Optimal Coordination and Pricing of a Network under Incomplete Information

Masaki Aoyagi[*]
Osaka University

September 12, 2011

Abstract

A monopolist sells a good whose value depends on the set (network) of buyers who adopt it as well as on their private types. This paper studies the seller's revenue maximization in this problem when he coordinates the buyers' adoption decisions based on their reported types. We characterize ex post implementable sales schemes, and identify the conditions under which the revenue maximizing scheme has the properties that a larger network is more affordable than a smaller network, and that the network size is maximized subject to the participation constraints.

Key words: adoption externalities, strategy-proof, revenue maximization.
Journal of Economic Literature Classification Numbers: C72, D82.

1 Introduction

Goods have network externalities when their value to any consumer depends on the consumption decision of other consumers. A classical example of a good with network externalities, or more simply a network good, is a telecommunication device whose value depends directly on the number of other people using the device. Other leading examples of network goods include the operating system (OS) of PC's, fuel-cell vehicles, social networking services, industrial parks, and so on. The nature of network externalities may be purely physical as in the case of the telecommunication device, but may also be market-based or psychological. Market-based externalities arise when more users of a good induces the market to provide complementary goods

[*]ISER, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan.

that enhance the value of the good. More users of a fuel-cell vehicle, for example, encourages entry into the market of charge stations, which leads to the increased value of such vehicles. On the other hand, much of bandwagon consumption in the fashion, toy and electronic industries can be explained through psychological externalities where consumers' tastes for a particular good are directly influenced by the size of its consumption. When all types of externalities are accounted for, it would be no exaggeration to say that a substantial fraction of consumption goods have network properties.

Despite their importance, network goods have received relatively little attention in economic theory.[1] Analysis of network goods in the literature has mostly been focused on the resolution of the coordination problem arising from the multiplicity of equilibria. When every consumer expects others to adopt the good, its expected value is high enough to render adoption a rational decision (at least for some price). On the other hand, when every consumer expects no other consumers to adopt, then its low expected value makes no adoption rational. Expectation is self-fulfilling in both cases, leading to multiple, Pareto-ranked equilibria. A subsidy scheme as proposed by Dybvig and Spatt (1983) is one way to eliminate the problem by promising to compensate the adopters when the number of adoptions is below some threshold. The existence of Pareto-ranked equilibria is also the main focus of the analysis of intertemporal patterns of adoption of a network good.[2] In contrast, the problem of revenue maximization by a monopolist has been analyzed only indirectly either through the analysis of subsidy schemes under the implicit assumption that higher participation implies higher revenue, or through the analysis of introductory prices, a common practice of setting a low price for early adopters and a higher, regular price for others (Cabral *et al.*, 1999).[3] The objective of this paper is to directly explore the revenue maximizing coordination and pricing of a network in the incomplete information environment.

In the present context, a *network* is the set of all adopting buyers. Each buyer i's valuation function v_i depends on a network, and also is an increasing function of his private type distributed over the unit interval. A *coordinating scheme* is a

[1]Rohlfs (1974) is the first to give a theoretical analysis of network goods.

[2]See Gale (1995, 2001), Ochs and Park (2010) and Shichijo and Nakayama (2009).

[3]Sekiguchi (2009) examines the monopolist's revenue in the dynamic setup as in Gale (1995) when the price is held constant over time and across consumers. Aoyagi (2010) analyzes a related but different problem in which a monopolist attempts to maximize revenue when the buyers' valuations mutually depend on one another's types.

pair of an assignment rule, which determines a network as a function of the buyers' reported types, and a pricing rule, which determines transfer from each buyer as a function of the realized network.

We analyze a revenue maximizing coordinating scheme that is *ex post implementable* in the sense that it is strategy-proof and ex post individually rational. Our first result on two buyers identifies all possible configurations for the optimal ex post implementable schemes under an arbitrary type distribution. It turns out that an optimal scheme can take diverse forms depending on the joint distribution of private types and the buyers' valuation functions. When the buyers' types are independent, however, one simple property of a coordinating scheme emerges essential as described below. Given the price of each network, consider the marginal type of buyer i who is just indifferent between adopting network a at price p_i and not adopting. We say that a network a priced at p_i is *more affordable* for buyer i than another network a' priced at p_i' if, whenever the marginal type of buyer i for network a is lower than the marginal type for network a'. In other words, a is more affordable than a' if any type who would accept a' at p_i' will surely accept a at p.[4] A coordinating scheme is *monotone* if (1) a larger network is always more affordable than a smaller network for every buyer, and (2) the assignment rule chooses the largest network as permitted by individual rationality. In other words, a monotone scheme has an important efficiency property that it does not exclude any buyer type who is willing to adopt the network for the given price.

When the type distribution satisfies the increasing hazard rate condition and the value functions satisfy some permissive conditions, the optimal scheme is monotone against two buyers with independent types. Against three or more ex ante symmetric buyers, we establish the following results under the same conditions on the distribution and value functions. First, when the externalities are sufficiently strong, there exists a monotone scheme that is optimal among the class of symmetric ex post implementable schemes. Second, we look at a stronger incentive compatibility condition under which no group deviations in reporting are profitable. Any coordinating scheme satisfying the condition is hence robust against buyer collusion.[5] We show that a monotone coordinating scheme is coalitionally ex post implementable in this sense. Furthermore, there exists a monotone scheme that is optimal among the class of symmetric coalitionally ex post implementable schemes.

[4]Note that this does not imply that $p_i < p_i'$ since the values of a and a' are different for i.

[5]Buyer collusion is a plausible concern for some type of networks. For example, potential buyers of an industrial park may come from the same industry and know each other well.

The idea of a coordinating scheme is a generalization of an inducement scheme proposed by Park (2004). An inducement scheme, which itself generalizes the subsidy schemes discussed above to the incomplete information environment, is a sales mechanism in which the transfer between the seller and buyers depends on the realized network. It first posts a price of each network, and then lets the buyers simultaneously decide whether to adopt or not. Because of this feature, the buyers' adoption decisions are independent of one another under an inducement scheme. In contrast, we model a seller who actively coordinates adoption decisions, and propose a sales scheme that works a coordinating device.

The perceived multiplicity of equilibria in network problems makes strategy-proofness a preferable incentive condition compared with Bayesian incentive compatibility. While strategy-proofness is independent of the type distribution by definition, the computation of expected revenues requires the specification of the distribution of buyer types. One unique aspect of the present analysis is that it combines these two elements together.[6] As will be seen, our conclusion that an optimal scheme is monotone is *distribution-free* in the sense that it is not sensitive to the specification of the type distribution as long as it satisfies the standard increasing hazard rate condition.

