Co-Diffusion of Technologies in Social Networks

Natalia Lazzati[*†]

December 17, 2019

Abstract

This paper studies the diffusion process of two complementary technologies among people who are connected through a social network. It characterizes adoption rates over time for different initial allocations and network structures. In doing so, we provide some microfoundations for the stochastic formation of consideration sets. We are particularly interested in the following question: Suppose we want to maximize technology diffusion and have a limited number of units of each of the two technologies to initially distribute. How should we allocate these units among people in the social network?

JEL codes: D83, O33

Keywords: Complements, Supermodularity, Quasi-supermodularity, Increasing and Supermodular Stochastic Order, Technology Diffusion, Stochastic Consideration Sets

[*] *e-mail:* nlazzati@ucsc.edu.

[†] The author thanks Rabah Amir, Dan Friedman, Bruno Strulovici, the Editor of AEJ: Microeconomics (Johannes Horner), and an anonymous referee for their useful comments.

1 Introduction

There is a large body of work in economics and marketing that studies the adoption process of a single technology (or good) among people who interact through a social network. A key question that emerged early in the literature is the problem of how to optimally *seed* the technology via the selection of some initial adopters.[1] This paper focuses on a variant of this model that, surprisingly, has received much less attention. It studies the co-diffusion process of two complementary technologies among people who interact through a social network. Interestingly, the underlying dynamics between the interacting technologies involves some rich consumer behavior that calls for a very different theoretical approach and opens new possibilities for effective seeding strategies to start the market up.[2]

There are plenty of examples that motivate our model of complements in the presence of network effects. Think, for instance, in the (possibility of) joint adoption of Internet and an antivirus program provided by the same firm. It is reasonable to assume that a person would be more willing to acquire the antivirus program if he has Internet access and vice versa. Moreover, it is also natural to think that the person would be more interested in, or aware of, each of the products if more of his friends already have one or the other. Other examples in the telecommunications industry include wireless voice and data services (Niculescu and Whang (2012)), personal computers and Internet (Dewan, Ganley, and Kraemer (2010)), and mobile operators and SMS messaging (Westland, Hao, Xiao, and Shan (2016)). By focusing on situations that display these two features, we make three interrelated contributions.

First, we build a formal model to study interdependent choices for complementary technologies based on stochastic consideration sets. We capture complementarities by assuming that each person′s preferences regarding the two technologies are quasi-supermodular. Intu-

[1] Aral, Muchnik, and Sundararajan (2013) state that "seeding or network targeting has emerged as the primary marketing strategy designed to leverage the word of mouth effect".

[2] Bayus, Kim, and Shocker (2000) explore various issues associated with the market dynamics of multiple, interacting products.

itively, this property states that if a person prefers to get one of the technologies when he has not adopted the other one, then the same has to be true if he acquires the latter. We model interdependences among people via a network structure using a mean field approach. At each period of time, each person randomly meets a few people and is more willing to pay attention to a particular technology if it is popular enough among the people he encounters. In other words, the popularity of each technology among peers affects his consideration set. Afterwards, the person decides the combination of technologies to adopt (if any). We characterize the equilibrium points of the dynamic system and describe the diffusion process of adoption rates for different initial allocations and network structures.

Second, we use this framework to address a question regarding optimal network seeding that has no counterpart in the case of a single technology. Suppose we have a limited number of units of each technology to distribute among people with the purpose of maximizing adoption rates over time. For instance, let us assume we have 10,000 one-month free Internet subscriptions and 20,000 one-month free antivirus trials to distribute. How should we allocate the two technologies among people? Specifically, should we maximize the number of people who get at least one of the two technologies, or, instead, we should maximize the number of people who get both of them together? We show that adoption rates are higher over time if we maximize the number of technologies that end up in the same hands. In fact, we show via example that this allocation can markedly dominate (in terms of adoption rates) an alternative mechanism that distributes twice as many initial technologies among people. (This result is also crucial to explain the source of inefficiency generated by Nash equilibrium behavior in an extension of the model that we offer at the end of the paper.)

Lastly, we show that (under our initial assumptions) some people might be worse-off by the allocation that maximizes adoption rates over time. As we mentioned above (for a fixed number of initial technologies) the allocation that maximizes diffusion is the one that maximizes the fraction of people who get the two technologies together. The fact that some people might be worse-off by this policy happens because this allocation also maximizes the fraction of people who get none of the two technologies. We then show that everyone is better-off (on average) if the utility function is supermodular instead of just quasi-supermodular.

We offer three extensions of our initial model: The first two extensions show that the main results are robust under alternative diffusion processes. These alternative set-ups allow for memory in choices and a friendship structure that can accommodate homophily. The last set-up extends the initial monopoly model to the case of two firms, each of which controls the diffusion of only one technology. We characterize the seeding game and state the precise sense in which any Nash equilibrium is inefficient. This last result builds directly on the main insights we develop for the monopoly case.

We finally relate our findings to the existing literature. The comparison of the diffusion process for different initial allocations relies on an association order on multivariate distributions that has recently received attention in Economics: The increasing and supermodular (ISPM) stochastic order. This order allows us to compare the expectation of functions that are both increasing and supermodular. It proves useful in our framework as we show that the mapping of adoption rates from one period of time to the next can be expressed as the expectation of a function that is increasing and supermodular. Thus, the mapping of adoption rates increases with respect to the ISPM order. We also show that the set of all adoption distributions ordered by the ISPM order is a complete lattice. (From a technical perspective, this finding is interesting as previous work has documented that the same set of bivariate distributions ordered by first-order stochastic dominance is not a lattice.[3]) These two results lead to a nice characterization of equilibrium points via Tarski′s fixed point theorem. They also facilitate the comparative statics analysis we perform in the paper.

The ISPM order is strictly weaker than both the standard first-order stochastic dominance and the stochastic supermodular order. For the case of bivariate distributions, it coincides

[3] Kamae, Krengel, and O'Brien (1977) offer a counterexample which shows that the set of bivariate Bernoulli distributions (partially) ordered by first-order stochastic dominance is not a lattice. (See also Echenique (2003).) We show that the same set, coupled with a different order, has indeed a lattice structure.

with the upper orthant order and thereby has a simple representation.[4] Meyer and Strulovici (2012, 2015) provide a useful characterization of this association order coupled with many interesting applications to Economics. Dziewulski and Quah (2014) use the ISPM order to derive a revealed preference test for production with complementarities.[5]

This paper is also related to the literature on stochastic consideration sets (or limited attention mechanisms). Roughly speaking, papers in this literature assume people are boundedly rational and do not consider all the alternatives at the moment of making a decision. Instead, each subset of alternatives has some ex-ante probability of ending up as the consideration set of the person. Once the consideration set is formed (or realized), the person chooses the alternative that maximizes his utility in that consideration set (see, e.g., Manzini and Mariotti (2014)). An important, and still open, question in the literature is how the ex-ante probabilities of the different subsets of alternatives emerge. In our model, there are four possible consideration sets: the empty one, the consideration set that only includes technology one, the consideration set that only includes technology two, and a last one that includes both technologies. We let the probability of paying attention to each of the two technologies depend on the number of friends that the person observes have adopted it. The probabilities at the level of each technology generate a probability distribution on the four sub-sets of alternatives. This probability distribution function varies with the choices of peers. In making this connection we offer a microfoundation for the formation of consideration sets.

Last, but not least, the paper relates to a large and growing literature on diffusion in social networks. This literature includes, but is not restricted to, Bala and Goyal (1998), Blume (1993, 1995), Galeotti and Goyal (2009), Golub and Jackson (2010), Jackson and Rogers (2007), Jackson and Yariv (2007), Vega-Redondo (2007), and Young (2009).[6] Many of these papers have used first- and second-order stochastic dominance to perform comparative static

[4] The connection of the ISPM order with the upper orthant order has been established by Scarsini (1998).

[5] The stronger concept of supermodular stochastic order has also been recently used by Amir and Lazzati (2016), Athey and Levin (2001), and Levin (2001), among others.

[6] Jackson, Rogers, and Zenou (2016) offer a nice overview of the leading papers.

analysis of resting and tipping points of the dynamic evolution of adoption rates.[7] We extend their analysis to the case of two complementary technologies: In doing so, we study a problem of *joint* allocation of the two technologies among people that has no parallel in the case of the single technology and use (to address this question) an association order (the ISPM stochastic order) that has not been used earlier in this literature. In an extension of the model, we also set up the basis for the study of competition regarding seeding strategies. Chen, Zenou, and Zhou (2018) characterize the Nash equilibrium for players embedded in a social network that choose the level of two interdependent activities. They show that previous results for quadratic payoffs aggregate nicely to multiple actions. Though slightly related to our work, the question we address and the method we use are rather different from theirs.

From a business perspective, the importance of complementarities in technology diffusion has already been highlighted in the marketing literature. Bayus, Kim, and Shocker (2000) review part of this literature, most of which builds on variants of the canonical "Bass model" (Bass (1969)). As far as we are concerned, the question of how to optimally allocate the two technologies among people in the social network to maximize diffusion has not been formally studied.[8]

The rest of the paper is as follows. Section 2 describes the model and the diffusion process. Section 3 presents the main results. Section 4 studies three extensions of the initial model and Section 5 concludes. The Appendix contains all proofs in the paper.

[7] See, for example, Jackson and Rogers (2007).

[8] Graham, Imbens, and Ridder (2014) study econometric methods for measuring the average output effect of reallocating an indivisible input across production units in the presence of complementarities.

2 The Model

2.1 Preferences, Network, and Limited Attention

A large population of agents form a social network. We study the diffusion process of two complementary technologies, 1 and 2, among people in the network.

The population is described by a distribution of types $\mathrm{H}(n)$ for $n = 1, 2, ..., \mathrm{N}$ with $\sum_{n=1}^{\mathrm{N}} \mathrm{H}(n) = 1$. As we explain next, each type in our model incorporates three dimensions: a utility function that captures the preferences of the person over the two technologies; the popularity of the person within the network; and a limited attention mechanism that connects the choice set of each person with the adoption rates of his friends or peers.

In particular, a type n person has associated a utility function $\mathrm{U}_n(x_1, x_2)$ for $(x_1, x_2) \in \{0,1\} \times \{0,1\}$, where $x_i = 1$ means that the person gets technology $i = 1, 2$ and $x_i = 0$ means he does not get it. We restrict attention to the class of quasi-supermodular (Q-SPM) utility functions that, for completeness, we define next.

Definition (Q-SPM) For each $(x_1', x_2') \in \{(0,1), (1,0)\}$ and $(x_1'', x_2'') \in \{(0,1), (1,0)\}$ with $(x_1', x_2') \neq (x_1'', x_2'')$, we have that

$$\mathrm{U}_n(x_1', x_2') \geq (>) \, \mathrm{U}_n(0,0) \Longrightarrow \mathrm{U}_n(1,1) \geq (>) \, \mathrm{U}_n(x_1'', x_2'').$$

Q-SPM captures the notion of complementarities between the two technologies.[9] In words, it states that if a person prefers to get one of the two technologies when he has not adopted the other one, then the same has to be true if he acquires the latter. Related to this concept is the (stronger) notion of supermodular (SPM) utility functions.

Definition (SPM) $\mathrm{U}_n(1,1) + \mathrm{U}_n(0,0) \geq \mathrm{U}_n(0,1) + \mathrm{U}_n(1,0)$.

The next example illustrates these two definitions.

[9] See Topkis (1998) for a nice overview of this class of functions and its role in optimization.

Example 1 Let the utility function of a type n person be as follows

$$\mathrm{U}_n(x_1, x_2) = (\nu_{1n} - p_1)\, x_1 + (\nu_{2n} - p_2)\, x_2 + \Delta_n x_1 x_2$$

where p_1 and p_2 are the prices of technologies 1 and 2, respectively. We often refer to ν_{1n} and ν_{2n} as the stand-alone values of each technology. The term Δ_n is the interaction effect among the technologies. In this specification, the utility is SPM (and thus Q-SPM) if $\Delta_n \geq 0$.

This quasilinear specification of the utility function has been used in many papers to capture preferences for complementary goods. For instance, Fox and Lazzati (2017) show how to recover this type of utility from consumption data exploiting price variation. $\square$

The example of Internet subscriptions and the antivirus software might naturally fit our definition of complements. Along the analysis we will also assume preferences are strict. That is, we let $\mathrm{U}_n(x_1', x_2') \neq \mathrm{U}_n(x_1'', x_2'')$ for all $(x_1', x_2') \neq (x_1'', x_2'')$.

To describe the social network we use a "mean field" approach. That is, we assume the population is large enough so that the effect of all the other individuals on any given person is approximated by an average effect. A type n person randomly meets d_n people in each period of time and observes their technology choices (i.e., we model a directed network). These people are random draws from the distribution of types in the population.[10] Roughly speaking, d_n captures the popularity of the person within the network.

The probability of paying attention to a particular technology increases if the person encounters more people adopting it. Let S_{1n} and S_{2n} be the total number of people (out of the d_n people he crossed paths with) that the person observes have adopted technologies 1 and 2, respectively. Formally, the probability that the person pays attention to technologies 1 and 2, are $\pi_{1n}(\mathrm{S}_{1n})$ and $\pi_{2n}(\mathrm{S}_{2n})$, respectively, with $\pi_{1n} : [0, d_n] \to [0, 1]$ and $\pi_{2n} : [0, d_n] \to [0, 1]$. We assume that π_{1n} and π_{2n} are both increasing functions. These probabilities capture the

[10] In Sub-Section 4.2 we extend this model to the case where each person that a type n person meets is a random draw from a distribution of types that might be different from the distribution of types in the population. This extension allow us to introduce, among other things, homophily into the model —i.e., a situation in which each person interacts more often with similar people.

limited attention mechanism in our model. At each time, the consideration set Γ_n (i.e., the subset of technologies he pays attention to) is distributed as follows

$$\Gamma_n = \begin{cases} \Gamma_{0n} & = & \{0\} \times \{0\} & \text{with} & \Pr(\Gamma_{0n}) & = & [1 - \pi_{1n}(\mathrm{S}_{1n})][1 - \pi_{2n}(\mathrm{S}_{2n})] \\ \Gamma_{1n} & = & \{0,1\} \times \{0\} & \text{with} & \Pr(\Gamma_{1n}) & = & \pi_{1n}(\mathrm{S}_{1n})[1 - \pi_{2n}(\mathrm{S}_{2n})] \\ \Gamma_{2n} & = & \{0\} \times \{0,1\} & \text{with} & \Pr(\Gamma_{2n}) & = & [1 - \pi_{1n}(\mathrm{S}_{1n})]\pi_{2n}(\mathrm{S}_{2n}) \\ \Gamma_{12n} & = & \{0,1\} \times \{0,1\} & \text{with} & \Pr(\Gamma_{12n}) & = & \pi_{1n}(\mathrm{S}_{1n})\pi_{2n}(\mathrm{S}_{2n}) \end{cases}.$$

As in Demuynck and Seel (2018), the formation of consideration sets starts at the individual good (or technology) level rather than at the bundle level. This modeling approach captures the idea that the person needs to see a given technology used by his peers to pay attention to it. In specific contexts, it would be reasonable to assume that, upon observing his peers using one technology, the person puts more effort in looking for related objects. In terms of our initial example, it could be the case that a person who observes his friends using a specific Internet subscription, also looks for related products, and comes across the antivirus program. Due to the complementary nature of the two technologies in our model, this other approach would likely facilitate diffusion.

