stated condition, the probability of $\Theta^H(n)$ approaches 1 as $n \to \infty$ (McDiarmid 1981, Theorem 4.7).[14] It follows that

$$\lim_{n \to \infty} P_p(\theta \in \Theta^\varepsilon(n)) \geq \lim_{n \to \infty} P_p(\theta \in \Theta^H(n)) = 1.$$

$\qquad \square$

Proof of Proposition 6. 1) Suppose first that agent $i \in F_s(\theta)$ for some $s \in S(\theta)$. If $\theta_i \neq \emptyset$, then $F_s(\theta)$ is a strong component of $D(\theta)$ by definition, and hence $u_i(g(\theta), \theta_i) = v$. If $\theta_i = \emptyset$, then $u_i(g(\theta), \theta_i) = 0$. Since $x_i(\theta) = -\frac{\varepsilon}{n}$ in both cases, i's payoff under θ equals

$$u_i(g(\theta), \theta_i) - x_i(\theta) = \begin{cases} \frac{\varepsilon}{n} & \text{if } \theta_i = \emptyset, \\ v + \frac{\varepsilon}{n} & \text{otherwise.} \end{cases}$$

If i reports θ_i' such that $i \in F_{s'}(\theta_i', \theta_{-i})$ for some $s' \in S(\theta_i', \theta_{-i})$, then his payoff equals

$$u_i(g(\theta_i', \theta_{-i}), \theta_i) - x_i(\theta_i', \theta_{-i}) = \begin{cases} \frac{\varepsilon}{n} & \text{if } \theta_i \cap G_{s'}(\theta_i', \theta_{-i}) = \emptyset, \\ v + \frac{\varepsilon}{n} & \text{otherwise.} \end{cases}$$

Since $\theta_i = \emptyset$ implies $\theta_i \cap G_{s'}(\theta_i', \theta_{-i}) = \emptyset$, it follows that reporting θ_i' is not a profitable deviation whether $\theta_i = \emptyset$ or not. If i reports θ_i' such that $i \in Y_{s'}(\theta_i', \theta_{-i})$ for some $s' \in S(\theta_i', \theta_{-i})$, then i's payoff equals

$$u_i(g(\theta_i', \theta_{-i}), \theta_i) - x_i(\theta_i', \theta_{-i}) = \begin{cases} 0 - (v - \frac{\varepsilon}{n}) & \text{if } \theta_i \cap G_{s'}(\theta_i', \theta_{-i}) = \emptyset, \\ v - (v - \frac{\varepsilon}{n}) & \text{otherwise.} \end{cases} \tag{8}$$

---

[14] See also Bollobás (2001, p217).

23

If i reports θ_i' such that $i \notin G(\theta_i', \theta_{-i})$, then his payoff equals $\frac{\varepsilon}{n}$. It follows that reporting θ_i' is not a profitable deviation in either case.

2) Suppose that agent $i \in Y_s(\theta)$ for some $s \in S(\theta)$. Since then $x_i(\theta) = v - \frac{\varepsilon}{n}$ and $u_i(g(\theta_i, \theta_{-i}), \theta_i) = v$, i's payoff under θ equals

$$u_i(g(\theta), \theta_i) - x_i(\theta) = v - (v - \frac{\varepsilon}{n}) = \frac{\varepsilon}{n}.$$

Furthermore, no $j \in \theta_i$ is strongly preference-connected to i under θ_{-i} since otherwise, i and j would be part of a strong component and hence $i \in F_s(\theta)$ would hold. If i reports θ_i' such that $i \in F_{s'}(\theta_i', \theta_{-i})$ for some $s' \in S(\theta_i', \theta_{-i})$, then $\theta_i \cap G_{s'}(\theta_i', \theta_{-i}) = \emptyset$ since by definition, $G_{s'}(\theta_i', \theta_{-i})$ contains only those agents who are strongly preference-connected to i under θ_{-i}, but as noted above, no agent $j \in \theta_i$ is strongly preference-connected to i under θ_{-i}. Hence, $u_i(g(\theta_i', \theta_{-i}), \theta_i) = 0$. It follows that his payoff equals

$$u_i(g(\theta_i', \theta_{-i}), \theta_i) - x_i(\theta_i', \theta_{-i}) = \frac{\varepsilon}{n}.$$