In line with the existing research on network goods, we suppose that each network is associated with a single (individualized) price. In other words, we analyze adoption-contingent pricing of a network, where the price depends only on the identities of adopters rather than their and others' reported types. Adoption-contingent pricing is extensively analyzed in contracting problems with externalities. In particular, the principal's optimization problem (such as revenue maximization and cost minimization) in various contracting settings are studied by Armstrong (2006), Bernstein and Winter (2010), and Segal (2003), among others. Compared with these models, the distinguishing feature of the present model is the presence of incomplete information about buyer types.

In network problems, only a subset of buyers may end up consuming the good. A similar framework is found in the problem of excludable public goods where the planner can exclude some agents from consumption. However, the public good literature typically assumes that the good's value depends on the amount of con-

[6]A similar approach is taken by Shao and Zhou (2008), who combine strategy-proofness and expected surplus maximization in an allocation problem of an indivisible good to two buyers. One interpretation is that the buyers have common knowledge about one another's type, but the seller only knows their distribution.

tributions from the agents rather than their adoption status, and focuses on the efficient cost sharing rather than revenue maximization.[7]

The paper is organized as follows: The next section introduces a coordinating scheme. Ex post implementable schemes are characterized in Section 3. We study the problem with two buyers in Section 4, and optimal symmetric schemes with a general number of ex ante symmetric buyers in Section 5. Subsection 5.1 analyzes the case of strong externalities, and Subsection 5.2 analyzes coalitionally strategy-proof schemes. We conclude in Section 6. All the proofs are collected in the Appendix.

2 Model

There are I potential buyers of a network good indexed by $i \in I = \{1, \ldots, I\}$.[8] Buyer i's decision is either to buy the good ($a_i = 1$), or not ($a_i = 0$). A *network* is a profile of adoption decisions $a = (a_i)_{i \in I}$, and an element of the set $A = \{0, 1\}^I$. Let A_i be the set of networks including buyer i: $A_i = \{a \in A : a_i = 1\}$.[9] The value of the good to buyer i, denoted $v_i(a, s_i)$, depends on the network a as well as his own private type s_i. The type profile $s = (s_i)_{i \in I}$ has a strictly positive joint density g over $S = \prod_{i \in I} S_i$, where S_i is the unit interval $[0, 1] \subset \boldsymbol{R}_+$. Throughout, we normalize the payoff from no-adoption to zero for any buyer type: $v_i(a, s_i) = 0$ for any $a \notin A_i$ and $s_i \in S_i$.

A *coordinating scheme* determines the network as a function of the private type profile, and monetary transfer as a function of the realized network. Formally, a coordinating scheme is a pair (f, t) of an *assignment rule* $f : S \to A$ and a *pricing rule* $t = (t_1, \ldots, t_I) : A \to \boldsymbol{R}^I$: $f(s) \in A$ is the network formed under the type profile $s \in S$, and $t_i(a) \in \boldsymbol{R}$ is the monetary transfer from buyer i when network a is realized.[10] A coordinating scheme (f, t) is *strategy-proof* if

$$v_i(f(s_i, s_{-i}), s_i) - t_i(f(s_i, s_{-i})) \geq v_i(f(s_i', s_{-i}), s_i) - t_i(f(s_i', s_{-i})) \quad \text{for every } i,\ s_i,\ s_i' \text{ and } s_{-i},$$

[7] See, for example, Moulin (1994), Deb and Razzolini (1999a, b), and Bag and Winter (1999).

[8] Note that symbol I is used to denote both the set and the number of buyers.

[9] In view of the one-to-one correspondence between a and the set $\{i \in I : a_i = 1\}$ of adopting agents, the term network is interchangeably used to imply the set of agents who adopt the good.

[10] In other words, a coordinating scheme is a social choice function $(f, \tau) : S \to A \times \boldsymbol{R}^I$ such that for any s, $s' \in S$, $f(s) = f(s')$ implies $\tau(s) = \tau(s')$.

and *ex post individually rational* if

$$v_i(f(s_i, s_{-i}), s_i) - t_i(f(s_i, s_{-i})) \geq 0 \quad \text{for any } i,\ s_i, \text{ and } s_{-i}.$$

A coordinating scheme (f, t) is *ex post implementable* if it is both strategy-proof and ex post individually rational.[11]

Given the concern for the multiplicity of equilibria in the network good problems, strategy-proofness is a particularly suitable requirement compared with Bayesian incentive compatibility, which does not address the multiplicity issue.[12]

We say that a coordinating scheme (f, t) is *constrained (ex post) efficient* if

$$\sum_i \{v_i(f(s), s_i) - t_i(f(s))\} \geq \sum_i \{v_i(a, s_i) - t_i(a)\}$$

for every $s \in [0,1]^I$ and $a \in A$. In other words, when the pricing rule t is given, no other network achieves a higher aggregate net welfare than $f(s)$ for any profile s.[13]

Let the seller's expected revenue per buyer under a coordinating scheme (f, t) be defined by

$$R(f,t) = \frac{1}{I} \sum_{i \in I} E_s[t_i(f(s))].$$

An ex post implementable coordinating scheme (f, t) is *optimal* if it maximizes the seller's expected revenue:

$$R(f,t) = \max\{R(f', t') : (f', t') \text{ is ex post implementable}\}.$$

3 Characterization of Ex Post Implementability

In this section, we present a simple characterization of ex post implementability that will later be used in the analysis of optimal schemes. We make the following assumptions on the valuation function $v_i : A \times S_i \to \boldsymbol{R}_+$:

[11] In line with the standard assumption of mechanism design, ex post individual rationality requires that each buyer i not reject the network $f(s)$ even if he is not assigned the good. Since $v_i(a, \cdot) = 0$ for $a \notin A_i$, it immediately follows that the transfer required for buyer i in such a case is non-positive.

[12] Park (2004) presents an analysis of Bayesian implementable sales mechanisms for a network good. His analysis shows that the optimal Bayesian implementable mechanism admits multiple equilibria.

[13] When unconstrained by t (*i.e.*, under $t(\cdot) \equiv 0$), only the maximal network I is ex post efficient for any s since $v_i(a, \cdot) \geq 0$ for any a.

Assumption 1 *For any* $i \in I$,

1) $v_i(a, 0) = 0$ *for any* $a \in A$,

2) $a \notin A_i \Rightarrow v_i(a, \cdot) \equiv 0$,

3) $a \in A_i \Rightarrow \frac{\partial v_i}{\partial s_i}(a, \cdot) > 0$,

4) For any a, $b \in A$ *and* s_i, $s'_i \in [0, 1]$, $\frac{\partial v_i}{\partial s_i}(a, s_i) < \frac{\partial v_i}{\partial s_i}(b, s_i) \Leftrightarrow \frac{\partial v_i}{\partial s_i}(a, s'_i) < \frac{\partial v_i}{\partial s_i}(a, s'_i)$.