The next example illustrates our last concept.

Example 2 The specification of consideration sets we just described includes, as a particular case, a simple threshold model. Specifically, it allows the possibility that

$$\Gamma_n = \begin{cases} \Gamma_{0n} & = & \{0\} \times \{0\} & \text{if} & \mathrm{S}_{1n} < \mathrm{k}_{1n} & \text{and} & \mathrm{S}_{2n} < \mathrm{k}_{2n} \\ \Gamma_{1n} & = & \{0,1\} \times \{0\} & \text{if} & \mathrm{S}_{1n} \geq \mathrm{k}_{1n} & \text{and} & \mathrm{S}_{2n} < \mathrm{k}_{2n} \\ \Gamma_{2n} & = & \{0\} \times \{0,1\} & \text{if} & \mathrm{S}_{1n} < \mathrm{k}_{1n} & \text{and} & \mathrm{S}_{2n} \geq \mathrm{k}_{2n} \\ \Gamma_{12n} & = & \{0,1\} \times \{0,1\} & \text{if} & \mathrm{S}_{1n} \geq \mathrm{k}_{1n} & \text{and} & \mathrm{S}_{2n} \geq \mathrm{k}_{2n} \end{cases}$$

where k_{1n} and k_{2n} are the corresponding cutoffs for considering each particular technology. In this case, $\pi_{1n}(\mathrm{S}_{1n}) = 1(\mathrm{S}_{1n} \geq \mathrm{k}_{1n})$ and $\pi_{2n}(\mathrm{S}_{2n}) = 1(\mathrm{S}_{2n} \geq \mathrm{k}_{2n})$ are indicator functions. Here, the person pays attention to a particular technology if, and only if, the technology is popular enough among his friends or peers. □

We finally assume the person selects the most preferred option among the ones he is actually considering. Formally, the optimal choice of the person is given by

$$\mathrm{BR}_{\bullet n} = \arg\max\nolimits_{x_1,x_2}\left\{\mathrm{U}_n\left(x_1,x_2\right) : x_1,x_2 \in \Gamma_{\bullet n}\right\}$$

for $\bullet = 0, 1, 2, 12$. Since consideration sets are (ex-ante) stochastic, the current adoption rates induce a distribution of choices that evolves over time.

Let us finally remark that this model allows for almost unrestricted heterogeneity across people. In particular, people might differ regarding preferences for the technologies, number of people they meet, and the shape of the probabilities that define their consideration sets. The model also allows for an arbitrary pattern of association between the distribution of preferences, degree links, and consideration sets in the population. We only assume, for tractability, that the set of types is finite.

2.2 Co-Diffusion Process

We consider a diffusion process governed by best-reply dynamics in discrete time. Initially, at time 0, the two technologies are allocated in the population according to the joint distribution function P^0 described by

$$\mathrm{P}^0_{00} \geq 0,\ \mathrm{P}^0_{10} \geq 0,\ \mathrm{P}^0_{01} \geq 0,\ \mathrm{P}^0_{11} \geq 0,\ \text{and } \mathrm{P}^0_{00} + \mathrm{P}^0_{10} + \mathrm{P}^0_{01} + \mathrm{P}^0_{11} = 1$$

where P^0_{00}, P^0_{10}, P^0_{01}, and P^0_{11} are the fraction of people who have no technology, only technology 1, only technology 2, and both of them, respectively. Bivariate Bernoulli distributions can also be characterized by three parameters: Two parameters that define the marginal distributions and a third parameter that establishes the dependence between the Bernoulli marginal distributions. Let $\mathrm{P}^0_1 = \mathrm{P}^0_{10} + \mathrm{P}^0_{11}$ and $\mathrm{P}^0_2 = \mathrm{P}^0_{01} + \mathrm{P}^0_{11}$ be the total fraction of people who have each technology at time 0. In words, P^0_1 and P^0_2 characterize the marginal distributions of P^0. We find it convenient to describe the joint distribution P^0 via the alternative specification P^0_1, P^0_2, and P^0_{11} with

$$1 \geq \mathrm{P}^0_1 \geq 0,\ 1 \geq \mathrm{P}^0_2 \geq 0,\ \text{and } \min\left\{\mathrm{P}^0_1, \mathrm{P}^0_2\right\} \geq \mathrm{P}^0_{11} \geq \max\left\{0, \mathrm{P}^0_1 + \mathrm{P}^0_2 - 1\right\}.$$

The timing of the model is as follows. At period of time $t > 0$, each person randomly meets a few people and observes how many of them have adopted each of the two technologies at time $t-1$. These observations affect the attention he pays to each of the two technologies, determining his consideration set. Then, the person decides what technologies to acquire (if any) according to his best-reply function. We next describe the evolution of technology adoption rates over time induced by this model. (Note that, in this model, the current distribution of consideration sets depends on the past distribution of consideration sets only through consumption choices. Sub-section 4.1 presents a variant of this initial model in which the current distribution of consideration sets depends directly on both the past distribution of consideration sets and consumption choices.[11])

Let $\left(x_{1i}^{t-1}, x_{2i}^{t-1}\right) \in \{0,1\} \times \{0,1\}$ with $i = 1, 2, ..., d_n$ be the adoption decisions of the people that a type n person encounters. Each $\left(x_{1i}^{t-1}, x_{2i}^{t-1}\right)$ is an independent random draw from the adoption distribution P^{t-1}. For instance, the probability of meeting a person that has adopted the two technologies, $\left(x_{1i}^{t-1} = 1, x_{2i}^{t-1} = 1\right)$, is P_{11}^{t-1}. Let $\mathrm{S}_{1n}^{t-1} = \sum_{i=1}^{d_n} x_{1i}^{t-1}$ and $\mathrm{S}_{2n}^{t-1} = \sum_{i=1}^{d_n} x_{2i}^{t-1}$ be the total number of people, among the ones that the person meets, who have adopted technologies 1 and 2, respectively. The distribution of consideration sets for type n people at period of time t is as follows

$$
\begin{aligned}
\Pr\left(\Gamma_{0n}^{t} \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\left[1 - \pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right]\left[1 - \pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right]\right\} \\
\Pr\left(\Gamma_{1n}^{t} \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\left[1 - \pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right]\right\} \\
\Pr\left(\Gamma_{2n}^{t} \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\left[1 - \pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right]\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \\
\Pr\left(\Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\}
\end{aligned}
$$

where E is the expectation operator over the random vector $\left(\mathrm{S}_{1n}^{t-1}, \mathrm{S}_{2n}^{t-1}\right)$. Since type n people have associated a best-reply function for each consideration set, the distribution of consideration sets (coupled with our large population assumption) induces a distribution of optimal

[11] Under this alternative specification each person's consideration set increases over time. We thank Bruno Strulovici for suggesting this possibility.

choices at period of time t, P_n^t, given by

$$\begin{aligned}\left(\mathrm{P}_{1n}^t, \mathrm{P}_{2n}^t\right) &= \mathrm{BR}_{1n} \Pr\left(\Gamma_{1n}^t \mid \mathrm{P}^{t-1}\right) + \mathrm{BR}_{2n} \Pr\left(\Gamma_{2n}^t \mid \mathrm{P}^{t-1}\right) + \mathrm{BR}_{12n} \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) \\ \mathrm{P}_{11n}^t &= 1\left(\mathrm{BR}_{12n} = (1,1)\right) \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right)\end{aligned}$$

where $1(\cdot)$ is the standard indicator function. Adding up across all types, the distribution of technology adoption at period of time t is given by

$$\mathrm{P}_1^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{1n}^t \mathrm{H}(n), \ \mathrm{P}_2^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{2n}^t \mathrm{H}(n), \text{ and } \mathrm{P}_{11}^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{11n}^t \mathrm{H}(n).$$

This model does not have closed form solution except for extremely simple representations of utility functions and network structures. (We will use these simple representations later to illustrate some of our results.) Nevertheless, as we show below, we can still evaluate the diffusion process of technology adoption P^t for different initial allocations and address its implications for optimal seeding strategies and people welfare.

3 Main Results

3.1 Equilibrium Structure

The results in this section are easier to follow if we express the diffusion process we described above using a specific matrix form. To this end, let $\mathrm{P}_n^t = \left(\mathrm{P}_{1n}^t,\ \mathrm{P}_{2n}^t,\ \mathrm{P}_{11n}^t\right)'$, $\mathrm{P}^t = \left(\mathrm{P}_1^t,\ \mathrm{P}_2^t,\ \mathrm{P}_{11}^t\right)'$, $1_1 = (1,0)'$, and $1_2 = (0,1)'$. In addition, let us define, for each type n, the next mapping

$$\mathrm{P}_n^t = \begin{pmatrix} \mathrm{BR}_{1n} 1_1 & 0 & \left(\mathrm{BR}_{12n} - \mathrm{BR}_{1n}\right) 1_1 \\ 0 & \mathrm{BR}_{2n} 1_2 & \left(\mathrm{BR}_{12n} - \mathrm{BR}_{2n}\right) 1_2 \\ 0 & 0 & 1\left(\mathrm{BR}_{12n} = (1,1)\right) \end{pmatrix} \begin{pmatrix} \Pr\left(\Gamma_{1n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) \\ \Pr\left(\Gamma_{2n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) \\ \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) \end{pmatrix} \tag{1}$$

where the probabilities over the union of consideration sets are given by

$$\begin{aligned}\Pr\left(\Gamma_{1n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\} \\ \Pr\left(\Gamma_{2n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \\ \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right) \pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\}.\end{aligned} \tag{2}$$

These probabilities represent (given the large population assumption) the fraction of type n people who consider technology 1, technology 2, and both of them, respectively, at period of time t given the adoption rates at time $t-1$. The diffusion process in Section 2.2 is just a weighted average of P_n^t across types. That is,

$$\mathrm{P}^t = \mathrm{M}\left(\mathrm{P}^{t-1}\right) = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_n^t \mathrm{H}\left(n\right). \tag{3}$$

Let us finally define the set of all bivariate Bernoulli distributions $\mathrm{P} = (\mathrm{P}_1, \mathrm{P}_2, \mathrm{P}_{11})$ as

$$\mathcal{P} = \left\{\mathrm{P} : \mathrm{P}_1 \in [0,1], \mathrm{P}_2 \in [0,1], \text{ and } \mathrm{P}_{11} \in \left[\max\left\{0, \mathrm{P}_1 + \mathrm{P}_2 - 1\right\}, \min\left\{\mathrm{P}_1, \mathrm{P}_2\right\}\right]\right\}.$$

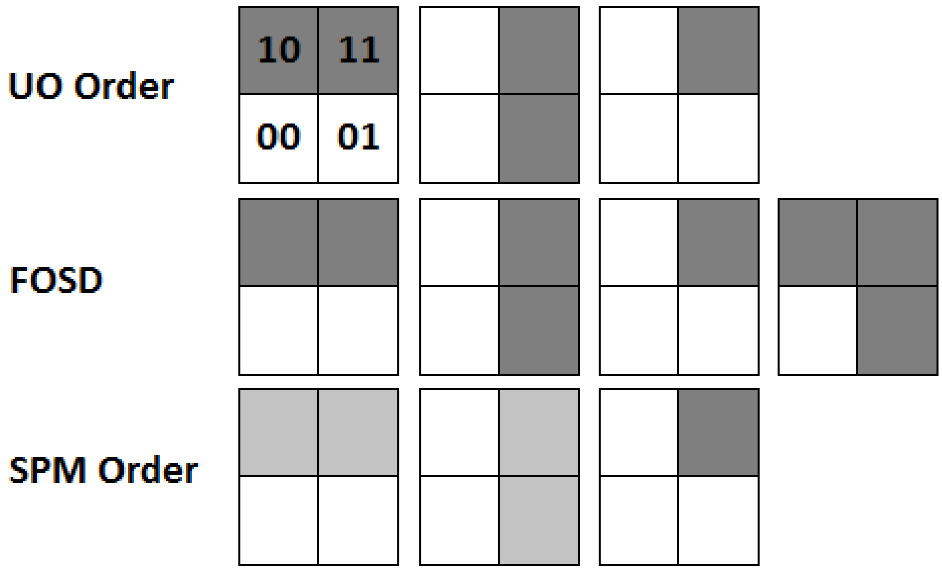
It follows that, by construction, the diffusion process M is a mapping from the set $\mathcal{P}$ into itself (i.e., M: $\mathcal{P} \to \mathcal{P}$). This section shows that the set of fixed points of M in (3) is non-empty and has a nice structure. This result builds on two lemmas that we discuss first.

The proof of existence of equilibrium points and the comparative statics analysis for the diffusion of a single technology have used first and second-order stochastic dominance (see, e.g., Jackson and Yariv (2007)). Our analysis of the case of two technologies relies on a stochastic order on multivariate distributions that is called Upper Orthant (UO) stochastic order. We specify this order next for the case of bivariate Bernoulli distributions $\overline{\mathrm{P}}$ and $\underline{\mathrm{P}}$. (See Müller and Stoyan (2002) for a fuller description of the UO stochastic order.)

Definition (UO order) We say $\overline{\mathrm{P}} \geq_{uo} \underline{\mathrm{P}}$ if $\overline{\mathrm{P}}_1 \geq \underline{\mathrm{P}}_1$, $\overline{\mathrm{P}}_2 \geq \underline{\mathrm{P}}_2$, and $\overline{\mathrm{P}}_{11} \geq \underline{\mathrm{P}}_{11}$.

According to the UO order, a probability distribution is larger than another one if it has higher probabilities for all *upper orthant* sets. In the case of bivariate Bernoulli distributions, the UO order coincides with the increasing and supermodular (ISPM) stochastic order. In turn, this order is weaker than the standard first-order stochastic dominance, which requires higher probabilities for all *upper level* sets. It is also weaker than the supermodular stochastic order, which requires the *same marginals* and higher probabilities of *upper orthant* sets. Figure 1 (below) provides a graphical representation of these statements.

Figure 1: Stochastic Orders



Each two-by-two square in Figure 1 represents the domain of probability distribution P, with 00, 10, 01, and 11 indicating no technology, only technology 1, only technology 2, and both of them, respectively. The disposition of these events in each two-by-two square is as in the two-by-two square located on the upper left corner. (Naturally, in terms of the probabilities of these events, we must have $P_{00} \geq 0$, $P_{10} \geq 0$, $P_{01} \geq 0$, $P_{11} \geq 0$, and $P_{00}+ P_{10}+ P_{01}+ P_{11} = 1$.) The first row in Figure 1 captures the UO stochastic order (UO order), the second row captures the first-order stochastic dominance (FOSD), and the last one refers to the supermodular order (SPM order). In each row, a given distribution increases (according to the corresponding order) when it has (weakly) higher probabilities for all the events shaded in dark grey and equal probabilities of the events shaded in light grey. It follows immediately from Figure 1 that the FOSD and the SPM order are more demanding than the UO order. For instance, the first three dark grey areas are identical for the UO order and the FOSD, but the latter also requires $P_{10}+P_{01}+P_{11}$ to increase (i.e., P_{00} to go down). In addition, the grey areas coincide for the UO order and the SPM order, but the latter requires the same marginals $P_1^0 = P_{10}^0 + P_{11}^0$ and $P_2^0 = P_{01}^0 + P_{11}^0$ instead of just weakly higher marginal probabilities.