If i reports θ_i' such that $i \in Y_{s'}(\theta_i', \theta_{-i})$ for some $s' \in S(\theta_i', \theta_{-i})$, then i's payoff is given by (8), and hence $\leq \frac{\varepsilon}{n}$. If i reports θ_i' such that $i \notin G(\theta_i', \theta_{-i})$, then his payoff equals $\frac{\varepsilon}{n}$. It follows that misreporting is not a profitable deviation in any of these cases.

3) Suppose that agent $i \notin G(\theta)$. In this case, $\theta_i = \emptyset$ and $x_i(\theta) = -\frac{\varepsilon}{n}$ so that $u_i(g(\theta_i, \theta_{-i}), \theta_i) - x_i(\theta_i) = \frac{\varepsilon}{n}$. If i reports θ_i' such that $i \in G_{s'}(\theta_i', \theta_{-i})$ for some $s' \in S'(\theta)$, then $u_i(g(\theta_i', \theta_{-i}), \theta_i) = 0$ and $x_i(\theta_i', \theta_{-i}) = v - \frac{\varepsilon}{n}$ or $-\frac{\varepsilon}{n}$ depending on whether i belongs to a strong component of $D(\theta_i', \theta_{-i})$ or not. In either case, $u_i(g(\theta_i', \theta_{-i}), \theta_i) - x_i(\theta_i', \theta_{-i}) \leq \frac{\varepsilon}{n}$ so that misreporting is not a profitable deviation.

The same argument as in the proof of Proposition 2 shows that Γ is uniquely acceptable. □

Proof of Lemma 1. Let Γ be a multiple group mechanism, and take any component C of $D(\theta)$ with two or more agents. Define $C_F = \{i \in C : \theta_i = \emptyset\}$ and $C_Y = \{i \in C : \theta_i \neq \emptyset\}$. Since $D(\theta)$ contains no strong component, $F(\theta)$ and $Y(\theta)$ in the definition of a multiple group mechanism satisfy $F(\theta) \cap C \subset C_F$ and $Y(\theta) \cap C = C_Y$. Denote $F(\theta) \cap C = \{h_1, \ldots, h_M\}$, and let $(Y_m(\theta))_{m=1}^M$ be the corresponding partition of C_Y according to g. By definition, if $(F_k, Y_k)_{k=1}^K$ is any collection such that $(Y_k)_{k=1}^K$ is a partition of C_Y, $(F_k)_{k=1}^K$ is a collection of distinct singleton subsets of C_F, and $Y_k \subset C_{\leadsto F_k}(Y_k, \theta)$ for every k, then

$$M \leq K. \tag{9}$$

Since $x_i(\theta) = -\frac{\varepsilon}{n}$ for every $i \in C_F$ and $x_i(\theta) = v - \frac{\varepsilon}{n}$ for every $i \in C_Y$, the platform's payoff under (g, x) from the component C is given by

$$|C_Y|\left(v - \frac{\varepsilon}{n} - c\right) - |C_F \cap F(\theta)|\left(\frac{\varepsilon}{n} + c\right) - |C_F \setminus F(\theta)|\frac{\varepsilon}{n}$$
$$= |C_Y|(v - c) - |C_F \cap F(\theta)|c - |C|\frac{\varepsilon}{n}.$$