That is, the value of the good equals zero (1) to a buyer of the lowest type $s_i = 0$, and (2) to a non-adopter. Moreover, (3) the value is strictly increasing with the private type. (4) is a single-crossing condition saying that if buyer i doesn't find two networks equivalent to each other, then the slope of his value function for one of the networks is always higher than that for the other network. We introduce some notation as follows. First, let

$$C_i(a) = \{a' \in A : v_i(a', \cdot) = v_i(a, \cdot)\}$$

be the set of networks among which buyer i is indifferent. For example, when the level of externalities depends only on the size of a network defined by $|a| = \sum_{i \in I} a_i$, then $C_i(a) = \{a' \in A_i : |a'| = |a|\}$ for $a \in A_i$. Next, fix any $s_{-i} \in S_{-i}$ and let

$$B_i(s_{-i}) = \{f(s_i, s_{-i}) : s_i \in S_i\}$$

be the set of possible networks that buyer i can achieve by changing his report when the type profile of other buyers is fixed at s_{-i}. Further, for any network $a \in A$ and profile $s_{-i} \in S_{-i}$, let

$$L_i(a, s_{-i}) = \mathrm{cl}\,\{s_i \in S_i : f(s_i, s_{-i}) = a\}$$

be the (closure of the) set of i's types that would lead to network a when others' type profile is fixed at s_{-i}, and for any network $a \in A$,

$$L_a = \mathrm{cl}\,\{s \in S : f(s) = a\}$$

be the (closure of the) set of type profiles that induce network a.

Now suppose that (f, t) is a coordinating scheme. Given any network $a \in A_i$, define $y_i^a \in [0, 1]$ to be the marginal type at which buyer i is indifferent between participating in network a for price $t_i(a)$, and not adopting:

$$v_i(a, y_i^a) - t_i(a) = 0. \tag{1}$$

Such a type y_i^a is unique by Assumption 1 if it exists. If $v_i(a,0) - t_i(a) > 0$, then let $y_i^a = 0$ and if $v_i(a,1) - t_i(a) < 0$, then let $y_i^a = 1$. Moreover, given any pair of networks a, $b \in A_i$ such that a has larger externalities than b: $\frac{\partial v_i}{\partial s_i}(a,\cdot) > \frac{\partial v_i}{\partial s_i}(b,\cdot)$, define $y_i^{ab} = y_i^{ba} \in [0,1]$ to be the marginal type at which buyer i is indifferent between network a at price $t_i(a)$ and network b at price $t_i(b)$:

$$v_i(a, y_i^{ab}) - t_i(a) = v_i(b, y_i^{ab}) - t_i(b). \quad (2)$$

Again, such a type y_i^{ab} is unique if it exists. If $v_i(a,0) - t_i(a) > v_i(b,0) - t_i(b)$, set $y_i^{ab} = 0$ and if $v_i(a,1) - t_i(a) < v_i(b,1) - t_i(b)$, set $y_i^{ab} = 1$.[14]

For each $i \in I$ and $a \in A_i$, we may restrict attention to the price $t_i(a)$ such that $0 \leq t_i(a) \leq v_i(a,1)$. Since there is a one-to-one correspondence between $t_i(a)$ and y_i^a for any such $t_i(a)$, we will interchangeably use the profile $y = (y_i^a)_{i \in I, a \in A_i}$ and the pricing rule t in what follows. We say that network a priced at $t_i(a)$ is *more affordable* for buyer i than network a' priced at $t_i(a')$ if $y_i^a \leq y_i^{a'}$. In other words, buyer i finds network a worth the purchase for a wider range of types than a'. Note that this is not equivalent to saying that a is less expensive than a', which is expressed as $t_i(a) \leq t_i(a')$.

Proposition 1 *A coordinating scheme (f,t) is ex post implementable if and only if the following holds. For any i and s_{-i}, if $a^1, \ldots, a^n \in A$ are all distinct networks such that*

a) $\frac{\partial v_i}{\partial s_i}(a^1,\cdot) < \cdots < \frac{\partial v_i}{\partial s_i}(a^n,\cdot)$, *and*

b) $\{a^1, \ldots, a^n\} \subset B_i(s_{-i}) \subset \bigcup_{k=1}^n C_i(a^k)$,

then for $k = 1, \ldots, n$,

1) $t_i(a) = t_i(a^k)$ *if* $a \in C_i(a^k) \cap B_i(s_{-i})$,

2) $t_i(a^1) \leq 0$,

3) $t_i(a^1) \leq \cdots \leq t_i(a^n)$.

4) $\bigcup_{a \in C_i(a^k) \cap B_i(s_{-i})} L_i(a, s_{-i}) = \left[y_i^{a^{k-1}a^k},\ y_i^{a^k a^{k+1}} \right]$, *where* $y_i^{a^0 a^1} = 0$.

[14] To summarize the three cases, y^a is defined so that buyer i prefers a to no adoption to the right of y^a and prefers no adoption to a to the left of it. Likewise, y^{ab} is defined so that buyer i prefers a to b to the right of y^{ab}, and b to a to the left of it.

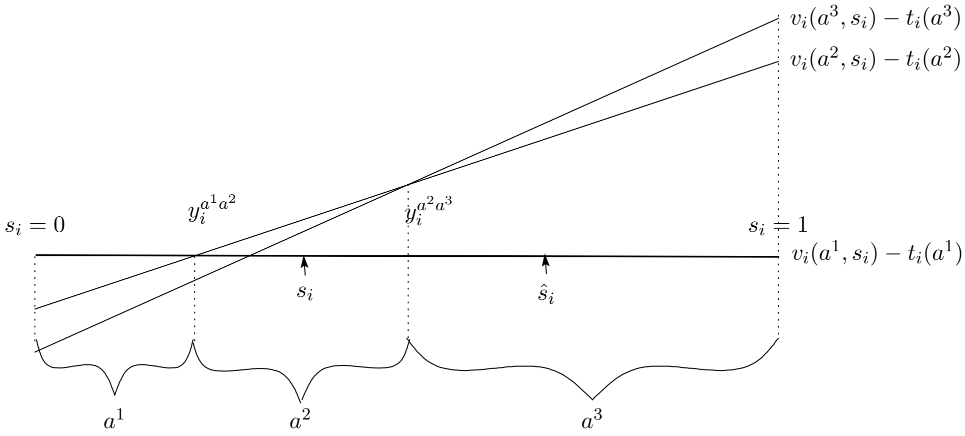

Figure 1: Illustration of ex post implementability

The above proposition can be illustrated as follows: Fix the type profile of buyers other than i. a^k is the network that may be chosen for some report by buyer i and represents the set $C_i(a^k)$ of networks equivalent to it for him. The networks that may be chosen for different reports of i's type should be lined up in the order of their externalities as measured by the marginal values $\frac{\partial v_i}{\partial s_i}(a, s_i)$. Note that network a^1 with the smallest externalities for i is typically the one that excludes buyer i such that $a^1 \notin A_i$ or $v_i(a^1, \cdot) = 0$. Proposition 1 states (1) for any equivalent networks, the required transfers are the same, (2) for the network a^1 that has the smallest externalities and is assigned to the lowest type, the required transfer is non-positive, (3) the transfers increase with the externalities, and (4) each equivalence class represented by a^k is assigned to the kth interval. Note that the relative ordering between any equivalent networks is indeterminate. An ex post implementable assignment rule is illustrated in Figure 1. In the figure, suppose for example that i's true type is s_i as indicated. If he reports it truthfully, then his payoff equals $v_i(a^2, s_i) - t_i(a^2)$, which is greater than $v_i(a^3, s_i) - t_i(a^3)$ that he would get by misreporting that his type is $\hat{s}_i$.