The next lemma states that the set of all bivariate Bernoulli distributions ordered by the UO order is a complete lattice.

Lemma 1 *$\mathcal{P}$ (partially) ordered by the UO order is a complete lattice.*

Remark Kamae, Krengel, and O'Brien (1977) offer a counterexample which shows that $\mathcal{P}$ (partially) ordered by first-order stochastic dominance is not a lattice (see also Echenique (2003)). Interestingly, we show that the same set of distributions, coupled with a different order, has indeed a lattice structure.

To offer a sketch of proof of Lemma 1, note that $\mathcal{P}$ is a proper subset of $\mathbb{R}^3$. In addition, for bivariate Bernoulli distributions characterized by the triplet $\mathrm{P} = (\mathrm{P}_1, \mathrm{P}_2, \mathrm{P}_{11})$, the UO order coincides with the standard coordinatewise order in $\mathbb{R}^3$.[12] Thus, completeness follows as $\mathcal{P}$ is a compact subset of $\mathbb{R}^3$. Showing that $\mathcal{P}$ is a lattice takes a few more steps. Given $\mathrm{P}' \in \mathcal{P}$ and $\mathrm{P}'' \in \mathcal{P}$, denote by $\inf_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'')$ and $\sup_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'')$ the infimum and supremum of P' and P'' in $\mathcal{P}$. We use $\inf(\mathrm{P}', \mathrm{P}'')$ and $\sup(\mathrm{P}', \mathrm{P}'')$ for the infimum and supremum of P' and P'' in $\mathbb{R}^3$. We show in the appendix that $\inf_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'') = \inf(\mathrm{P}', \mathrm{P}'')$ and

$$\sup_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'') = \left(\max\{\mathrm{P}'_1, \mathrm{P}''_1\}, \max\{\mathrm{P}'_2, \mathrm{P}''_2\}, \max\{\max\{\mathrm{P}'_1, \mathrm{P}''_1\} + \max\{\mathrm{P}'_2, \mathrm{P}''_2\} - 1, \mathrm{P}'_{11}, \mathrm{P}''_{11}\}\right).$$

That is, the infimum of P' and P'' in $\mathcal{P}$ is simply the coordinatewise infimum of P' and P'' in $\mathbb{R}^3$. For this part of the proof we show that $\inf(\mathrm{P}', \mathrm{P}'') \in \mathcal{P}$.

Alternatively, the supremum of P' and P'' in $\mathcal{P}$ corrects $\sup(\mathrm{P}', \mathrm{P}'')$ up to make sure that the third term (P_{11}) is above its lower bound ($\mathrm{P}_{11} \geq \mathrm{P}_1 + \mathrm{P}_2 - 1$). This correction works as any other upper bound for P' and P'' would require higher marginals (P_1 and P_2), and these higher marginals would push the lower bound for the third term even higher. The next example shows that this adjustment is indeed needed.

[12] Let $(a_i)_{i=1}^n \in \mathbb{R}^n$ and $(b_i)_{i=1}^n \in \mathbb{R}^n$ be two vectors in $\mathbb{R}^n$. The standard coordinatewise order in $\mathbb{R}^n$ states that $(a_i)_{i=1}^n \geq (b_i)_{i=1}^n$ if $a_i \geq b_i$ for all $i = 1, ..., n$.

Example 3 Let $\mathrm{P}' = (0.8, 0.2, 0.1)$ and $\mathrm{P}'' = (0.2, 0.8, 0.2)$. Note that $\mathrm{P}' \in \mathcal{P}$ and $\mathrm{P}'' \in \mathcal{P}$. We have that $\inf(\mathrm{P}', \mathrm{P}'') = (0.2, 0.2, 0.1)$ is in $\mathcal{P}$. Thus, $\inf_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'') = (0.2, 0.2, 0.1)$. On the other hand, $\sup(\mathrm{P}', \mathrm{P}'') = (0.8, 0.8, 0.2)$ is not in $\mathcal{P}$ because $0.2 \ngeq 0.8 + 0.8 - 1$. In this case, $\sup_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'') = (0.8, 0.8, 0.6)$. □

The following lemma states that technology adoption rates increase with respect to the UO stochastic order when the initial adoption rates increase regarding the same order.

Lemma 2 *Mapping M: $\mathcal{P} \to \mathcal{P}$ is monotone increasing (in the UO order).*

We provide a sketch of proof of Lemma 2 and connect our result with the ISPM stochastic order. The proof of Lemma 2 (in the appendix) relies on three intermediate results. Let $\overline{\mathrm{P}}^{t-1} \geq_{uo} \underline{\mathrm{P}}^{t-1}$. In the first two intermediate results, we show that the three probabilities in (2) are higher when the consideration sets are generated by $\overline{\mathrm{P}}^{t-1}$ as compared to $\underline{\mathrm{P}}^{t-1}$. That is, higher adoption distributions induce higher distributions of consideration sets (where higher is with respect to the UO order). (From a technical perspective, we show that the sum of d random draws from two bivariate Bernoulli distributions that are ordered regarding the UO order, preserve the same ordering.)

The third intermediate result shows that when we restrict attention to Q-SPM utility functions, the following is true: $\mathrm{BR}_{12n} \geq \mathrm{BR}_{1n} + \mathrm{BR}_{2n}$. As a consequence, all terms in the 3×3 matrix of P_n^t in expression (1) are (weakly) positive for each $n = 1, 2, ..., \mathrm{N}$.

These three intermediate results together imply that the probabilities P_{1n}^t, P_{2n}^t, and P_{11n}^t generated by $\overline{\mathrm{P}}^{t-1}$ are larger than the ones generated by $\underline{\mathrm{P}}^{t-1}$. Finally, note that

$$\mathrm{P}_1^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{1n}^t \mathrm{H}(n), \ \mathrm{P}_2^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{2n}^t \mathrm{H}(n), \text{ and } \mathrm{P}_{11}^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{11n}^t \mathrm{H}(n).$$

Thus, the claim in the lemma follows because P_1^t, P_2^t, and P_{11}^t are weighted averages of P_{1n}^t, P_{2n}^t, and P_{11n}^t across types $n = 1, 2, ..., \mathrm{N}$.

Interestingly, there is an alternative way of proving Lemma 2 that exploits the equivalence between the UO and the ISPM orders for the case of bivariate distributions established by

Scarsini (1998). This result adds to the uses of the ISPM order in Economics that have been documented by Meyer and Strulovici (2012, 2015).

Connection of Lemma 2 with the ISPM Stochastic Order: Let $\overline{\mathrm{P}}^{t-1} \geq_{uo} \underline{\mathrm{P}}^{t-1}$. Recall that $\mathrm{P}_1^t = \sum_{n=1}^{\mathrm{N}} \mathrm{P}_{1n}^t \mathrm{H}(n)$, $\mathrm{P}_2^t = \sum_{n=1}^{\mathrm{N}} \mathrm{P}_{2n}^t \mathrm{H}(n)$, and $\mathrm{P}_{11}^t = \sum_{n=1}^{\mathrm{N}} \mathrm{P}_{11n}^t \mathrm{H}(n)$, with

$$\begin{aligned} \left(\mathrm{P}_{1n}^t, \mathrm{P}_{2n}^t\right) &= \mathrm{BR}_{1n} \Pr\left(\Gamma_{1n}^t \mid \mathrm{P}^{t-1}\right) + \mathrm{BR}_{2n} \Pr\left(\Gamma_{2n}^t \mid \mathrm{P}^{t-1}\right) + \mathrm{BR}_{12n} \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) \\ \mathrm{P}_{11n}^t &= 1\left(\mathrm{BR}_{12n} = (1,1)\right) \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right). \end{aligned}$$

Let us define a partial order on the subindices of $\mathrm{BR}_{\bullet n}$ and $\Gamma_{\bullet n}$ as follows: $12 \geq 1$, $12 \geq 2$, $2 \geq 0$, and $1 \geq 0$. (We let 1 and 2 be unordered.) When preferences are Q-SPM, we get that $\mathrm{BR}_{12n} \geq \mathrm{BR}_{1n} + \mathrm{BR}_{2n}$. Also, $\mathrm{BR}_{0n} = (0,0)$. It follows that

$$\mathrm{BR}_{12n} + \mathrm{BR}_{0n} \geq \mathrm{BR}_{1n} + \mathrm{BR}_{2n}.$$

That is, $\mathrm{BR}_{\bullet n}$ is SPM in $\bullet$. In addition, $\mathrm{BR}_{12n} \geq \mathrm{BR}_{1n}$, $\mathrm{BR}_{12n} \geq \mathrm{BR}_{2n}$, $\mathrm{BR}_{1n} \geq \mathrm{BR}_{0n}$, and $\mathrm{BR}_{2n} \geq \mathrm{BR}_{0n}$. Thus, $\mathrm{BR}_{\bullet n}$ is also increasing in $\bullet$. It follows that vector $\left(\mathrm{P}_{1n}^t, \mathrm{P}_{2n}^t\right)$ is the expectation of an increasing and supermodular function. As we mentioned earlier, $\Pr\left(\Gamma_{1n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right)$, $\Pr\left(\Gamma_{2n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right)$ and $\Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right)$ are higher when the consideration sets are generated by $\overline{\mathrm{P}}^{t-1}$ as compared to $\underline{\mathrm{P}}^{t-1}$. Thus, $\Pr\left(\Gamma_n^t \mid \overline{\mathrm{P}}^{t-1}\right)$ dominates $\Pr\left(\Gamma_n^t \mid \underline{\mathrm{P}}^{t-1}\right)$ with respect to the UO stochastic order. In the case of bivariate distributions, the UO order coincides with the ISPM stochastic order. As the ISPM stochastic order allows us to compare the expectation of functions that are increasing and supermodular, we get that $\left(\mathrm{P}_{1n}^t, \mathrm{P}_{2n}^t\right)$ is higher when it is generated by $\overline{\mathrm{P}}^{t-1}$ as compared to $\underline{\mathrm{P}}^{t-1}$.[13] The fact that P_{11n}^t is also higher follows by simple inspection. The result then follows as P_1^t, P_2^t, and P_{11}^t are weighted averages of P_{1n}^t, P_{2n}^t, and P_{11n}^t across types $n = 1, 2, ..., \mathrm{N}$. □

We now use Lemmas 1 and 2 to characterize the set of fixed points of mapping M.

[13] Note that $\mathrm{BR}_{\bullet n}$: $\Gamma_{\bullet n} \rightarrow \{0,1\} \times \{0,1\}$. Though the main characterization of the ISPM stochastic order involves comparing the expectation of *real valued functions* that are increasing and supermodular, the characterization easily extends to this case.

Theorem 3 *The set of fixed points of M is a non-empty complete lattice.*

As we anticipated, Theorem 3 states that the set of fixed points of mapping M has a nice structure. In particular, this set has a largest and a smallest element that, in our model, correspond to the largest and the smallest resting points in terms of diffusion rates. Its proof uses Lemmas 1 and 2 to show the main result via Tarski′s fixed point theorem.

Note that if $\pi_{1n}(0) = 0$ and $\pi_{2n}(0) = 0$, then P characterized by $\mathrm{P}_1 = \mathrm{P}_2 = \mathrm{P}_{11} = 0$ (i.e., $\mathrm{P}_{00} = 1$) is always a fixed point. This corresponds to a situation in which people do not pay any attention to a particular technology unless they see other people using it. This observation highlights the relevance of seeding strategies to start the market up! The next example illustrates the main ideas in this section and motivates the one that follows.

Example 4 Suppose there is only one type of person (to avoid using subindex n) with a utility function specified as follows

$$\mathrm{U}(0,0) = 10,\ \mathrm{U}(1,0) = 5,\ \mathrm{U}(0,1) = 6, \text{ and } \mathrm{U}(1,1) = 15.$$

This function is SPM (and thus Q-SPM). Moreover, $\mathrm{BR}_1 = (0,0)$, $\mathrm{BR}_2 = (0,0)$, and $\mathrm{BR}_{12} = (1,1)$. In addition, let us assume $d = 1$, $\pi_1(\mathrm{S}_1) = 1(\mathrm{S}_1 \geq 1)$, and $\pi_2(\mathrm{S}_2) = 1(\mathrm{S}_2 \geq 1)$. (This limited attention mechanism is like the one in Example 2 but with $\mathrm{k}_1 = \mathrm{k}_2 = 1$.) Under this specification, the distribution of consideration sets in period of time t for a choice distribution P^{t-1} at $t-1$ is

$$\Pr\left(\Gamma_1^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}\right) = \mathrm{P}_1^{t-1},\ \Pr\left(\Gamma_2^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}\right) = \mathrm{P}_2^{t-1}, \text{ and } \Pr\left(\Gamma_{12}^t \mid \mathrm{P}^{t-1}\right) = \mathrm{P}_{11}^{t-1}.$$

Thus, mapping M takes the form of

$$\mathrm{P}^t = \mathrm{M}\left(\mathrm{P}^{t-1}\right) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{P}_1^{t-1} \\ \mathrm{P}_2^{t-1} \\ \mathrm{P}_{11}^{t-1} \end{pmatrix}.$$

It follows that the set of fixed points is given by

$$\{\mathrm{P} \in \mathcal{P} : \mathrm{P}_1 = \mathrm{P}_2 = \mathrm{P}_{11}\}.$$

As Theorem 3 states, this set is indeed a lattice. In particular, the largest fixed point is $\mathrm{P}_1 = \mathrm{P}_2 = \mathrm{P}_{11} = 1$. The smallest fixed point is $\mathrm{P}_1 = \mathrm{P}_2 = \mathrm{P}_{11} = 0$ (i.e., $\mathrm{P}_{00} = 1$). Moreover, the latter distribution is reached for any initial allocation P^0 that involves $\mathrm{P}^0_{11} = 0$. $\square$

3.2 Increasing Co-Diffusion and Optimal Seeding Strategy

By construction, each fixed point of M in (3) is a resting point of P^t for some P^0. The results in this section compare the diffusion process over time and the location of the resting points for different initial allocations of the two technologies and network structures.

The first proposition is as follows.

Proposition 4 *If $\overline{P}^0 \geq_{uo} \underline{P}^0$, then $\overline{P}^t \geq_{uo} \underline{P}^t$ for all $t > 0$. If, in addition, $\overline{P}^t$ and $\underline{P}^t$ converge weakly to $\overline{P}$ and $\underline{P}$, respectively, then $\overline{P} \geq_{uo} \underline{P}$.*

Remark In particular, if $\overline{\mathrm{P}}^0 = (1,1,1)$ and $\underline{\mathrm{P}}^0 = (0,0,0)$, then $\overline{\mathrm{P}}^t$ and $\underline{\mathrm{P}}^t$ converge to the largest and the smallest fixed points of M, respectively.