On the other hand, we claim that the first-best payoff for the platform under complete information from the component C is given by

$$|C_Y|(v - c) - |C_F \cap F(\theta)|c.$$

Let (F^*, Y^*) be the pair of disjoint subsets of C that solves (4). Since $\theta_i \neq \emptyset$ for every $i \in Y^*$ by definition, we have $Y^* \subset C_Y$. If $Y^* \subsetneq C_Y$, let $Y' = Y^* \cup \{i\}$ for $i \in C_Y \setminus Y^*$ and $F' = F^* \cup \{j\}$ for $j \in C_F$ such that $j \in \theta_i$. Since $v - 2c > 0$, we then have

$$|Y'|(v - c) - |F'|c \geq (|Y^*| + 1)(v - c) - (|F^*| + 1)c$$
$$= |Y^*|(v - c) - |F^*|c + v - 2c$$
$$> |Y^*|(v - c) - |F^*|c,$$

which contradicts our assumption that (F^*, Y^*) maximizes (4). We hence have $Y^* = C_Y = C \cap Y(\theta)$. Denote $F^* = \{i_1, \ldots, i_K\}$ and note that there exists a partition $(Y_k^*)_{k=1}^K$ of C_Y such that $Y_k^* \subset C_{\leadsto i_k}(Y_k^*, \theta)$ for every k.[15] Indeed, if we define $Y_1^* = C_{\leadsto i_1}(\theta)$, and $Y_k^* = C_{\leadsto i_k}(\theta) \setminus (\cup_{\ell=1}^{k-1} Y_\ell^*)$ for $k = 2, \ldots, K$, then $Y_k^* \subset C_{\leadsto i_k}(Y_k^*, \theta)$ for every k.[16] Since the collection $(\{i_k\}, Y_k^*)_{k=1}^K$ is such that $(Y_k^*)_{k=1}^K$ is a partition of C_Y, and $Y_k^* \subset C_{\leadsto}(Y_k^*, \theta)$ for every k, we have by (9) that $M \leq K$. This shows that $(C \cap F(\theta), C \cap Y(\theta))$ maximizes (4) so that $|Y^*|(v - c) - |F^*|c = |C \cap Y(\theta)|(v - c) - |C \cap F(\theta)|c$.

Under θ, hence, the platform's payoff from the component C under the multiple group mechanism is within $|C|\frac{\varepsilon}{n}$ of the corresponding first-best payoff. Hence, the difference in total payoff under θ is at most ε. $\qquad\square$

Proof of Proposition 7. Consider first the probability that $D(\theta)$ contains no pair (i, j) of agents who are strongly preference-connected to each other (*i.e.*, both links $i \to j$ and $j \to i$ exist). We have

$$P_p(D(\theta) \text{ contains no strongly preference-connected } \{i, j\}) = (1 - p^2)^{\binom{n}{2}}.$$

---

[15]It may be the case that $Y_k^* = \emptyset$ for some k.

[16]Suppose to the contrary that there exists $i \in Y_k^*$ such that for any directed path from i to i_k, there exists $j \neq i_k$ such that $j \notin Y_k$ on the path. By definition, $j \in \cup_{\ell=1}^{k-1} Y_\ell$ so that $i \in Y_\ell^*$ for some $\ell < k$, which is a contradiction. We hence have $Y_k^* \subset C_{\leadsto i_k}(Y_k^*, \theta)$.

If we let $k = \frac{n^2}{\tau^2}$, then $k \to \infty$ as $n \to \infty$, and the above probability can be evaluated as

$$\left(1 - p^2\right)^{\binom{n}{2}} > \left(1 - \frac{\tau^2}{n^2}\right)^{\frac{n^2}{2}} = \left(1 - \frac{1}{k}\right)^{\frac{\tau^2 k}{2}} = \left(\left(1 - \frac{1}{k}\right)^k\right)^{\frac{\tau^2}{2}} \to 1 \quad \text{as } n \to \infty.$$