4 Optimal Schemes against Two Buyers

Suppose now that there are only two buyers $I = \{1, 2\}$. The set of possible networks in this case is given by

$$A = \{11, 10, 01, 00\},$$

where

$$11 = (1,1),\ 10 = (1,0),\ 01 = (0,1) \text{ and } 00 = (0,0).$$

We assume positive network externalities as follows.

Assumption 2 *For each $a \in A_i$, $v_i(a,0) = 0$. Furthermore,*

$$\frac{\partial v_1}{\partial s_1}(11,\cdot) > \frac{\partial v_1}{\partial s_1}(10,\cdot), \quad \frac{\partial v_2}{\partial s_2}(11,\cdot) > \frac{\partial v_2}{\partial s_2}(01,\cdot).$$

The following theorem characterizes the optimal schemes in a general environment with two buyers.

Theorem 1 *If (f,t) is an optimal ex post implementable coordinating scheme against two buyers under Assumption 2, then it takes one of the following forms.*

(A) $y_1^{11},\ y_2^{11} < 1,\ y_1^{10},\ y_2^{01} \in (0,1),\ \begin{cases} \textit{(A0) } y_1^{11} \le y_1^{10},\ y_2^{11} \le y_2^{01} \\ \textit{(A1) } y_1^{11} \le y_1^{10},\ y_2^{11} > y_2^{01}, \\ \textit{(A2) } y_1^{11} > y_1^{10},\ y_2^{11} \le y_2^{01} \end{cases}$

$$\begin{cases} L_{11} = [y_1^{11},\ 1] \times [y_2^{11},\ 1] \\ L_{10} = [y_1^{10},\ 1] \times [0,\ y_2^{11}] \\ L_{01} = [0, y_1^{11}] \times [y_2^{01}, 1]. \end{cases}$$

(B1) $0 < y_1^{10} < y_1^{11} < y_1^{11,10} < 1,\ y_2^{11} < 1,\ y_2^{01} \in (0,1),$

$$\begin{cases} L_{11} = [y_1^{11,10},\ 1] \times [y_2^{11},\ 1] \\ L_{10} = [y_1^{10},\ 1] \times [0,1] \setminus \operatorname{int} L_{11} \\ L_{01} = [0, y_1^{10}] \times [y_2^{01}, 1]. \end{cases}$$

(B2) $0 < y_2^{01} < y_2^{11} < y_2^{11,01} < 1,\ y_1^{10} \in (0,1),\ y_1^{11} < 1,$

$$\begin{cases} L_{11} = [y_1^{11},\ 1] \times [y_2^{11,01},\ 1] \\ L_{10} = [y_1^{10},\ 1] \times [0,\ y_2^{01}] \\ L_{01} = [0,1] \times [y_2^{01}, 1] \setminus \operatorname{int} L_{11}. \end{cases}$$

(C1) $y_1^{10} < y_1^{11} < 1,\ 0 < y_2^{01} < y_2^{11} < 1,$

$$\begin{cases} L_{11} = [y_1^{11},\ 1] \times [y_2^{11},\ 1] \\ L_{10} = [y_1^{10},\ 1] \times [0,\ y_2^{11}] \\ L_{01} = [0, y_1^{10}] \times [y_2^{01}, 1]. \end{cases}$$

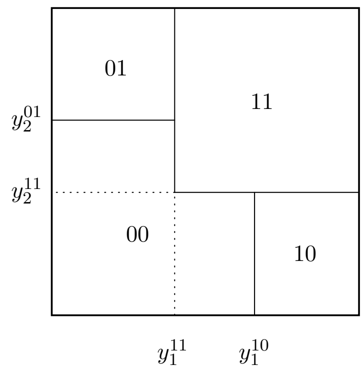

Figure 2: Configuration (A0)

(C2) $0 < y_1^{10} < y_1^{11} < 1,\ y_2^{01} < y_2^{11} < 1,$

$$\begin{cases} L_{11} = [y_1^{11},\ 1] \times [y_2^{11},\ 1] \\ L_{10} = [y_1^{10},\ 1] \times [0,\ y_2^{01}] \\ L_{01} = [0, y_1^{11}] \times [y_2^{01}, 1]. \end{cases}$$

These configurations are depicted in Figures 2, 3, 4 and 5. Note that Theorem 1 states that an optimal scheme has one of these configurations for any distribution of buyer types whether they are independent or correlated. Specification of the distribution and the value functions is required to pin down which one of the various assignment rules is indeed optimal as well as the exact locations of the marginal types y_i^a. Note that in most of these configurations, some types are precluded from adoption even though they are willing to adopt. In Configuration (B2), for example, buyer 1 is not assigned the good when his type $s_1 \geq y_1^{11}$ and buyer 2's type $s_2 \in (y_2^{01}, y_2^{11,01})$. This occurs when the seller finds it more profitable to charge buyer 2 a higher price $y_2^{11,01}$ for joint adoption than to charge him a lower price y_2^{11} and allow joint adoption whenever feasible.[15] As depicted, hence, no configuration other than (A0) is constrained efficient.[16]

4.1 Independent Types

A sharper characterization of an optimal scheme is possible under some additional assumptions on the valuation functions and the type distribution. Assume specifi-

[15]Note that the latter would correspond to Configuration (A1).

[16]In the degenerate cases, (A1) and (A2) are constrained efficient when $y_1^{11} = y_1^{10}$ and $y_2^{11} = y_2^{01}$, respectively.