The first claim in the proposition follows by applying the result of Lemma 2 t times. It states that the diffusion process is higher at each period of time if we start with a higher initial allocation, where higher is with respect to the UO order. The second claim follows from the previous one and the fact that the UO order is closed under weak convergence. Interestingly, this result could be used to construct an empirical test of complementarities among the two technologies. The reason is that, if the preferences of people for the two technologies were instead independent, then the main matrix in the diffusion process of Eq. (1) in Sub-Section 3.1 would be diagonal. In turn, this would imply that, if $\overline{\mathrm{P}}^0 \geq_{uo} \underline{\mathrm{P}}^0$ holds with $\overline{\mathrm{P}}^0_1 > \underline{\mathrm{P}}^0_1$ and $\overline{\mathrm{P}}^0_2 = \underline{\mathrm{P}}^0_2$ (that is, the firm only increases the initial allocation of technology 1), then the steady state equilibrium would display $\overline{\mathrm{P}}_1 \geq \underline{\mathrm{P}}_1$ and $\overline{\mathrm{P}}_2 = \underline{\mathrm{P}}_2$. (In the model of complements, Proposition 4 allows for $\overline{\mathrm{P}}_2 \geq \underline{\mathrm{P}}_2$.) This argument is similar to the idea (in a static model) of using variation in the price of one of the two goods to check for complementarities in preferences via variation in the units acquired of the other good (see, e.g., Athey and Stern (1998), and the literature therein).

From an economic perspective, Proposition 4 has important implications for marketing strategies, as we show next. (It is also crucial to explain the source of inefficiency in the seeding game we cover in Sub-Section 4.3.)

Optimal Seeding Strategy for Diffusion of Complements: Suppose we want to allocate a limited number of units of each technology, captured by P_1^0 and P_2^0, among people at period of time 0. We want to do so with the purpose of maximizing the adoption rates over time. Given Proposition 4 we should distribute the initial technologies maximizing the chance that each person gets *both* technologies together. Recall that, for fixed marginals, we have that

$$\mathrm{P}_{11}^0 \in \left[\max\left\{0, \mathrm{P}_1^0 + \mathrm{P}_2^0 - 1\right\}, \min\left\{\mathrm{P}_1^0, \mathrm{P}_2^0\right\}\right]$$

Thus, the optimal allocation involves setting $\mathrm{P}_{11}^0 = \min\left\{\mathrm{P}_1^0, \mathrm{P}_2^0\right\}$.

Note that, by definition,

$$\mathrm{P}_1^0 + \mathrm{P}_2^0 - \mathrm{P}_{11}^0 + \mathrm{P}_{00}^0 = 1.$$

Thus, by selecting P_{11}^0 as large as possible, we are also increasing the fraction of people who receive none of the technologies, P_{00}^0. This observation highlights the difference between the UO order and the first-order stochastic dominance we explained earlier. It also anticipates our finding regarding the effect of maximal diffusion on people welfare: it might hurt some people in the population. □

The next example illustrates the practical importance of the optimal seeding strategy to maximize technology diffusion.

Example 4 (continued) We keep the same model we described earlier. Thus, mapping M takes the form of

$$\mathrm{P}^t = \mathrm{M}\left(\mathrm{P}^{t-1}\right) = \left(\mathrm{P}_{11}^{t-1}, \mathrm{P}_{11}^{t-1}, \mathrm{P}_{11}^{t-1}\right).$$

Let us consider an initial allocation $\mathrm{P}_1^0 = 1/2$, $\mathrm{P}_2^0 = 1/2$, and $\mathrm{P}_{11}^0 = 0$. That is, we give technology 1 to half of the population and technology 2 to the other half. Given our previous characterization,

$$\mathrm{P}_1^t = 0,\ \mathrm{P}_2^t = 0, \text{ and } \mathrm{P}_{11}^t = 0 \text{ for all } t \geq 1.$$

That is, from period of time 1 the market collapses for the two technologies. Let us consider next distributing half as many technologies, but let us allocate them together. That is, $\mathrm{P}_1^0 = 1/4$, $\mathrm{P}_2^0 = 1/4$, and $\mathrm{P}_{11}^0 = 1/4$. It follows that

$$\mathrm{P}_1^t = 1/4,\ \mathrm{P}_2^t = 1/4, \text{ and } \mathrm{P}_{11}^t = 1/4 \text{ for all } t \geq 1.$$

Thus, while the first allocation distributes twice as many units of the two technologies among people, the second one generates much larger adoption rates over time. □

We finally evaluate the diffusion process of technology adoption for different network structures. To this end, let $d = (d_n)_{n=1}^{\mathrm{N}}$ and $\pi_1(\cdot),\ \pi_2(\cdot) = (\pi_{1n}(\cdot),\ \pi_{2n}(\cdot))_{n=1}^{\mathrm{N}}$ be partially ordered by the natural product order and product order on functions (i.e., $f(\cdot) \geq g(\cdot)$ if $f(x) \geq g(x)$ for each x), respectively. The next result states that technology diffusion increases with respect to the UO order when each person interacts with more people or when the likelihood of considering each technology increases. It extends to the case of two technologies a similar finding in the single technology diffusion model; while our result relies on the UO order, the latter uses first-order stochastic dominance (see, e.g., Proposition 3 in Jackson and Yariv (2007)).

Proposition 5 *For a given P^0, let $\overline{P}^t$ and $\underline{P}^t$ be generated by $\overline{d}, \overline{\pi}_1(\cdot), \overline{\pi}_2(\cdot)$ and $\underline{d}, \underline{\pi}_1(\cdot), \underline{\pi}_2(\cdot)$, respectively, with $\overline{d}, \overline{\pi}_1(\cdot), \overline{\pi}_2(\cdot) \geq \underline{d}, \underline{\pi}_1(\cdot), \underline{\pi}_2(\cdot)$. Thus, $\overline{P}^t \geq_{uo} \underline{P}^t$ for all $t > 0$. If, in addition, $\overline{P}^t$ and $\underline{P}^t$ converge weakly to $\overline{P}$ and $\underline{P}$, respectively, then $\overline{P} \geq_{uo} \underline{P}$.*

In particular, the smallest and the largest fixed points of M generated by $\overline{d}, \overline{\pi}_1(\cdot), \overline{\pi}_2(\cdot)$ are higher than the ones generated by $\underline{d}, \underline{\pi}_1(\cdot), \underline{\pi}_2(\cdot)$.

3.3 Increasing Co-Diffusion and Consumer Welfare

We now turn attention to the impact of the diffusion process on people welfare. In particular, we want to learn whether people are better-off with the initial allocation that maximizes technology diffusion. Recall that (for fixed marginals) the initial allocation that maximizes diffusion is the one that maximizes the fraction of people who get the two technologies together. Recall also that, by definition,

$$\mathrm{P}_1 + \mathrm{P}_2 - \mathrm{P}_{11} + \mathrm{P}_{00} = 1.$$

Thus, as we mentioned earlier, keeping P_1 and P_2 fixed, the allocation that maximizes diffusion also increases the fraction of people who get none of the two technologies. We first show (via example) that this possibility can lead to situations in which some people are indeed worse-off with the initial allocation that maximizes diffusion. We then show that this does not happen if utilities are SPM instead of just Q-SPM.

To elaborate on this point let $\mathrm{V}_{\bullet n} = \mathrm{U}_n(\mathrm{BR}_{\bullet n})$, for $\bullet = 0, 1, 2, 12$, be the indirect utility for a type n person under different consideration sets. The expected indirect utility of this type of person in period of time t can be expressed as follows

$$\begin{aligned} \mathrm{E}\left(\mathrm{V}_n \mid \mathrm{P}^{t-1}\right) &= \mathrm{V}_{0n} + (\mathrm{V}_{1n} - \mathrm{V}_{0n}) \Pr\left(\Gamma^t_{1n} \cup \Gamma^t_{12n} \mid \mathrm{P}^{t-1}\right) + (\mathrm{V}_{2n} - \mathrm{V}_{0n}) \Pr\left(\Gamma^t_{2n} \cup \Gamma^t_{12n} \mid \mathrm{P}^{t-1}\right) \\ &\quad + (\mathrm{V}_{12n} + \mathrm{V}_{0n} - \mathrm{V}_{1n} - \mathrm{V}_{2n}) \Pr\left(\Gamma^t_{12n} \mid \mathrm{P}^{t-1}\right). \end{aligned} \tag{4}$$

The next example shows that, without extra assumptions, a higher initial allocation (in terms of the UO order) can indeed decrease the expected indirect utility of certain type of people.

Example 5 Suppose there is only one type of person with a utility function specified as follows

$$\mathrm{U}(0,0) = 0,\ \mathrm{U}(1,0) = 10,\ \mathrm{U}(0,1) = 15, \text{ and } \mathrm{U}(1,1) = 20.$$

Note that this utility is Q-SPM (but not SPM). In addition, $\mathrm{BR}_1 = (1,0)$, $\mathrm{BR}_2 = (0,1)$, and $\mathrm{BR}_{12} = (1,1)$. Let us also assume $d = 1$, $\pi_1(\mathrm{S}_1) = 1(\mathrm{S}_1 \geq 1)$, and $\pi_2(\mathrm{S}_2) = 1(\mathrm{S}_2 \geq 1)$.

Under this specification, as in Example 4, we have that

$$\Pr\left(\Gamma_1^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}\right) = \mathrm{P}_1^{t-1},\ \Pr\left(\Gamma_2^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}\right) = \mathrm{P}_2^{t-1}, \text{ and } \Pr\left(\Gamma_{12}^t \mid \mathrm{P}^{t-1}\right) = \mathrm{P}_{11}^{t-1}.$$

Thus, mapping M takes the form of

$$\mathrm{P}^t = \mathrm{M}\left(\mathrm{P}^{t-1}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{P}_1^{t-1} \\ \mathrm{P}_2^{t-1} \\ \mathrm{P}_{11}^{t-1} \end{pmatrix}.$$

It follows that the set of fixed points of M is the whole $\mathcal{P}$. Moreover, for any P^0,

$$\mathrm{E}\left(\mathrm{V} \mid \mathrm{P}^{t-1}\right) = 10\mathrm{P}_1^0 + 15\mathrm{P}_2^0 - 5\mathrm{P}_{11}^0 \text{ for all } t \geq 1.$$

Since $\mathrm{E}\big(\mathrm{V} \mid \mathrm{P}^{t-1}\big)$ decreases in P_{11}^0, a higher initial allocation (in terms of the UO order) decreases the expected indirect utility of each person for any $t \geq 1$. □

We finally show that the expected indirect utility of a given type of people increases with technology diffusion if its corresponding utility is SPM instead of just Q-SPM. This result is important as it provides precise conditions under which the preferences of both the firm and the consumers are aligned regarding the initial allocation strategy. We will invoke this proposition later to state a similar claim for the seeding game we study in Sub-Section 4.3.

Proposition 6 *Assume that, for some $n \in \{1, 2, ..., N\}$, U_n is SPM. If $\overline{P}^0 \geq_{uo} \underline{P}^0$ then $E\left(V_n \mid \overline{P}^{t-1}\right) \geq E\left(V_n \mid \underline{P}^{t-1}\right)$ for all $t > 0$. If, in addition, $\overline{P}^t$ and $\underline{P}^t$ converge weakly to $\overline{P}$ and $\underline{P}$, respectively, then $E\left(V_n \mid \overline{P}\right) \geq E\left(V_n \mid \underline{P}\right)$.*

The proof of Proposition 6 is as follows. We explained earlier that the three probabilities in expression (4) increase in P^{t-1}. In addition, by the optimality principle, $\mathrm{V}_{1n} \geq \mathrm{V}_{0n}$ and $\mathrm{V}_{2n} \geq \mathrm{V}_{0n}$. The claim follows as we show (in the appendix) that if the utility is SPM, then the corresponding indirect utility satisfies

$$\mathrm{V}_{12n} + \mathrm{V}_{0n} - \mathrm{V}_{1n} - \mathrm{V}_{2n} \geq 0.$$

This last part is the one that can fail, as Example 5 shows, if the utility is just Q-SPM.

4 Some Extensions of the Model

4.1 Allowing for Memory

So far, we studied a model in the spirit of the limited attention mechanisms. In particular, we assumed that, at each period of time, each person fully revises his consideration set based on other people′s choices. In doing so, the model allows for the possibility that a person who is currently considering the two technologies ($\Gamma = \Gamma_{12}$), switches to only considering technology 2 in the next period ($\Gamma = \Gamma_2$), as technology 1 is no longer popular among his peers. A similar idea appears in models of Bayesian learning without recall (see, e.g., Amin Rahimian and Jadbabaie (2017)). Though this model captures interesting possibilities, alternative specifications might also be attractive. This sub-section considers one of them. In particular, we study a model in which, if a person considers (or becomes aware of) a technology at some point in time, then that technology is part of his consideration sets in all subsequent periods. As we next show, our main results remain valid under this alternative process.

The co-diffusion process in this section is as the one in Section 2.2 except for a single aspect: We assume that the observation of other people choices can only enlarge the consideration set of any given person. To formalize this idea, let us denote by

$$\mathrm{Q}_n^t = \left(\Pr\left(\Gamma_{1n}^t \cup \Gamma_{12n}^t\right), \Pr\left(\Gamma_{2n}^t \cup \Gamma_{12n}^t\right), \Pr\left(\Gamma_{12n}^t\right)\right)$$

the distribution of consideration sets for type n people at period of time t. Note that, by construction, $\mathrm{Q}_n^t \in \mathcal{P}$ for all $t \geq 0$. Initially, at time 0, the two technologies are allocated in the population according to the joint distribution function P^0. If this allocation is the same across types (i.e., random initial allocation), then it also defines the initial distribution of consideration sets. That is, $\mathrm{P}^0 = \mathrm{Q}_n^0$ for all $n = 1, ..., \mathrm{N}$. For each $t > 0$, the mapping of technology diffusion, for each type n, takes the form of

$$\mathrm{P}_n^t = \begin{pmatrix} \mathrm{BR}_{1n}1_1 & 0 & (\mathrm{BR}_{12n} - \mathrm{BR}_{1n})\,1_1 \\ 0 & \mathrm{BR}_{2n}1_2 & (\mathrm{BR}_{12n} - \mathrm{BR}_{2n})\,1_2 \\ 0 & 0 & 1\,(\mathrm{BR}_{12n} = 1, 1) \end{pmatrix} \begin{pmatrix} \Pr\left(\Gamma_{1n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) \\ \Pr\left(\Gamma_{2n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) \\ \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) \end{pmatrix}.$$

This linear system differs from expression (1) as now current choices directly depend on both the previous adoption rates and the previous distribution of consideration sets. It takes just a bit of algebra to show that the distribution of consideration sets for type n people at period of time t, Q_n^t, can be conveniently expressed as follows

$$\begin{aligned} \Pr\left(\Gamma_{1n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) &= \Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) + \left[1 - \Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\} \\ \Pr\left(\Gamma_{2n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) &= \Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) + \left[1 - \Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{2n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \end{aligned}$$

$$\Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) = \begin{cases} \Pr\left(\Gamma_{12n}^{t-1}\right) + \\ \left[\Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) - \Pr\left(\Gamma_{12n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} + \\ \left[\Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) - \Pr\left(\Gamma_{12n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\} + \\ \left[1 - \Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) - \Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) + \Pr\left(\Gamma_{12n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \end{cases}$$

That is, the fraction of people who are currently considering technology 1, technology 2, and both of them is the same as in the previous period of time plus a few terms that reflect the social learning that takes place by observing other people′s choices.