We next consider the probability that $D(\theta)$ contains a strongly preference-connected set of three or more agents. For this, consider an auxiliary undirected graph $U(\theta)$ obtained from $D(\theta)$ by ignoring direction of links in $D(\theta)$ and if there are two links $i \to j$ and $j \to i$ between i and j, replacing them by a single undirected link. Then

$$P_p(U(\theta) \text{ contains an (undirected) link } ij) = 1 - (1 - p)^2 < 2p,$$

and hence the stated condition on τ implies (Diestel 2000, p. 248) that

$$P_p(U(\theta) \text{ contains an (undirected) cycle of length} \geq 3) \to 0 \quad \text{as } n \to \infty.$$

Since

$$P_p(G(\theta) \text{ contains a directed cycle of length} \geq 3)$$
$$\leq P_p(U(\theta) \text{ contains an undirected cycle of length} \geq 3),$$

we conclude that $P_p(G(\theta)$ has a directed cycle of length $\geq 3) \to 0$ as $n \to \infty$. Combining the above arguments, we have

$$P_p(G(\theta) \text{ contains a strongly preference-connected set of agents}) \to 0 \quad \text{as } n \to \infty.$$

$\square$

References

Alcalde, J. & Revilla, P. (2004), 'Researching with whom? stability and manipulation', *Journal of Mathematical Economics* **40**(8), 869 – 887.

Aoyagi, M. (2013), 'Coordinating adoption decisions under externalities and incomplete information', *Games and Economic Behavior* **77**(1), 77 – 89.

Aoyagi, M. (2018), 'Bertrand competition under network externalities', *Journal of Economic Theory* **178**(C), 517–550.

Bloch, F. & Quérou, N. (2013), 'Pricing in social networks', *Games and Economic Behavior* **80**, 243 – 261.

Board, S. (2009), 'Monopolistic group design with peer effects', *Theoretical Economics* **4**(1), 89–125.

Bollobás, B. (2001), *Random Graphs*, 2nd edn, Cambridge University Press.

Caillaud, B. & Jullien, B. (2003), 'Chicken & egg: Competition among intermediation service providers', *The RAND Journal of Economics* **34**(2), 309–328.

Candogan, O., Bimpikis, K. & Ozdaglar, A. (2012), 'Optimal pricing in networks with externalities', *Operations Research* **60**(4), 883–905.

Cechlárová, K. & Romero-Medina, A. (2001), 'Stability in coalition formation games', *International Journal of Game Theory* **29**(4), 487–494.

Chen, Y., Zenou, Y. & Zhou, J. (2018), 'Competitive pricing strategies in social networks', *RAND Journal of Economics* **49**(3), 672–705.

Chen, Y., Zenou, Y. & Zhou, J. (2020), Network topology and market structure. discussion paper, CEPR.

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

Diestel, R. (2000), *Graph Theory*, 2nd edn, Springer.

Dimitrov, D., Borm, P., Hendrick, R. & Shao, C. S. (2006), 'Simple priorities and core stability in hedonic games', *Social Choice and Welfare* **26**(2), 421–433.

Dybvig, P. H. & Spatt, C. S. (1983), 'Adoption externalities as public goods', *Journal of Public Economics* **20**(2), 231 – 247.

Fainmesser, I. P. & Galeotti, A. (2016), 'Pricing Network Effects', *Review of Economic Studies* **83**(1), 165–198.

Gomes, R. & Pavan, A. (2016), 'Many-to-many matching and price discrimination', *Theoretical Economics* **11**(3), 1005–1052.

Hoppe, H. C., Moldovanu, B. & Ozdenoren, E. (2011), 'Coarse matching with incomplete information', *Economic Theory* **47**(1), 75–104.

McDiarmid, C. (1981), 'General percolation and random graphs', *Advances in applied probability* **13**(1), 40–60.

Myerson, R. B. (1982), 'Optimal coordination mechanisms in generalized principal-agent problems', *Journal of Mathematical Economics* **10**(1), 67–81.

Rodríguez-Álvarez, C. (2009), 'Strategy-proof coalition formation', *International Journal of Game Theory* **38**(3), 431–452.

Veiga, A. (2013), Platforms exclude but do not segregate. working paper, Toulouse School of Economics.