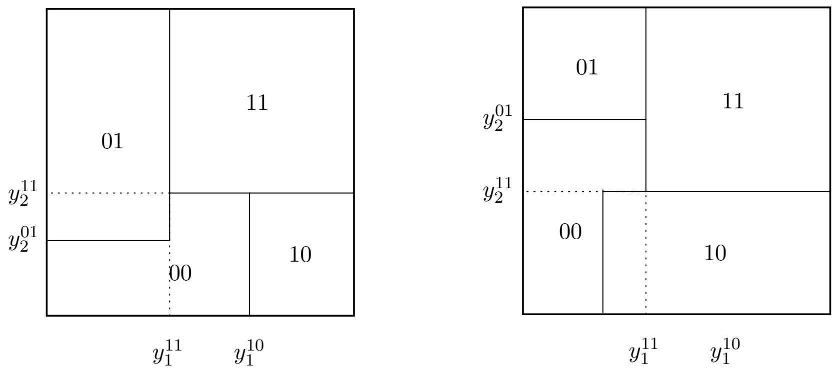

Figure 3: Configurations (A1) (left) and (A2) (right)

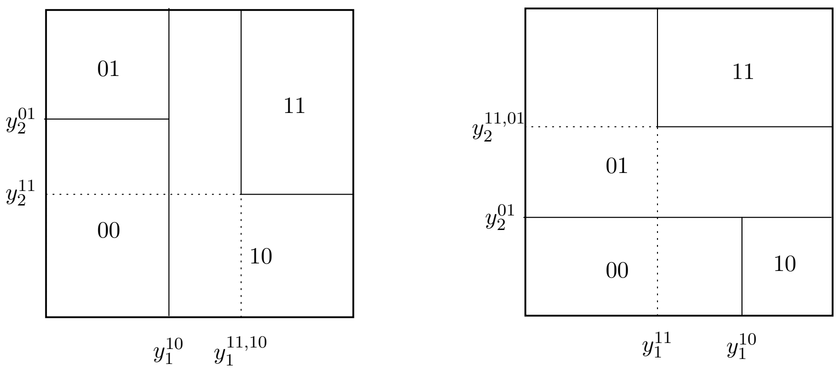

Figure 4: Configurations (B1) (left) and (B2) (right)

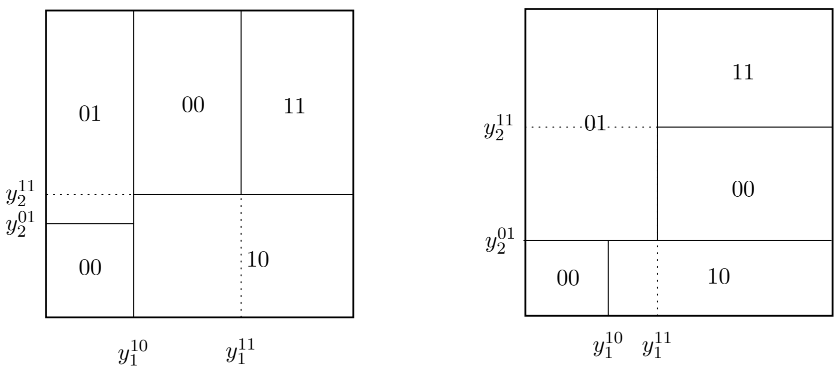

Figure 5: Configurations (C1) (left) and (C2) (right)

cally that the types s_1 and s_2 are independent. Let G_i be the cumulative distribution function of s_i, and for $i \in I$, $a \in A$ and $s_i \in S_i$, define

$$r_i(a, s_i) = \{1 - G_i(s_i)\}\, v_i(a, s_i)$$

to be the seller's expected revenue from buyer i when he offers network a for price $v_i(a, s_i)$. We make the following assumptions.

Assumption 3 *1) $v_i(a, \cdot)$ is strictly log-concave for each $a \in A_i$.*[17]

2) $\frac{v_1(11,\cdot)}{v_1(10,\cdot)}$ and $\frac{v_2(11,\cdot)}{v_2(01,\cdot)}$ are weakly decreasing.

3) $\frac{g_i(\cdot)}{1-G_i(\cdot)}$ is strictly increasing.

In Assumption 3, the only requirement on the distribution is the increasing hazard rate condition in (3), which is known to hold for most distributions. The first two conditions concern the functional form of the valuation function. (1) requires that it be not too convex as a function of the type, and (2) is a single-crossing condition on log v_i since its alternative expression is $\frac{\partial(\log v_1)}{\partial s_1}(11, \cdot) \leq \frac{\partial(\log v_1)}{\partial s_1}(10, \cdot)$. In other words, the log-value of the larger network increases at a lower rate than that of the smaller network. For example, (2) holds with equality for a multiplicatively separable valuation function $v_i(a, s_i) = \gamma_i(a)\, h_i(s_i)$ in which the effect of the network $\gamma_i(a)$ is separated from that of the type $h_i(s_i)$.[18] As summarized by the following lemma, Assumption 3 implies that the graph of $r_1(a, \cdot)$ has a single peak when $a = 11$ or 10 and that the peak of $r_1(11, \cdot)$ is located to the left of that of $r_1(10, \cdot)$ (Figure 6).

Lemma 1 *Suppose that Assumptions 2 and 3 hold. Then*

1) For each $a \in A_i$, $r_i(a, \cdot)$ is strictly log-concave with the (unique) maximizer $\bar{z}_i^a$ which satisfies $\bar{z}_1^{11} \leq \bar{z}_1^{10}$ and $\bar{z}_2^{11} \leq \bar{z}_2^{01}$.

2)

$$\frac{\partial r_1}{\partial s_1}(11, s_1) < \frac{\partial r_1}{\partial s_1}(10, s_1) \quad \textit{for } s_1 > \bar{z}_1^{10},$$
$$\frac{\partial r_2}{\partial s_2}(11, s_2) < \frac{\partial r_2}{\partial s_2}(01, s_2) \quad \textit{for } s_2 > \bar{z}_2^{01}.$$

[17] That is, log $v_i(a, \cdot)$ is strictly concave for each a.

[18] More generally, (2) holds if $v_i(a, s_i) = \gamma_i(a) h_i(s_i) + c_i(s_i)$, $\frac{c_i(s_i)}{h_i(s_i)}$ is weakly increasing in s_i, and $\gamma_i(a)$ increases with the network size. It also holds when $\frac{\partial v_1}{\partial s_1}(11, \cdot) / \frac{\partial v_1}{\partial s_1}(10, \cdot)$ and $\frac{\partial v_2}{\partial s_2}(11, \cdot) / \frac{\partial v_2}{\partial s_2}(01, \cdot)$ are weakly decreasing.

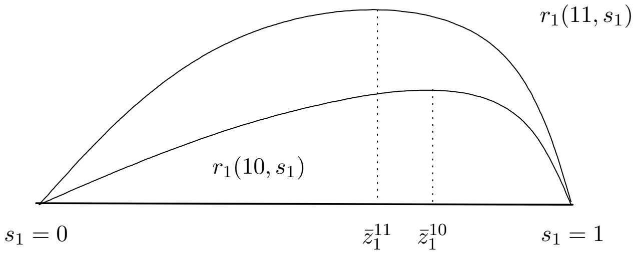

Figure 6: Functions $r_1(10, \cdot)$ and $r_1(11, \cdot)$.