Under this specification, the mapping of technology diffusion from a period of time to the next takes the form of

$$\left(\mathrm{P}^t, \left(\mathrm{Q}_n^t\right)_{n=1}^{\mathrm{N}}\right) = \mathrm{M}^*\left(\mathrm{P}^{t-1}, \left(\mathrm{Q}_n^{t-1}\right)_{n=1}^{\mathrm{N}}\right) = \left(\sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_n^t \mathrm{H}\left(n\right), \left(\mathrm{Q}_n^t\right)_{n=1}^{\mathrm{N}}\right).$$

This mapping incorporates direct information about both adoption rates and consideration sets across types. Recall that $\mathrm{Q}_n^t \in \mathcal{P}$ for all $t \geq 0$. Thus, the diffusion process M^* is a mapping from the set $\mathcal{P}^{\mathrm{N}+1}$ into itself (i.e., M^*: $\mathcal{P}^{\mathrm{N}+1} \rightarrow \mathcal{P}^{\mathrm{N}+1}$). The next result states that the claims in Lemmas 1 and 2 extend to this alternative specification. All other propositions remain valid because they are based on these two lemmas and the 3×3 matrix in expression (1), which has not changed.

Lemma 7 *$\mathcal{P}^{N+1}$ (partially) ordered by the UO order is a complete lattice and mapping M^*: $\mathcal{P}^{N+1} \to \mathcal{P}^{N+1}$ is monotone increasing (in the UO order).*

The next example shows that, as expected, this second model facilitates diffusion.

Example 6 Let us use the specification of Example 4 with the new diffusion process. That is, suppose there is only one type of person with a utility function specified as follows

$$\mathrm{U}(0,0) = 10,\ \mathrm{U}(1,0) = 5,\ \mathrm{U}(0,1) = 6,\ \text{and}\ \mathrm{U}(1,1) = 15.$$

This function is SPM (and thus Q-SPM). Moreover, $\mathrm{BR}_1 = (0,0)$, $\mathrm{BR}_2 = (0,0)$, and $\mathrm{BR}_{12} = (1,1)$. In addition, let us assume $d = 1$, $\pi_1(\mathrm{S}_1) = 1(\mathrm{S}_1 \geq 1)$, and $\pi_2(\mathrm{S}_2) = 1(\mathrm{S}_2 \geq 1)$.

Under the new diffusion model, the distribution of consideration sets in period t is

$$\begin{aligned}
\Pr\left(\Gamma_1^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}^{t-1}\right) &= \Pr\left(\Gamma_1^{t-1} \cup \Gamma_{12}^{t-1}\right) + \left[1 - \Pr\left(\Gamma_1^{t-1} \cup \Gamma_{12}^{t-1}\right)\right] \mathrm{P}_1^{t-1} \\
\Pr\left(\Gamma_2^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}^{t-1}\right) &= \Pr\left(\Gamma_2^{t-1} \cup \Gamma_{12}^{t-1}\right) + \left[1 - \Pr\left(\Gamma_1^{t-1} \cup \Gamma_{12}^{t-1}\right)\right] \mathrm{P}_2^{t-1} \\
\Pr\left(\Gamma_{12}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}^{t-1}\right) &= \begin{cases} \Pr\left(\Gamma_{12}^{t-1}\right) + \\ \left[\Pr\left(\Gamma_1^{t-1} \cup \Gamma_{12}^{t-1}\right) - \Pr\left(\Gamma_{12}^{t-1}\right)\right] \mathrm{P}_2^{t-1} + \\ \left[\Pr\left(\Gamma_2^{t-1} \cup \Gamma_{12}^{t-1}\right) - \Pr\left(\Gamma_{12}^{t-1}\right)\right] \mathrm{P}_1^{t-1} + \\ \left[1 - \Pr\left(\Gamma_1^{t-1} \cup \Gamma_{12}^{t-1}\right) - \Pr\left(\Gamma_2^{t-1} \cup \Gamma_{12}^{t-1}\right) + \Pr\left(\Gamma_{12}^{t-1}\right)\right] \mathrm{P}_{11}^{t-1} \end{cases}
\end{aligned}$$

In addition,

$$\mathrm{P}^t = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Pr\left(\Gamma_1^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}^{t-1}\right) \\ \Pr\left(\Gamma_2^t \cup \Gamma_{12}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}^{t-1}\right) \\ \Pr\left(\Gamma_{12}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}^{t-1}\right) \end{pmatrix}.$$

In this simple specification, there are (only) two fixed points in terms of adoption rates

$$\mathrm{P}_1 = \mathrm{P}_2 = \mathrm{P}_{11} = 0 \text{ and } \mathrm{P}_1 = \mathrm{P}_2 = \mathrm{P}_{11} = 1.$$

Moreover, for any initial allocation, diffusion is larger under this specification as compared to the one we considered earlier in Example 4. □

Note that in this alternative specification of the co-diffusion process, many initial allocations would often converge to a situation in which all the population ends up considering the two technologies (i.e., $\Gamma = \Gamma_{12}$). Even in this case, the optimal seeding strategy is important as it guarantees higher diffusion rates at *each* period of time.

4.2 Alternative Friendship Structure

In our limited attention mechanism, a type n person crosses paths with d_n people at each time and observes their choices. Each person in this group of people is a random draw from the full population. Under this specification, the probability over the union of consideration sets

$$\begin{aligned} \Pr\left(\Gamma_{1n}^{t} \cup \Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\} \\ \Pr\left(\Gamma_{2n}^{t} \cup \Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \\ \Pr\left(\Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}\right) &= \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \end{aligned} \tag{5}$$

is obtained by taking the expectation with respect to the full distribution of adoption rates at time $t-1$. Coupled with information about best-replies, these probabilities characterize the diffusion process at the level of each type. The diffusion process at the population level is just a weighted average of the latter across all types (see Eq. (1) and (3) in Sub-Section 3.1).

In some cases, it might be reasonable to think that a type n person only interacts with people who belong to a few other types or that he interacts with people from all other types but with a frequency that differs from the distribution of types in the population. To capture these possibilities we can postulate that each of the d_n people with whom a type n person crosses paths with is a random draw from a distribution

$$\mathrm{G}_n\left(n'\right) \text{ for } n' = 1, 2, ..., \mathrm{N} \text{ with } \sum\nolimits_{n'=1}^{\mathrm{N}} \mathrm{G}_n\left(n'\right) = 1.$$

Note that this distribution could sharply differ from the distribution of types in the population, $\mathrm{H}(\cdot)$. In addition, note that this friendship structure can easily accommodate homophily by letting $\mathrm{G}_n\left(n'\right)$ be very large for $n' = n$.

From a technical perspective, this alternative specification of the model would affect the distribution of adoption rates that we use to calculate the expectations in Eq. (5) above. Since all our results are initially built at the level of each type and then added up at the population level, the main insights of our work would remain. This richer definition of types would be at the cost of some extra notation.

4.3 Competing Firms

We have focused on the case of a single firm (i.e., a Monopoly) that controls two complementary technologies and aims to allocate them among people in the network to maximize diffusion. We now assume that these two technologies belong to different firms, and each of them controls the distribution of its own technology. We let each firm decide how many units (of its own technology) to initially allocate in the population, taking as given the seeding decision of the other firm. We use Nash equilibrium as our solution concept. Below, we characterize this game and state the precise sense in which any Nash equilibrium is inefficient.[14]

There are two firms: Firm 1 provides technology 1, and Firm 2 provides technology 2. The preferences, network structure, and limited attention mechanism are as described in Section 2.1. Firm 1 decides the fraction of people that initially receive technology 1, i.e., P_1^0. In a similar way, Firm 2 decides about P_2^0. As this is a competitive setting, we assume that these two decisions are made independently. Thus, for each pair of marginals, P_1^0 and P_2^0, the induced joint distribution of initial technologies (P^0) is as follows

$$\mathrm{P}_{00}^0 = \left(1 - \mathrm{P}_1^0\right)\left(1 - \mathrm{P}_2^0\right), \mathrm{P}_{01}^0 = \left(1 - \mathrm{P}_1^0\right)\mathrm{P}_2^0, \mathrm{P}_{10}^0 = \mathrm{P}_1^0\left(1 - \mathrm{P}_2^0\right), \text{ and } \mathrm{P}_{11}^0 = \mathrm{P}_1^0\mathrm{P}_2^0.$$

In particular, this means that the fraction of people who get the two technologies is just the product of the two marginals. (Naturally, these four probabilities add up to 1.)

[14] Our results apply to any Nash equilibrium when there is at least one. So, they are robust to multiplicity of equilibria. Along the analysis we will abstract from showing equilibrium existence; doing so would require imposing extra restrictions (e.g., convexity/concavity assumptions on the profit functions).

Let the induced diffusion process P^t corresponding to the competitive (or non-cooperative $\mathcal{NC}$) initial allocation converge to the joint distribution $\mathrm{P}^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ with corresponding marginals $\mathrm{P}_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ and $\mathrm{P}_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$. By definition, these marginals are the fraction of people that adopt technologies 1 and 2 in the long run, given the initial decision of each firm. We model the profits of these firms at the steady state equilibrium as

$$\Pi_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{b}_1 \mathrm{P}_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) - \mathrm{C}_1\left(\mathrm{P}_1^0\right) \text{ and } \Pi_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{b}_2 \mathrm{P}_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) - \mathrm{C}_2\left(\mathrm{P}_2^0\right)$$

where $\mathrm{b}_1, \mathrm{b}_2 > 0$ are the benefits of a one percent increase in the fraction of people who get each technology, and $\mathrm{C}_1(\cdot)$ and $\mathrm{C}_2(\cdot)$ are the costs of the initial allocation for Firms 1 and 2, respectively. We will refer to this game as the Seeding Game (SG). A Nash equilibrium of the SG is a pair P_1^*, P_2^* that satisfies

$$\mathrm{P}_1^* \in \arg\max_{\mathrm{P}_1^0}\left\{\Pi_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^*\right) : \mathrm{P}_1^0 \in [0,1]\right\} \text{ and } \mathrm{P}_2^* \in \arg\max_{\mathrm{P}_2^0}\left\{\Pi_2^{\mathcal{NC}}\left(\mathrm{P}_1^*, \mathrm{P}_2^0\right) : \mathrm{P}_2^0 \in [0,1]\right\}.$$

Our first result establishes an interesting characteristic of this game.

Lemma 8 *The SG has positive spillovers, i.e.,* $\Pi_i^{\mathcal{NC}}\left(P_1^0, P_2^0\right)$ *increases in* P_j^0 *for each* $i \neq j$.

This result states that each firm benefits when the other firm initially distributes a larger number of technologies among people. From a technical perspective, the proof of Lemma 8 follows from Proposition 4. Specifically, by Proposition 4, we know that $\mathrm{P}_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ increases in P_2^0. Since $\mathrm{b}_1 > 0$, the net benefits of Firm 1 increase with P_2^0. The result follows because P_2^0 does not enter the cost of initial allocation for Firm 1. From an economic perspective, this result derives from the complementarity among the two technologies.

We finally show that any Nash equilibrium of the SG is inefficient in a precise way: The firms would benefit from targeting the *same* consumers for the initial allocation. We also show that if the preferences of the consumers with respect to the two technologies are supermodular (instead of just Q-SPM), then the coordinating strategy also favors them. Thus, in this case, incentives are aligned between the two firms and the consumers. These results build directly on Propositions 4 and 6 in Section 3.

Let P_1^0 and P_2^0 be the initial allocations of Firms 1 and 2, respectively. But suppose now that these two firms coordinate their actions. In particular, suppose they agree on allocating the two technologies in the same hands as much as possible. If they do so, then

$$\mathrm{P}_{11}^0 = \min\left\{\mathrm{P}_1^0, \mathrm{P}_2^0\right\}.$$

Let the induced diffusion process P^t of the cooperative $(\mathcal{C})$ strategy converge to the joint adoption distribution $\mathrm{P}^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ with corresponding marginals $\mathrm{P}_1^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ and $\mathrm{P}_2^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right).$ The profits of these firms at the steady state equilibrium are

$$\Pi_1^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{b}_1\mathrm{P}_1^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) - \mathrm{C}_1\left(\mathrm{P}_1^0\right) \text{ and } \Pi_2^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{b}_2\mathrm{P}_2^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) - \mathrm{C}_2\left(\mathrm{P}_2^0\right).$$

Note that, for each pair P_1^0, P_2^0, it is always the case that

$$\min\left\{\mathrm{P}_1^0, \mathrm{P}_2^0\right\} \geq \mathrm{P}_1^0\mathrm{P}_2^0.$$

Thus, the allocation of technologies of the cooperative strategy is higher (with respect to the UO order) than the one generated by simple competition. By Proposition 4, we get that

$$\mathrm{P}_1^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) \geq \mathrm{P}_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) \text{ and } \mathrm{P}_2^{\mathcal{C}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) \geq \mathrm{P}_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right).$$

It follows that the benefits of the firms are also higher under the coordinated strategy. Since the two strategies initially allocate the same number of technologies, they share the same cost of diffusion. This result holds for any pair of initial allocations, including the equilibrium ones. The next proposition formalizes these claims and states that consumers are also better off when the firms coordinate strategies if their preferences are supermodular (instead of just Q-SPM).

Proposition 9 *For any Nash equilibrium P_1^*,P_2^* of the SG, we have that*

$$\Pi_1^{\mathcal{C}}\left(P_1^*, P_2^*\right) \geq \Pi_1^{\mathcal{NC}}\left(P_1^*, P_2^*\right) \text{ and } \Pi_2^{\mathcal{C}}\left(P_1^*, P_2^*\right) \geq \Pi_2^{\mathcal{NC}}\left(P_1^*, P_2^*\right).$$

Also, assume that, for some $n \in \{1, 2, ..., N\}$, U_n is SPM. Then we have that

$$E\left(V_n \mid P_1^*, P_2^*, \min\left\{P_1^*, P_2^*\right\}\right) \geq E\left(V_n \mid P_1^*, P_2^*, P_1^*P_2^*\right).$$

We finally extend Example 4 to contemplate competition between the two firms and illustrate the statements in Lemma 8 and Proposition 9.