We say that a coordinating scheme (f, y) against two buyers is *monotone* if

1) $y_1^{11} \leq y_1^{10}$, $y_2^{11} \leq y_2^{01}$, and

2) $f_1(s) = \begin{cases} 1 & \text{if } s_1 \geq y_1^{10}, \text{ or } s \geq (y_1^{11}, y_2^{11}), \\ 0 & \text{otherwise,} \end{cases}$ and

$f_2(s) = \begin{cases} 1 & \text{if } s_2 \geq y_2^{01}, \text{ or } s \geq (y_1^{11}, y_2^{11}), \\ 0 & \text{otherwise.} \end{cases}$

Under a monotone scheme, hence, the larger network $a = 11$ is more affordable than the smaller network $a = 10$ or 01, and the network size is maximized subject to the individual rationality constraints. The second property can also be interpreted as saying that the good is allocated to a single buyer only when the other buyer's type is too low for joint adoption. Configuration (A0) in Figure 1 corresponds to a monotone scheme. If a monotone scheme is optimal, then $(y_1^{10}, y_2^{01}) = (\bar{z}_1^{10}, \bar{z}_2^{01})$ and

$$(y_1^{11}, y_2^{11}) \in \operatorname{argmax} \{1 - G_2(y_2^{11})\} \, r_1(11, y_1^{11}) + \{1 - G_1(y_1^{11})\} \, r_2(11, y_2^{11}) + G_2(y_2^{11}) \, r_1(10, \bar{z}_1^{10}) + G_1(y_1^{11}) \, r_2(01, \bar{z}_2^{01}). \tag{3}$$

It is clear from the definition that a monotone scheme is constrained efficient.[19]

The following theorem characterizes optimal schemes against two buyers with independent types.

Theorem 2 *Suppose that (s_1, s_2) is independent and that Assumptions 2 and 3 hold. If (f, y) is an optimal ex post implementable coordinating scheme against two buyers, then it is monotone.*

[19] A partial converse of this also holds.

Proof. See the Appendix. ■

When $y_1^{11} < y_1^{10}$ and $y_2^{11} < y_2^{01}$, it is impossible to replicate the monotone scheme by any scheme in which the buyers' decisions are based only on their own types or on the decisions of other buyers: In any such scheme, at least one buyer's decision (*e.g.,* the first-mover's decision) must be independent of other buyers' types.

Example: Suppose that s_i has the uniform distribution $G_i(s_i) = s_i$, and that the buyers' valuation functions are given by

$$\begin{array}{ll} v_1(10, s_1) = \gamma s_1 & v_2(01, s_2) = \gamma s_2 \\ v_1(11, s_1) = \delta s_1, & v_2(11, s_2) = \delta s_2, \end{array}$$

where $0 < \gamma < \delta$. Given that the optimal scheme is monotone, the marginal type for the single adoption 10 or 01 equals $y_1^{10} = y_2^{01} = \frac{1}{2}$. By (3) and symmetry, the marginal type $y_1^{11} = y_2^{11}$ for the joint adoption 11 solves

$$y_1^{11} = y_2^{11} \in \underset{x}{\operatorname{argmax}}\ \delta x(1-x)^2 + \frac{\gamma}{4}x.$$

Solving this, we get[20]

$$y_1^{11} = y_2^{11} = \frac{1}{3\delta}\left\{2\delta - \sqrt{\delta^2 - \frac{3}{4}\gamma\delta}\right\}.$$

We can confirm that $y_1^{11} = y_2^{11} < \frac{1}{2} = y_1^{10} = y_2^{01}$ if and only if $\gamma < \delta$. Consider now the price of each network associated with these marginal types. They are given by

$$t_1(10) = t_2(01) = \frac{\gamma}{2}, \quad \text{and} \quad t_1(11) = t_2(11) = \frac{1}{3}\left\{2\delta - \sqrt{\delta^2 - \frac{3}{4}\gamma\delta}\right\}.$$

From these, we can check that the price of the size 2 network 11 is higher than that of the size 1 network if and only if

$$\frac{\delta}{\gamma} > \frac{3}{4},$$

which is true since $\delta > \gamma$. In this example, the larger network is more expensive than the smaller network although it is more affordable in the aforementioned sense.

[20] As seen, analytical derivation of an optimal scheme is possible only under very limited specifications of the distribution and values.

5 Optimal Symmetric Schemes

With more than two buyers, the problem of identifying all the ex post implementable schemes becomes intractable. In this section, we focus on an optimal symmetric scheme when the buyers are ex ante symmetric. We show that the optimal scheme is monotone when the network externalities are strong, or when a stronger notion of incentive compatibility is imposed.

Suppose that the types $s_1, \ldots, s_I$ are independent and identically distributed, and denote by g the density of s_i and by G the corresponding cumulative distribution. The valuation functions are symmetric in the sense that

$$v_i(a, s_i) = v_j(a', s_j)$$

for any $i \neq j$, $s_i = s_j \in [0,1]$ and a, $a' \in A$ such that $(a_i, a_j, a_{-i-j}) = (a'_j, a'_i, a'_{-i-j})$. The symmetry condition implies that the network externalities depend only on the *size* of the network $a \in A$, defined by $|a| = \sum_{j \in I} a_j$. In what follows, hence, we refer to any network of size k as network k, and denote the set of networks by $N \equiv \{0, 1, \ldots, I\} = I \cup \{0\}$. For any network $k \in N$, denote the valuation function of any buyer by $v^k : [0,1] \to \boldsymbol{R}_+$.

We say that a coordinating scheme (f, t) is *symmetric* if for any $i \neq j$,

$$(f_i(s), f_j(s), f_{-i-j}(s)) = (f_j(s'), f_i(s'), f_{-i-j}(s'))$$

for any s, $s' \in S$ such that $(s_i, s_j, s_{-i-j}) = (s'_j, s'_i, s_{-i-j})$, and

$$(t_i(a), t_j(a), t_{-i-j}(a)) = (t_j(a'), t_i(a'), t_{-i-j}(a'))$$

for any a, $a' \in A$ such that $(a_i, a_j, a_{-i-j}) = (a'_j, a'_i, a'_{-i-j})$. That is, when (f, t) is symmetric, swapping the private types of any pair of buyers results in the swapping of their assignments and transfers but does not affect those of any other buyers.[21] When the scheme (f, t) is symmetric, the transfer depends on the network only through its size. That is, $t_i(a) = t_j(a')$ for any i, $j \in I$ and any $a \in A_i$, $a' \in A_j$ such that $|a| = |a'|$. Hence, we let t^k denote the price of network k for any single buyer. We make the following assumptions as in the previous section.

Assumption 4 *1)* $v^1(0) = \cdots = v^I(0) = 0$ *and* $(v^1)'(\cdot) < \cdots < (v^I)'(\cdot)$.

2) $v^1, \ldots, v^I$ *are strictly log-concave.*

[21] In the social choice literature, this property is often called anonymity.

3) $\frac{v^n(\cdot)}{v^m(\cdot)}$ *is weakly decreasing if* $m < n$.