Example 7 Let us use again the specification of Example 4, but this time with the duopoly model. To this end, let $\mathrm{b}_1 = 4$, $\mathrm{b}_2 = 4$, $\mathrm{C}_1\left(\mathrm{P}_1^0\right) = \frac{1}{4}\mathrm{P}_1^0 + \left(\mathrm{P}_1^0\right)^2$ and $\mathrm{C}_2\left(\mathrm{P}_2^0\right) = \frac{1}{4}\mathrm{P}_2^0 + \left(\mathrm{P}_2^0\right)^2$. The other parts of the model remain as before. Thus, mapping M takes the form of

$$\mathrm{P}^t = \mathrm{M}\left(\mathrm{P}^{t-1}\right) = \left(\mathrm{P}_{11}^{t-1}, \mathrm{P}_{11}^{t-1}, \mathrm{P}_{11}^{t-1}\right).$$

It follows, as we showed earlier, that

$$\mathrm{P}_1^t = \mathrm{P}_{11}^0,\ \mathrm{P}_2^t = \mathrm{P}_{11}^0, \text{ and } \mathrm{P}_{11}^t = \mathrm{P}_{11}^0 \text{ for all } t \geq 1.$$

Recall that, for each pair of actions P_1^0, P_2^0, the initial allocation in the SG entails $\mathrm{P}_{11}^0 = \mathrm{P}_1^0\mathrm{P}_2^0$. Therefore, the steady state distribution of adoptions in the non-cooperative setting is given by

$$\mathrm{P}_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{P}_1^0\mathrm{P}_2^0,\ \mathrm{P}_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{P}_1^0\mathrm{P}_2^0, \text{ and } \mathrm{P}_{11}^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{P}_1^0\mathrm{P}_2^0.$$

Finally, the steady state equilibrium profits are

$$\Pi_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = 4\mathrm{P}_1^0\mathrm{P}_2^0 - \left(\frac{1}{4}\mathrm{P}_1^0 + \left(\mathrm{P}_1^0\right)^2\right) \text{ and } \Pi_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = 4\mathrm{P}_1^0\mathrm{P}_2^0 - \left(\frac{1}{4}\mathrm{P}_2^0 + \left(\mathrm{P}_2^0\right)^2\right).$$

Note that $\partial\Pi_i^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)/\partial\mathrm{P}_j^0 = 4\mathrm{P}_i^0 \geq 0$ for $i \neq j$, as Lemma 8 states.

The first order conditions of the maximization problems of the firms are[15]

$$\frac{\partial\Pi_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)}{\partial\mathrm{P}_1^0} = 4\mathrm{P}_2^0 - \frac{1}{4} - 2\mathrm{P}_1^0 = 0 \text{ and } \frac{\partial\Pi_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)}{\partial\mathrm{P}_2^0} = 4\mathrm{P}_1^0 - \frac{1}{4} - 2\mathrm{P}_2^0 = 0$$

Then, the best-reply functions of these firms are

$$\mathrm{P}_1^0 = 2\mathrm{P}_2^0 - \frac{1}{8} \text{ and } \mathrm{P}_2^0 = 2\mathrm{P}_1^0 - \frac{1}{8}.$$

[15] Note that, in this example, $\partial\Pi_i\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)/\partial\mathrm{P}_1^0\partial\mathrm{P}_2^0 = 4 > 0$ for $i = 1, 2$. That is, the SG is a supermodular game. If this feature were robust under other specifications, then we could use Tarski's fixed point theorem to establish equilibrium existence. We do not think this feature immediately extends to the general model.

These best-replies lead to a unique Nash equilibrium

$$\mathrm{P}_1^* = \frac{1}{8} \text{ and } \mathrm{P}_2^* = \frac{1}{8},$$

and steady-state equilibrium profits of the SG are

$$\Pi_1^{\mathcal{NC}}\left(\frac{1}{8}, \frac{1}{8}\right) = \frac{1}{64} \text{ and } \Pi_2^{\mathcal{NC}}\left(\frac{1}{8}, \frac{1}{8}\right) = \frac{1}{64}.$$

If these firms agree on implementing the cooperative allocation, then $\mathrm{P}_{11}^{\mathcal{C}0} = \frac{1}{8}$. In this case, the long run distribution of adoption rates would be

$$\mathrm{P}_1^{\mathcal{C}} = \frac{1}{8},\ \mathrm{P}_2^{\mathcal{C}} = \frac{1}{8}, \text{ and } \mathrm{P}_{11}^{\mathcal{C}} = \frac{1}{8},$$

which leads to the following steady state cooperative payoffs

$$\Pi_1^{\mathcal{C}}\left(\frac{1}{8}, \frac{1}{8}\right) = \frac{29}{64} \text{ and } \Pi_2^{\mathcal{C}}\left(\frac{1}{8}, \frac{1}{8}\right) = \frac{29}{64}.$$

As Proposition 9 states, the profits of the firms are higher under coordination.

Finally, recall that in Example 4 there is only one type of person with a utility function specified as follows

$$\mathrm{U}(0,0) = 10,\ \mathrm{U}(1,0) = 5,\ \mathrm{U}(0,1) = 6, \text{ and } \mathrm{U}(1,1) = 15.$$

Note that this utility function is indeed SPM. Given the previous results, we get that

$$\mathrm{E}\left(\mathrm{V}_n \mid \mathrm{P}_1^*, \mathrm{P}_2^*, \min\{\mathrm{P}_1^*, \mathrm{P}_2^*\}\right) = \frac{15}{8} > \frac{15}{64} = \mathrm{E}\left(\mathrm{V}_n \mid \mathrm{P}_1^*, \mathrm{P}_2^*, \mathrm{P}_1^*\mathrm{P}_2^*\right).$$

As Proposition 9 also states, the weighted average indirect utility of the population is higher when the firms coordinate their diffusion strategies. Thus, incentives are aligned for firms and consumers in this setting. $\square$

5 Conclusion

This paper studies the co-diffusion process of two complementary technologies among people who are connected through a social network. In doing so, we address a question on optimal network seeding that has no counterpart in the case of a single technology. Specifically, we show that (for fixed marginals) adoption rates are maximized over time when the initial fraction of people who get the two technologies is maximized. From a theoretical standpoint, our result relates to the increasing and supermodular (ISPM) stochastic order. From a modelling perspective, our approach connects stochastic choice sets with network or peer effects. We believe this connection opens new possibilities for the empirical identification of random consideration sets. Finally, we show that our results prove useful to establish the source of inefficiency of Nash equilibrium behavior in an extension of the model that accommodates strategic interactions in seeding strategies among two firms.

Appendix: Proofs

Proof of Lemma 1: Let the set of all bivariate Bernoulli distributions $\mathrm{P} = (\mathrm{P}_1, \mathrm{P}_2, \mathrm{P}_{11})$ be

$$\mathcal{P} = \{\mathrm{P} : \mathrm{P}_1 \in [0,1], \mathrm{P}_2 \in [0,1], \text{ and } \mathrm{P}_{11} \in [\max\{0, \mathrm{P}_1 + \mathrm{P}_2 - 1\}, \min\{\mathrm{P}_1, \mathrm{P}_2\}]\}.$$

By construction, $\mathcal{P} \subset \mathbb{R}^3$. We will use the coordinatewise order in the 3-dimensional Euclidean space to compare elements. Note that when we restrict attention to elements in $\mathcal{P}$, the coordinatewise order coincides with the UO order. We want to show that $\mathcal{P}$ (partially) ordered by the UO order is a complete lattice (i.e., each non-empty subset has an infimum and a supremum). Since $\mathcal{P}$ is a compact subset of the Euclidean space with the usual coordinatewise order, it suffices to show that $\mathcal{P}$ (partially) ordered by the UO order is a lattice.

Given $\mathrm{P}' \in \mathbb{R}^3$ and $\mathrm{P}'' \in \mathbb{R}^3$, denote by $\inf(\mathrm{P}', \mathrm{P}'')$ and $\sup(\mathrm{P}', \mathrm{P}'')$ the coordinatewise infimum and supremum of P' and P'' in $\mathbb{R}^3$, that is

$$\begin{aligned} \inf(\mathrm{P}', \mathrm{P}'') &= (\min\{\mathrm{P}'_1, \mathrm{P}''_1\}, \min\{\mathrm{P}'_2, \mathrm{P}''_2\}, \min\{\mathrm{P}'_{11}, \mathrm{P}''_{11}\}) \text{ and} \\ \sup(\mathrm{P}', \mathrm{P}'') &= (\max\{\mathrm{P}'_1, \mathrm{P}''_1\}, \max\{\mathrm{P}'_2, \mathrm{P}''_2\}, \max\{\mathrm{P}'_{11}, \mathrm{P}''_{11}\}). \end{aligned}$$

Given $\mathrm{P}' \in \mathcal{P}$ and $\mathrm{P}'' \in \mathcal{P}$, denote $\inf_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'')$ and $\sup_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'')$ the coordinatewise infimum and supremum of P' and P'' in $\mathcal{P}$. We show below that $\inf_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'') = \inf(\mathrm{P}', \mathrm{P}'')$ and

$$\sup_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'') = (\max\{\mathrm{P}'_1, \mathrm{P}''_1\}, \max\{\mathrm{P}'_2, \mathrm{P}''_2\}, \max\{\max\{\mathrm{P}'_1, \mathrm{P}''_1\} + \max\{\mathrm{P}'_2, \mathrm{P}''_2\} - 1, \mathrm{P}'_{11}, \mathrm{P}''_{11}\}).$$

(Thus, $\mathcal{P}$ is a lattice.) We divide the proof in two parts.

<u>Part 1</u>: Let $\mathrm{P}' \in \mathcal{P}$ and $\mathrm{P}'' \in \mathcal{P}$. We next show that $\inf_{\mathcal{P}}(\mathrm{P}', \mathrm{P}'') = \inf(\mathrm{P}', \mathrm{P}'')$. Since $\inf(\mathrm{P}', \mathrm{P}'')$ is the coordinatewise infimum of P' and P'' in $\mathbb{R}^3$ and $\mathcal{P} \subset \mathbb{R}^3$, we only need to show that $\inf(\mathrm{P}', \mathrm{P}'') \in \mathcal{P}$. Since $\mathrm{P}' \in \mathcal{P}$ and $\mathrm{P}'' \in \mathcal{P}$

$$\mathrm{P}'_1 \in [0,1], \mathrm{P}'_2 \in [0,1], \mathrm{P}''_1 \in [0,1], \text{ and } \mathrm{P}''_2 \in [0,1]$$

Thus,

$$\min\{\mathrm{P}'_1, \mathrm{P}''_1\} \in [0,1] \text{ and } \min\{\mathrm{P}'_2, \mathrm{P}''_2\} \in [0,1].$$

We next show that

$\min\{\mathrm{P}'_{11},\mathrm{P}''_{11}\}\in[\max\{0,\min\{\mathrm{P}'_1,\mathrm{P}''_1\}+\min\{\mathrm{P}'_2,\mathrm{P}''_2\}-1\},\min\{\min\{\mathrm{P}'_1,\mathrm{P}''_1\},\min\{\mathrm{P}'_2,\mathrm{P}''_2\}\}]$.

We further divide Part 1 in two parts.

Part 1(i): We next show that $\min\{\mathrm{P}'_{11},\mathrm{P}''_{11}\}\geq\max\{0,\min\{\mathrm{P}'_1,\mathrm{P}''_1\}+\min\{\mathrm{P}'_2,\mathrm{P}''_2\}-1\}$. Since $\mathrm{P}'\in\mathcal{P}$ and $\mathrm{P}''\in\mathcal{P}$, then $\mathrm{P}'_{11}\geq 0$ and $\mathrm{P}''_{11}\geq 0$. Thus, $\min\{\mathrm{P}'_{11},\mathrm{P}''_{11}\}\geq 0$. In addition,

$$\mathrm{P}'_{11}\geq\mathrm{P}'_1+\mathrm{P}'_2-1 \text{ and } \mathrm{P}''_{11}\geq\mathrm{P}''_1+\mathrm{P}''_2-1.$$

It follows that

$$\min\{\mathrm{P}'_{11},\mathrm{P}''_{11}\}\geq\min\{\mathrm{P}'_1,\mathrm{P}''_1\}+\min\{\mathrm{P}'_2,\mathrm{P}''_2\}-1.$$

The claim follows by combining the last two results.

Part 1(ii): We next show that $\min\{\mathrm{P}'_{11},\mathrm{P}''_{11}\}\leq\min\{\min\{\mathrm{P}'_1,\mathrm{P}''_1\},\min\{\mathrm{P}'_2,\mathrm{P}''_2\}\}$. Since $\mathrm{P}'\in\mathcal{P}$ and $\mathrm{P}''\in\mathcal{P}$, then

$$\mathrm{P}'_{11}\leq\mathrm{P}'_1,\mathrm{P}'_{11}\leq\mathrm{P}'_2,\mathrm{P}''_{11}\leq\mathrm{P}''_1, \text{ and } \mathrm{P}''_{11}\leq\mathrm{P}''_2.$$

Thus,

$$\min\{\mathrm{P}'_{11},\mathrm{P}''_{11}\}\leq\min\{\mathrm{P}'_1,\mathrm{P}'_2,\mathrm{P}''_1,\mathrm{P}''_2\}=\min\{\min\{\mathrm{P}'_1,\mathrm{P}''_1\},\min\{\mathrm{P}'_2,\mathrm{P}''_2\}\}.$$

This last claim completes Part 1.

Part 2: Let $\mathrm{P}'\in\mathcal{P}$ and $\mathrm{P}''\in\mathcal{P}$. We next show that

$\sup_{\mathcal{P}}(\mathrm{P}',\mathrm{P}'')=(\max\{\mathrm{P}'_1,\mathrm{P}''_1\},\max\{\mathrm{P}'_2,\mathrm{P}''_2\},\max\{\max\{\mathrm{P}'_1,\mathrm{P}''_1\}+\max\{\mathrm{P}'_2,\mathrm{P}''_2\}-1,\mathrm{P}'_{11},\mathrm{P}''_{11}\})$.

Since $\mathrm{P}'\in\mathcal{P}$ and $\mathrm{P}''\in\mathcal{P}$

$$\mathrm{P}'_1\in[0,1],\mathrm{P}'_2\in[0,1],\mathrm{P}''_1\in[0,1], \text{ and } \mathrm{P}''_2\in[0,1]$$

Thus,

$$\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}\in[0,1] \ \text{ and } \ \max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}\in[0,1].$$

Note also that any other upper bound $\widehat{\mathrm{P}}$ for P' and P'', would require $\widehat{\mathrm{P}}_1\geq\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}$ and $\widehat{\mathrm{P}}_2\geq\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}$. Thus, we only need to show that

$$\max\left\{\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}+\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}-1,\mathrm{P}_{11}',\mathrm{P}_{11}''\right\}\leq\min\left\{\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\},\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}\right\}.$$

(The fact that this inequality suffices follows by the next argument. First, since $\mathrm{P}_{11}'\geq 0$ and $\mathrm{P}_{11}''\geq 0$,

$$\max\left\{\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}+\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}-1,\mathrm{P}_{11}',\mathrm{P}_{11}''\right\}\geq 0.$$

Also, by construction,

$$\max\left\{\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}+\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}-1,\mathrm{P}_{11}',\mathrm{P}_{11}''\right\}\geq\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}+\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}-1.$$

Second, as we mentioned above, any other upper bound $\widehat{\mathrm{P}}$ for P' and P'' would require $\widehat{\mathrm{P}}_1\geq\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}$ and $\widehat{\mathrm{P}}_2\geq\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}$ which would push the lower bound for the last term $\widehat{\mathrm{P}}_{11}$ up.)