4) $\frac{g(\cdot)}{1-G(\cdot)}$ *is strictly increasing.*

Again, the increasing hazard rate in (4) is the only requirement on the distribution. (1) says that the value is zero to the lowest type, and that the externalities as measured by the derivatives increase strictly with the network size. (2) and (3) are exact counterparts of those in the previous section.

We adapt y_i^a and y_i^{ab} defined in (1) and (2) to the symmetric environment by writing y^n for the marginal type that is indifferent between network n priced at t^n and no-adoption, and y_i^{mn} for the marginal type that is indifferent between m at priced at t^m and n priced at t^n.[22] Just as in the general formulation of Section 3, restricting the range of the pricing rule t^n to $[0, v^n(1)]$ for each $n \in N$ entails no loss of generality as far as the expected revenue is concerned. Given the one-to-one correspondence between such a pricing rule $t = (t^1, \dots, t^n)$ and the profile of marginal types $y = (y^1, \dots, y^I)$, we again use t and y interchangeably when describing a coordinating scheme.

Let $\lambda^0 = \lambda^0_{I-1} = 1$, and for each $k = 1, \dots, I-1$, let $\lambda^k = \lambda^k_{I-1}$ be the kth highest value among $I-1$ types $s_{-i} = (s_j)_{j \neq i}$. A symmetric coordinating scheme (f, t) is *monotone* if

1) $y^I \leq \cdots \leq y^1$, and

2) $$f_i(s) = \begin{cases} 1 & \text{if } s_i \geq y^n \text{ and } \lambda^{n-1} \geq y^n \text{ for some } n \in N, \\ 0 & \text{otherwise.} \end{cases}$$

As before, in a monotone scheme, (1) a larger network is more affordable than a smaller network, and (2) the maximal network is chosen subject to individual rationality: for any $n \in N$, $|f(s)| = n$ if and only if $|\{i \in I : s_i \geq y^n\}| = n$.[23] It is not difficult to see from Proposition 1 that a monotone scheme is strategy-proof.[24] Moreover, a monotone scheme is constrained efficient in the sense defined in Section 2.

[22] As before, $y^n = 0$ if $t^n < 0$, and $y^n = 1$ if $t^n > v^n(1)$. Likewise, $y^{mn} = 0$ if $t^n - t^m < 0$, and $y^{mn} = 1$ if $t^n - t^m > v^n(1) - v^m(1)$.

[23] To see that a monotone scheme (f, y) has this property, note that it is clear from the definition that $|f(s)| = n$ if $|\{i \in I : s_i \geq y^n\}| = n$. For the other implication, suppose that $|f(s)| = n$. Then IR implies that $|\{i \in I : s_i \geq y^n\}| \geq n$. If the inequality is strict, then take any i such that $s_i \geq y^n$. For this i, $\lambda^n \geq y^n \geq y^{n+1}$ so that $|f(s)| \geq n+1$ must hold by definition, a contradiction.

[24] Proposition 2 below proves that it satisfies a stronger condition of coalitional strategy-proofness.

As is the case with two buyers, we begin by examining the seller's expected revenue from a single buyer i. Specifically, take any set $K \subset N$ of networks. Let also the marginal types $y = (y^1, \ldots, y^I) \in S$ be given. Suppose now that the seller offers buyer i a menu consisting of networks in K. That is, the menu lists each network $k \in K$ for price $t^k = v^k(y^k)$. Letting $y^K = (y^k)_{k \in K}$, we will denote by $r^K(y^K)$ the seller's expected revenue from offering this menu to buyer i. In Figure 1, for example, the seller's expected revenue from buyer i is given by $r^K(y^K)$, where $K = \{a^1, a^2, a^3\}$, and equals the sum of the probability that each network is chosen multiplied by its price. When the menu contains a single item $K = \{k\}$, we denote $r^K(y^K) = r^k(y^k)$, and when it contains two items $K = \{k, \ell\}$, we denote $r^K(y^K) = r^{k\ell}(y^k, y^\ell)$. We see that

$$r^k(y^k) = \{1 - G(y^k)\}\, v^k(y^k),$$

and that when $k < \ell$ and $y^k < y^\ell < y^{k\ell} < 1$,

$$r^{k\ell}(y^k, y^\ell) = \{1 - G(y^{k\ell})\}\, v^\ell(y^\ell) + \{G(y^{k\ell}) - G(y^k)\}\, v^k(y^k).$$

A general formula of $r^K(y^K)$ is presented in the Appendix. $r^K(y^K)$ plays a key role in what follows since a general coordinating scheme requires a buyer to make a choice from more than two networks even if the type profile of other buyers is fixed.[25] On the other hand, as seen from Figure 2, a monotone scheme requires a buyer to make only a binary choice between network 0 and network k (≥ 1) for any fixed type profile of other buyers.

Under Assumption 4, we have:

Lemma 2 *Suppose that Assumption 4 holds. Then the following hold.*

1) *For each $n \in N$, r^n is strictly log-concave with the (unique) maximizer $\bar{z}^n$ which satisfies $1 > \bar{z}^1 \geq \cdots \geq \bar{z}^I > 0$.*

2) *If $m < n$, $y^m < y^n$, and $(r^n)'(y^{mn}) \geq (r^m)'(y^{mn})$, then $r^n(y^n) > r^{mn}(y^m, y^n)$.*

In other words, (1) each r^n is single-peaked and when $k < \ell$, r^ℓ peaks earlier than r^k, and (2) for the seller, offering a menu containing two networks m and n

[25] For example, in configuration (B1) in Figure 4, when $s_2 > y_2^{01}$, buyer 1 chooses among networks 01, 10 and 11.

is dominated by offering just the larger of the two for some y^m and y^n.[26] Let the marginal types $y = (y^1, \dots, y^I)$ and set $K \subset N$ of networks be given, and define

$$S_{-i}(K, y) = \{s_{-i} \in S_{-i} : \min_{k \in K} (\lambda^{k-1} - y^k) \geq 0, \max_{k \notin K} (\lambda^{k-1} - y^k) < 0\}. \tag{4}$$

$S_{-i}(K, y)$ is the set of type profiles of buyers other than i such that when $s_{-i} \in S_{-i}(K, y)$, it is possible to include i to form any network $k \in K$ without violating any buyer's individual rationality as long as $s_i \geq y^k$, but no other network including i can be formed for any s_i. Since $\lambda^0 = 1$ and $y^k \leq 1$ for any k, for any y, if $S_{-i}(K, y) \neq \emptyset$, then $1 \in K$.[27] Let

$$Q^K(y) = P(s_{-i} \in S_{-i}(K, y))$$

be the probability that s_{-i} is such a type profile.