Since $\mathrm{P}'\in\mathcal{P}$ and $\mathrm{P}''\in\mathcal{P}$

$$\mathrm{P}_{11}'\leq\mathrm{P}_1',\mathrm{P}_{11}'\leq\mathrm{P}_2',\mathrm{P}_{11}''\leq\mathrm{P}_1'', \text{ and } \mathrm{P}_{11}''\leq\mathrm{P}_2''.$$

Thus,

$$\max\left\{\mathrm{P}_{11}',\mathrm{P}_{11}''\right\}\leq\min\left\{\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\},\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}\right\}.$$

The fact that

$$\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}+\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}-1\leq\min\left\{\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\},\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}\right\}.$$

follows because $\max\left\{\mathrm{P}_1',\mathrm{P}_1''\right\}\in[0,1]$ and $\max\left\{\mathrm{P}_2',\mathrm{P}_2''\right\}\in[0,1]$. ■

The proof of Lemma 2 relies on three intermediate results that we cover in the next three lemmas.

Lemma 10 *Let* $S_1 = \sum_{i=1}^{d} x_{1i}$, $S_2 = \sum_{i=1}^{d} x_{2i}$, $T_1 = \sum_{i=1}^{d} y_{1i}$, *and* $T_2 = \sum_{i=1}^{d} y_{2i}$. *In addition, suppose that each* (x_{1i}, x_{2i}) *is an independent random draw from a bivariate Bernoulli distribution* $\overline{P}$ *and each* (y_{1i}, y_{2i}) *is an independent random draw from a bivariate Bernoulli distribution* $\underline{P}$, *for* $i = 1, 2, ..., d$. *If* $\overline{P} \geq_{uo} \underline{P}$, *then* $\Pr(S_1 \geq k_1, S_2 \geq k_2) \geq \Pr(T_1 \geq k_1, T_2 \geq k_2)$ *for any* k_1, k_2.

Proof of Lemma 10: The result follows by induction on d. Let $d = 1$. Note that

$$\Pr(x_{11} \geq k_1, x_{21} \geq k_2) \geq \Pr(y_{11} \geq k_1, y_{21} \geq k_2)$$

follows by definition, given that $\overline{\mathrm{P}} \geq_{uo} \underline{\mathrm{P}}$. Suppose the claim is true for d, so that

$$\Pr(S_1 \geq k_1, S_2 \geq k_2) \geq \Pr(T_1 \geq k_1, T_2 \geq k_2).$$

We need to show that the last inequality holds for $d+1$. To this end, note that

$$\begin{aligned}
& \Pr(S_1 + x_{1d+1} \geq k_1, S_2 + x_{2d+1} \geq k_2) \\
= & \sum\nolimits_{a,b} \Pr(S_1 + x_{1d+1} \geq k_1, S_2 + x_{2d+1} \geq k_2 \mid x_{1d+1} = a, x_{2d+1} = b) \Pr(x_{1d+1} = a, x_{2d+1} = b) \\
\geq & \sum\nolimits_{a,b} \Pr(T_1 \geq k_1 - x_{1d+1}, T_2 \geq k_2 - x_{2d+1} \mid x_{1d+1} = a, x_{2d+1} = b) \Pr(x_{1d+1} = a, x_{2d+1} = b) \\
= & \sum\nolimits_{a,b} \Pr(T_1 + x_{1d+1} \geq k_1, T_2 + x_{2d+1} \geq k_2 \mid x_{1d+1} = a, x_{2d+1} = b) \Pr(x_{1d+1} = a, x_{2d+1} = b) \\
= & \Pr(T_1 + x_{1d+1} \geq k_1, T_2 + x_{2d+1} \geq k_2) \\
= & \sum\nolimits_{c,d} \Pr(x_{1d+1} \geq k_1 - T_1, x_{2d+1} \geq k_2 - T_2 \mid T_1 = c, T_2 = d) \Pr(T_1 = c, T_2 = d) \\
\geq & \sum\nolimits_{c,d} \Pr(y_{1d+1} \geq k_1 - T_1, y_{2d+1} \geq k_2 - T_2 \mid T_1 = c, T_2 = d) \Pr(T_1 = c, T_2 = d) \\
= & \Pr(T_1 + y_{1d+1} \geq k_1, T_2 + y_{2d+1} \geq k_2)
\end{aligned}$$

where the inequalities hold because $\overline{\mathrm{P}} \geq_{uo} \underline{\mathrm{P}}$ and (x_{1d+1}, x_{2d+1}) and (y_{1d+1}, y_{2d+1}) are independent of (S_1, S_2) and (T_1, T_2), respectively. In addition, the equalities hold by the Law of Total Probability. ■

Lemma 11 *Let* $S_1 = \sum_{i=1}^{d} x_{1i}$, $S_2 = \sum_{i=1}^{d} x_{2i}$, $T_1 = \sum_{i=1}^{d} y_{1i}$, *and* $T_2 = \sum_{i=1}^{d} y_{2i}$. *Assume* $\Pr(S_1 \geq k_1, S_2 \geq k_2) \geq \Pr(T_1 \geq k_1, T_2 \geq k_2)$ *for any* k_1, k_2. *Then,* *1.* $E\{\pi_{1n}(S_1)\,\pi_{2n}(S_2)\} \geq E\{\pi_{1n}(T_1)\,\pi_{2n}(T_2)\}$

2. $E\{\pi_{1n}(S_1)\} \geq E\{\pi_{1n}(T_1)\}$

3. $E\{\pi_{2n}(S_2)\} \geq E\{\pi_{2n}(T_2)\}$

Proof of Lemma 11: Recall that $\pi_{1n}(\cdot)$ and $\pi_{2n}(\cdot)$ are non-negative and increasing functions. Then, Point 1 follows by Theorem 3.3.16 in Müller and Stoyan, 2002. In addition, Points 2 and 3 follow by the standard characterization (in terms of the expectation of increasing functions) of first-order stochastic dominance. ■

Lemma 12 *$BR_{12n} \geq BR_{1n}+ BR_{2n}$ for each n = 1, 2, ..., N.*

Proof of Lemma 12: Note that BR_{1n} is either $(0,0)$ or $(1,0)$ and BR_{2n} is either $(0,0)$ or $(0,1)$. In addition, $\mathrm{BR}_{12n} \in \{(0,0),(1,0),(0,1),(1,1)\}$. To prove this lemma we consider four cases.

Case 1: Let $\mathrm{BR}_{1n} = (0,0)$ and $\mathrm{BR}_{2n} = (0,0)$. Then, $\mathrm{BR}_{12n} \geq \mathrm{BR}_{1n}+ \mathrm{BR}_{2n} = (0,0)$.

Case 2: Let $\mathrm{BR}_{1n} = (0,0)$ and $\mathrm{BR}_{2n} = (0,1)$. Since the utility is Q-SPM, we know that

$$\mathrm{U}_n(0,1) > \mathrm{U}_n(0,0) \Longrightarrow \mathrm{U}_n(1,1) > \mathrm{U}_n(1,0).$$

The left hand side holds here as $\mathrm{BR}_{2n} = (0,1)$ and there are no ties. Thus, BR_{12n} is either $(0,1)$ or $(1,1)$. In both cases, $\mathrm{BR}_{12n} \geq \mathrm{BR}_{1n}+ \mathrm{BR}_{2n} = (0,1)$.

Case 3: Let $\mathrm{BR}_{1n} = (1,0)$ and $\mathrm{BR}_{2n} = (0,0)$. This case is similar to the last one, thus we omit it.

Case 4: Let $\mathrm{BR}_{1n} = (1,0)$ and $\mathrm{BR}_{2n} = (0,1)$. Since the utility is Q-SPM, we know that

$$\mathrm{U}_n(0,1) > \mathrm{U}_n(0,0) \Longrightarrow \mathrm{U}_n(1,1) > \mathrm{U}_n(1,0).$$

The left hand side holds here as $\mathrm{BR}_{2n} = (0,1)$ and there are no ties. Since the utility is Q-SPM, we know that

$$\mathrm{U}_n(1,0) > \mathrm{U}_n(0,0) \Longrightarrow \mathrm{U}_n(1,1) > \mathrm{U}_n(0,1).$$

The left hand side holds here as $\mathrm{BR}_{1n} = (1,0)$ and there are no ties. Given the last two results, $\mathrm{BR}_{2n} = (1,1) = \mathrm{BR}_{1n} + \mathrm{BR}_{2n} = (1,1)$. ∎

Proof of Lemma 2: Let $\overline{\mathrm{P}}^{t-1} \geq_{uo} \underline{\mathrm{P}}^{t-1}$. By Lemmas 10 and 11, we have that the three probabilities in expression (2) are higher when the consideration sets are generated by $\overline{\mathrm{P}}^{t-1}$ as compared to $\underline{\mathrm{P}}^{t-1}$. By Lemma 12, $\mathrm{BR}_{12n} \geq \mathrm{BR}_{1n} + \mathrm{BR}_{2n}$. As a consequence, all terms in the 3×3 matrix in expression (1) are (weakly) positive for each $n = 1, 2, ..., \mathrm{N}$. The three lemmas together imply that the probabilities P^t_{1n}, P^t_{2n}, and P^t_{11n} generated by $\overline{\mathrm{P}}^{t-1}$ are larger than the ones generated by $\underline{\mathrm{P}}^{t-1}$. Then, the claim holds as

$$\mathrm{P}^t_1 = \sum_{n=1}^{\mathrm{N}} \mathrm{P}^t_{1n}\mathrm{H}(n),\ \mathrm{P}^t_2 = \sum_{n=1}^{\mathrm{N}} \mathrm{P}^t_{2n}\mathrm{H}(n),\ \text{and}\ \mathrm{P}^t_{11} = \sum_{n=1}^{\mathrm{N}} \mathrm{P}^t_{11n}\mathrm{H}(n)$$

are simple weighted averages of P^t_{1n}, P^t_{2n}, and P^t_{11n} across $n = 1, 2, ..., \mathrm{N}$. ∎

Proof of Theorem 3: By construction, M: $\mathcal{P} \rightarrow \mathcal{P}$. By Lemma 1, $\mathcal{P}$ ordered by the UO order is a complete lattice. By Lemma 2, mapping M is increasing. Thus, by Tarski′s fixed point theorem, the set of fixed points is a non-empty complete lattice. ∎

Proof of Proposition 4: By Lemma 2, we know that, for each $t > 0$, if $\overline{\mathrm{P}}^{t-1} \geq_{uo} \underline{\mathrm{P}}^{t-1}$ then $\overline{\mathrm{P}}^{t} \geq_{uo} \underline{\mathrm{P}}^{t}$. Let $t = 1$, we have that $\overline{\mathrm{P}}^{0} \geq_{uo} \underline{\mathrm{P}}^{0}$ implies $\overline{\mathrm{P}}^{1} \geq_{uo} \underline{\mathrm{P}}^{1}$. Applying Lemma 2 $t-1$ more times we get $\overline{\mathrm{P}}^{t} \geq_{uo} \underline{\mathrm{P}}^{t}$. This shows the first claim.

The second claim follows from the previous one and the fact that the ISPM order is closed with respect to weak convergence (see Müller and Stoyan, 2002, Theorem 3.9.12). ∎

Proof of Proposition 5: We first show that, for each P^{t-1}, P^t increases (w.r.t. the UO order) in d and $\pi_1(\cdot), \pi_2(\cdot)$. It is clear that, for each P^{t-1}, the three lines in (2) increase in

d and π_1, π_2. Then, the probabilities P_{1n}^t, P_{2n}^t, and P_{11n}^t increase in d and π_1, π_2. The claim holds as

$$\mathrm{P}_1^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{1n}^t \mathrm{H}(n), \ \mathrm{P}_2^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{2n}^t \mathrm{H}(n), \text{ and } \mathrm{P}_{11}^t = \sum\nolimits_{n=1}^{\mathrm{N}} \mathrm{P}_{11n}^t \mathrm{H}(n)$$

are simple weighted averages of P_{1n}^t, P_{2n}^t, and P_{11n}^t across $n = 1, 2, ..., \mathrm{N}$.

Fix some P^0 and let $\overline{\mathrm{P}}^t$ and $\underline{\mathrm{P}}^t$ be generated by $\overline{d}, \overline{\pi}_1, \overline{\pi}_2$ and $\underline{d}, \underline{\pi}_1, \underline{\pi}_2$. By our initial result, we know that $\overline{\mathrm{P}}^1 \geq_{uo} \underline{\mathrm{P}}^1$. Let $\mathrm{M}^{\overline{d},\overline{\pi}_1,\overline{\pi}_2}\left(\mathrm{P}^{t-1}\right)$ and $\mathrm{M}^{\underline{d},\underline{\pi}_1,\underline{\pi}_2}\left(\mathrm{P}^{t-1}\right)$ be mapping M in (3) for $d, \pi_1, \pi_2 = \overline{d}, \overline{\pi}_1, \overline{\pi}_2$ and $d, \pi_1, \pi_2 = \underline{d}, \underline{\pi}_1, \underline{\pi}_2$, respectively. We have that

$$\overline{\mathrm{P}}^2 = \mathrm{M}^{\overline{d},\overline{\pi}_1,\overline{\pi}_2}\left(\overline{\mathrm{P}}^1\right) \geq_{uo} \mathrm{M}^{\overline{d},\overline{\pi}_1,\overline{\pi}_2}\left(\underline{\mathrm{P}}^1\right) \geq_{uo} \mathrm{M}^{\underline{d},\underline{\pi}_1,\underline{\pi}_2}\left(\underline{\mathrm{P}}^1\right) = \underline{\mathrm{P}}^2$$

where the first inequality follows by Lemma 2 and the second one by our initial claim. The first part of the proposition follows by applying this idea $t-2$ more times.

The second claim follows from the previous one and the fact that the ISPM order is closed with respect to weak convergence (see Müller and Stoyan, 2002, Theorem 3.9.12).

The last part of the proposition follows by Theorem 2.5.2 in Topkis (1998). ■

Proof of Proposition 6: Let $\mathrm{V}_{\bullet n} = \mathrm{U}_n(\mathrm{BR}_{\bullet n})$, for $\bullet = 0, 1, 2, 12$, be the indirect utility for a type n person under different consideration sets. The expected indirect utility of this type of person in period of time t can be expressed as follows

$$\begin{aligned} \mathrm{E}\left(\mathrm{V}_n \mid \mathrm{P}^{t-1}\right) \;=\;& \mathrm{V}_{0n} + (\mathrm{V}_{1n} - \mathrm{V}_{0n}) \Pr\left(\Gamma_{1n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) + (\mathrm{V}_{2n} - \mathrm{V}_{0n}) \Pr\left(\Gamma_{2n}^t \cup \Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right) \\ & + (\mathrm{V}_{12n} + \mathrm{V}_{0n} - \mathrm{V}_{1n} - \mathrm{V}_{2n}) \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}\right). \end{aligned}$$

Let $\overline{\mathrm{P}}^{t-1} \geq_{uo} \underline{\mathrm{P}}^{t-1}$. By Lemmas 10 and 11, we have that the three probabilities in the last expression are higher when the consideration sets are generated by $\overline{\mathrm{P}}^{t-1}$ as compared to $\underline{\mathrm{P}}^{t-1}$. Also, by the optimization principle we have that $\mathrm{V}_{1n} - \mathrm{V}_{0n} \geq 0$ and $\mathrm{V}_{2n} - \mathrm{V}_{0n} \geq 0$. Thus, we only need to show that if U_n is SPM then

$$\mathrm{V}_{12n} + \mathrm{V}_{0n} - \mathrm{V}_{1n} - \mathrm{V}_{2n} \geq 0.$$

There are two cases to consider.

Case 1: Assume that either BR_{1n} or BR_{2n} is $(0,0)$. In this case the result follows as, by the optimization principle, $\mathrm{V}_{12n}-\mathrm{V}_{1n} \geq 0$ and $\mathrm{V}_{12n}-\mathrm{V}_{2n} \geq 0$.

Case 2: Assume $\mathrm{BR}_{1n} = (1,0)$ and $\mathrm{BR}_{2n} = (0,1)$. In this case, by Lemma 12 we know that $\mathrm{BR}_{12n} = (1,1)$. By SPM we know that

$$\mathrm{V}_{12n} + \mathrm{V}_{0n} = \mathrm{U}_n(1,1) + \mathrm{U}_n(0,0) \geq \mathrm{U}_n(1,0) + \mathrm{U}_n(0,1) = \mathrm{V}_{1n} + \mathrm{V}_{2n}$$

The rest of the proof is similar to the proof of Proposition 5, thus we omit it. ■

Proof of Lemma 7: By Lemma 1, $\mathcal{P}$ (partially) ordered by the UO order is a complete lattice. The first claim follows as the product of complete lattices is a complete lattice.

Given the proof of Lemma 2, to show the monotonicity of M^* we only need to prove that Q_n^t increases with P^{t-1} and Q_n^{t-1} (with respect to the UO order). Let us first consider

$$\Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) = \begin{cases} \Pr\left(\Gamma_{12n}^{t-1}\right) + \\ \left[\Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) - \Pr\left(\Gamma_{12n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} + \\ \left[\Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) - \Pr\left(\Gamma_{12n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\} + \\ \left[1 - \Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) - \Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) + \Pr\left(\Gamma_{12n}^{t-1}\right)\right] \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \end{cases}$$

Note that, by definition,

$$\begin{aligned} &\mathrm{E}\left\{\left(1-\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right)\left(1-\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right)\right\} + \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\left(1-\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right)\right\} + \\ &\mathrm{E}\left\{\left(1-\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} + \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} = 1. \end{aligned}$$

Re-arranging and simplifying terms, we obtain

$$1-\mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\}-\mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\}+\mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} = \mathrm{E}\left\{\left(1-\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right)\left(1-\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right)\right\} \geq 0.$$

Thus, keeping fixed $\Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right)$ and $\Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1}\right)$, we have that

$$\partial \Pr\left(\Gamma_{12n}^t \mid \mathrm{P}^{t-1}, \mathrm{Q}_n^{t-1}\right) / \partial \Pr\left(\Gamma_{12n}^{t-1}\right) = 1-\mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\}-\mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\}+\mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \geq 0.$$

In addition,

$$\partial \Pr\left(\Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}, \mathrm{Q}_{n}^{t-1}\right) / \partial \Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) = \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\right\} - \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \geq 0$$
$$\partial \Pr\left(\Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}, \mathrm{Q}_{n}^{t-1}\right) / \partial \Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1}\right) = \mathrm{E}\left\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} - \mathrm{E}\left\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\right\} \geq 0.$$

Thus, $\Pr\left(\Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}, \mathrm{Q}_{n}^{t-1}\right)$ increases in Q_{n}^{t-1}. The fact that $\Pr\left(\Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}, \mathrm{Q}_{n}^{t-1}\right)$ is also increasing in P^{t-1} holds as (by Lemma 11) an increase in P^{t-1} positively affects $\mathrm{E}\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\}$, $\mathrm{E}\{\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\}$, and $\mathrm{E}\{\pi_{1n}\left(\mathrm{S}_{1n}^{t-1}\right)\pi_{2n}\left(\mathrm{S}_{2n}^{t-1}\right)\}$, and these three expectations are pre-multiplied by positive terms. Thus, $\Pr\left(\Gamma_{12n}^{t} \mid \mathrm{P}^{t-1}, \mathrm{Q}_{n}^{t-1}\right)$ increases when either P^{t-1} or Q_{n}^{t-1} increase with respect to the UO order. The fact that $\Pr\left(\Gamma_{1n}^{t-1} \cup \Gamma_{12n}^{t-1} \mid \mathrm{P}^{t-1}, \mathrm{Q}_{n}^{t-1}\right)$ and $\Pr\left(\Gamma_{2n}^{t-1} \cup \Gamma_{12n}^{t-1} \mid \mathrm{P}^{t-1}, \mathrm{Q}_{n}^{t-1}\right)$ increase when either P^{t-1} or Q_{n}^{t-1} increase (with respect to the UO order) follows by similar arguments. ■

Proof of Lemma 8: Recall that in the SG each pair of marginals P_1^0, P_2^0 selected by the firms induces the next joint distribution of initial technologies (P^0)

$$\mathrm{P}_{00}^0 = \left(1 - \mathrm{P}_1^0\right)\left(1 - \mathrm{P}_2^0\right), \mathrm{P}_{01}^0 = \left(1 - \mathrm{P}_1^0\right)\mathrm{P}_2^0, \mathrm{P}_{10}^0 = \mathrm{P}_1^0\left(1 - \mathrm{P}_2^0\right), \text{ and } \mathrm{P}_{11}^0 = \mathrm{P}_1^0\mathrm{P}_2^0.$$

Let the induced diffusion process P^t converge to the joint distribution $\mathrm{P}^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ with corresponding marginals $\mathrm{P}_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ and $\mathrm{P}_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$. Note that, if P_1^0 increases, then $\mathrm{P}_2^0 = \mathrm{P}_{01}^0 + \mathrm{P}_{11}^0$ remains at the same level and P_{11}^0 increases. Thus, when Firm 1 increases its initial allocation, the induced joint distribution of initial technologies P^0 increases with respect to the UO order. Thus, by Proposition 4, we have that $\mathrm{P}_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ and $\mathrm{P}_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ increase. In turn, this means that

$$\Pi_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) = \mathrm{b}_2\mathrm{P}_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right) - \mathrm{C}_2\left(\mathrm{P}_2^0\right)$$

moves up. Thus, $\Pi_2^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ increases in P_1^0. A similar argument could be used to show that $\Pi_1^{\mathcal{NC}}\left(\mathrm{P}_1^0, \mathrm{P}_2^0\right)$ increases in P_2^0. ■

Proof of Proposition 9: Let P_1^*, P_2^* be a Nash equilibrium of the SG. This pair of choices generate an initial allocation P^0 characterized by

$$\mathrm{P}_1^0 = \mathrm{P}_1^*, \mathrm{P}_2^0 = \mathrm{P}_2^*, \text{ and } \mathrm{P}_{11}^0 = \mathrm{P}_1^*\mathrm{P}_2^*.$$

Let the induced diffusion process P^t converge to the joint distribution $P^{\mathcal{NC}}(P_1^*, P_2^*)$ with corresponding marginals $P_1^{\mathcal{NC}}(P_1^*, P_2^*)$ and $P_2^{\mathcal{NC}}(P_1^*, P_2^*)$.

Consider next a coordinated initial allocation P^0 with

$$P_1^0 = P_1^*,\ P_2^0 = P_2^*, \text{ and } P_{11}^0 = \min\{P_1^*, P_2^*\}.$$

Let the induced diffusion process P^t converge to the joint distribution $P^{\mathcal{C}}(P_1^*, P_2^*)$ with corresponding marginals $P_1^{\mathcal{C}}(P_1^*, P_2^*)$ and $P_2^{\mathcal{C}}(P_1^*, P_2^*)$. Note that the non-cooperative and cooperative diffusion process start with the same marginals P_1^*, P_2^* but differ regarding the joint allocation. In particular,

$$\min\{P_1^*, P_2^*\} \geq P_1^* P_2^*.$$

Thus, $P^{\mathcal{C}}(P_1^*, P_2^*)$ is higher than $P^{\mathcal{NC}}(P_1^*, P_2^*)$ with respect to the UO order. It follows by Proposition 4, that

$$P_1^{\mathcal{C}}(P_1^*, P_2^*) \geq P_1^{\mathcal{NC}}(P_1^*, P_2^*) \text{ and } P_2^{\mathcal{C}}(P_1^*, P_2^*) \geq P_2^{\mathcal{NC}}(P_1^*, P_2^*).$$

Since $b_1 > 0$, we get that

$$\Pi_1^{\mathcal{C}}(P_1^*, P_2^*) = b_1 P_1^{\mathcal{C}}(P_1^*, P_2^*) - C_1(P_1^*) \geq \Pi_1^{\mathcal{NC}}(P_1^*, P_2^*) = b_1 P_1^{\mathcal{NC}}(P_1^*, P_2^*) - C_1(P_1^*).$$

By a similar argument, we get that $\Pi_2^{\mathcal{C}}(P_1^*, P_2^*) \geq \Pi_2^{\mathcal{NC}}(P_1^*, P_2^*)$.

The last claim in Proposition 9 follows directly from Proposition 6 and the fact that the cooperative initial allocation is higher than the competitive one (with respect to the UO order). ■

References

[1] Amin Rahimian, M., and A. Jadbabaie (2017): "Bayesian Learning Without Recall," *IEEE Transactions on Signal and Information Processing over Networks*, 3, 592–606.

[2] Amir, R., and N. Lazzati (2016): "Endogenous Information Acquisition in Bayesian Games with Strategic Complementarities," *Journal of Economic Theory*, 163, 684–698.

[3] Aral, S., L. Muchnik, and A. Sundararajan (2013): "Engineering Social Contagions: Optimal Network Seeding in the Presence of Homophily," *Network Science*, 1, 125–153.

[4] Athey, S., and J. Levin (2001): "The Value of Information in Monotone Decision Problems," Working Paper, Stanford University.

[5] Athey, S., and S. Stern (1998): "An Empirical Framework for Testing Theories about Complementarity in Organizational Design," Working paper.

[6] Bala, V., and S. Goyal (1998): "Learning from Neighbors," *Review of Economic Studies*, 65, 595–621.

[7] Bass, F. (1969): "A New Product Growth Model for Consumer Durables," *Management Science*, 15, 215–27.

[8] Bayus, B., N. Kim, and A. Shocker (2000): "Growth Models for Multi-Product Interactions: Current Status and New Directions," New-Product Diffusion Models, Ed. by Mahajan, V., E. Muller, and Y. Wind, Kluwer Academic Publishers, Boston, MA.

[9] Blume, L. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5, 387–424.

[10] Blume, L. (1995): "The Statistical Mechanics of Best-Response Strategy Revision," *Games and Economic Behavior*, 11, 111–45.

[11] Chen, Y.-J., Y. Zenou, and J. Zhou (2018): "Multiple Activities in Networks," *American Economic Journal: Microeconomics*, 3, 34-85.

[12] Demuynck, T., and C. Seel (2018): "Revealed Preference with Limited Consideration," *American Economic Journal: Microeconomics,* 10, 102–131.

[13] Dewan, S., D. Ganley, and K. Kraemer (2010): "Complementarities in the Diffusion of Personal Computers and the Internet: Implications for the Global Digital Divide," *Information Systems Research*, 21, 925–940.

[14] Dziewulski, P., and J. Quah (2014): "Testing for Production with Complementarities," Working Paper.

[15] Echenique, F. (2003): "Mixed Equilibria in Games of Strategic Complementarities," *Economic Theory*, 22, 33–44.

[16] Fox, J., and N. Lazzati (2017): "A Note on Identification of Discrete Choice Models for Bundles and Binary Games," *Quantitative Economics*, 8, 1021-1036.

[17] Galeotti, A., and S. Goyal (2009): "Influencing the Influencers: A theory of Strategic Diffusion," *RAND Journal of Economics*, 40, 509–532.

[18] Golub, B., and M. Jackson (2010): "Naive Learning in Social Networks: Convergence, Influence, and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2, 112–149.

[19] Graham, B., G. Imbens, and G. Ridder (2014): "Complementarity and Aggregate Implications of Assortative Matching: A Nonparametric Analysis," *Quantitative Economics*, 5, 29–66.

[20] Jackson, M., and B. Rogers (2007): "Relating Network Structure to Diffusion Properties through Stochastic Dominance," *Advances in Theoretical Economics*, 7, 1–13.

[21] Jackson, M., B. Rogers, and Y. Zenou (2016): "Networks: An Economic Perspective," *Oxford Handbook of Social Network Analysis, R. Light and J. Moody (Eds.), Oxford University Press.*

[22] Jackson, M., and L. Yariv (2007): "Diffusion of Behavior and Equilibrium Properties in Network Games," *American Economic Review*, 97, 92–98.

[23] Kamae, T., U. Krengel, and G. O'Brien (1977): "Stochastic Inequalities on Partially Ordered Spaces," *The Annals of Probability*, 5, 899–912.

[24] Levin, J. (2001): "Information and the Market for Lemons," *RAND Journal of Economics*, 32, 657–666.

[25] Manzini, P., and M. Mariotti (2014): "Stochastic Choice and Consideration Sets," *Econometrica*, 82, 1153–1176.

[26] Meyer, M., and B. Strulovici (2012): "Increasing Interdependence of Multivariate Distributions," *Journal of Economic Theory,* 147, 1460–1489.

[27] Meyer, M., and B. Strulovici (2015): "Beyond Correlation: Measuring Interdependence Through Complementarities," Working Paper.

[28] Müller, A., and D. Stoyan (2002): Comparison Methods for Stochastic Models and Risk, Wiley Series in Probability and Statistics.

[29] Niculescu, M., and S. Whang (2012): "Codiffusion of Wireless Voice and Data Services: An Empirical Analysis of the Japanese Mobile Telecommunications Market," *Information Systems Research*, 23, 260–279.

[30] Scarsini, M. (1998): "Multivariate Convex Orderings, Dependence, and Stochastic Equality," *Journal of Applied Probability*, 35, 93–103.

[31] Topkis, D. (1998): Supermodularity and Complementarity, Princeton University Press.

[32] Vega-Redondo, F. (2007): Complex Social Networks. Cambridge: Cambridge University Press.Valente

[33] Westland, J., J. Xing, H. Xiao, and S. Shan (2016): "Substitutes, Complements and Network Effects in Instant Messaging Services," *Networks and Spatial Economics*, 16, 525–543.

[34] Young, P. (2009): "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *American Economic Review*, 99, 1899–1924.