5.1 Optimal Symmetric Scheme under Strong Externalities

In this section, we analyze optimal schemes when a larger network has significantly stronger externalities than a smaller network. Specifically, we suppose that for each positive network $k = 1, \dots, I$, there exist $\underline{v}^k : [0, 1] \to \boldsymbol{R}_+$ and $0 < \rho^1 \leq \cdots \leq \rho^I$ such that

$$v^k(s_i) = \rho^k \underline{v}^k(s_i).$$

$\underline{v}^k$ is the base valuation function for network k satisfying Assumption 4, and we use $\rho^1, \dots, \rho^I$ to vary the externality levels. Throughout this subsection, we assume that the density g is continuous and strictly positive over $[0, 1]$.

Given the marginal types $y = (y^1, \dots, y^I)$ and the set K of networks, let $S_{-i}(K, y) \subset S_{-i}$ be as defined in (4), and $Q^K(y)$ be the probability that $s_{-i} \in S_{-i}(K, y)$. Define $w : S \to \boldsymbol{R}_+$ by

$$w(y) = \sum_{1 \in K \subset N} Q^K(y) \max_{\emptyset \neq L \subset K} r^L(y^L).$$

To interpret w, suppose that $s_{-i} \in S_{-i}(K, y)$ so that the seller can include i to form only those networks in K. Since his expected revenue from buyer i by offering him a menu L of networks is $r^L(y^L)$, the maximal revenue from buyer i conditional on $s_{-i} \in S_{-i}(K, y)$ cannot exceed $\max_{\emptyset \neq L \subset K} r^L(y^L)$. In this sense, $w(y)$ presents an

[26] In the proof, it is shown that $(r^n)'(y^{mn}) \geq (r^m)'(y^{mn})$ holds only if $y^{mn} < \bar{z}^n$.

[27] In other words, regardless of others' profile, it is possible to form network 1 if i's signal $s_i \geq y^1$.

upper bound on the seller's revenue from buyer i.[28] With symmetry, this also equals an upper bound on the expected revenue (per buyer) under an ex post implementable coordinating scheme (f, y). We can also make the following observation. Suppose that (f, y) is monotone. Suppose further that $s_{-i} \in S_{-i}(K, y)$ for some i and $K \subset N$. Since $y^{\max K} \leq y^k$ for any $k \in K$ by monotonicity, if $s_i < y^{\max K}$, then $s_i < y^k$ for any $k \in K$ so that i is not included in the network: $f_i(s_i, s_{-i}) = 0$. On the other hand, since the network is maximized subject to IR, if $s_i \geq y^{\max K}$, then $|f(s_i, s_{-i})| = \max K$ should hold. Therefore, the expected revenue from buyer i conditional on $s_{-i} \in S_{-i}(K, y)$ equals $r^{\max K}(y^{\max K})$, and the unconditional expected revenue from buyer i under a monotone scheme equals

$$R(f, y) = \sum_{1 \in K \subset N} Q^K(y)\, r^{\max K}(y^{\max K}). \tag{5}$$

For any K, hence, if offering any menu L of networks in K is dominated by offering just the maximal network $\max K$, or equivalently, if $r^{\max K}(y^{\max K}) = \max_{\emptyset \neq L \subset K} r^L(y^L)$, then we have $w(y) = R(f, y)$ by the definition of $w(y)$ and (5). The following lemma summarizes this observation.

Lemma 3 *Suppose that y satisfies $y^I \leq \cdots \leq y^1$, and*

$$r^{\max K}(y^{\max K}) = \max_{\emptyset \neq L \subset K} r^L(y^L) \text{ for any } 1 \in K \subset N.$$

If (f, y) is a symmetric monotone scheme, then $R(f, y) = w(y)$.

Recall from Lemma 2 that when $y^1, \ldots, y^I$ satisfy certain conditions, offering a menu containing two networks is dominated by offering just the larger of the two. Under strong externalities, we can repeat this argument to show that offering a menu of any subset of networks in K is dominated by offering just the largest network $\max K$. In this case, strong externalities further guarantee that any maximizer y of w must satisfy $y^I \leq \cdots \leq y^1$ as would be required by a monotone scheme. The following theorem combines these observations to prove that any maximizer y of w satisfies the conditions of Lemma 3.

Theorem 3 *Suppose that $v^k(s_i) = \rho^k \underline{v}^k(s_i)$ for every $k = 1, \ldots, I$, where $0 < \rho^1 \leq \cdots \leq \rho^I$ and $\underline{v}^k : [0, 1] \to \boldsymbol{R}_+$ satisfies Assumption 4. Then there exists $\varepsilon > 0$ such that if $\max_{2 \leq k \leq I} \frac{\rho^{k-1}}{\rho^k} < \varepsilon$, then the optimal symmetric coordinating scheme is monotone.*

[28] It is only an upper bound because while it takes into account buyer i's IC and all buyers' IR, it does not take into account other buyers' IC.

When the externalities are positive but weak, preliminary analysis indicates that an optimal ex post implementable symmetric scheme is not monotone. Full characterization of an optimal scheme in such an environment appears extremely difficult as it entails a very complex assignment rule. As seen in the next section, however, requiring a stronger version of incentive compatibility recovers the monotonicity of an optimal scheme for any positive degree of externalities.

5.2 Optimal Symmetric Scheme under Coalitional Implementability

Given a coordinating scheme (f, t), a subset $J \subset I$ of buyers, and type profiles $s = (s_J, s_{-J})$ and $\hat{s}_J$, $\hat{s}_J$ is a *profitable deviation* for the coalition J at s if

$$v_i(f(\hat{s}_J, s_{-J}), s_i) - t_i(f(\hat{s}_J, s_{-J})) \geq v_i(f(s), s_i) - t_i(f(s)) \text{ for every } i \in J, \text{ and}$$
$$v_i(f(\hat{s}_J, s_{-J}), s_i) - t_i(f(\hat{s}_J, s_{-J})) > v_i(f(s), s_i) - t_i(f(s)) \text{ for some } i \in J.$$

(f, t) is *coalitionally strategy-proof* if no coalition of buyers has a profitable deviation at any type profile. Coalitional strategy-proofness is hence a strong robustness requirement since even if there exists a group of buyers who share the information about their private types and jointly misreport them, the deviation is not profitable.[29] (f, t) is *coalitionally ex post implementable* if it is coalitionally strategy-proof and ex post individually rational. The following proposition shows that a monotone scheme has this robustness property.

Proposition 2 *A monotone scheme (f, t) is coalitionally ex post implementable.*

Given the marginal types $y = (y^1, \ldots, y^I)$, define

$$M(y) = \{m : m = 1, \ldots, I - 1, y^m < \max_{\ell > m} y^\ell\}.$$

$M(y)$ is the set of networks that are more affordable than some of the larger networks. Also, let

$$K(f) = \{n \in N : |f(s)| = n \text{ for some } s \in S\}$$

be the set of networks that may be formed under f. If (f, y) is a monotone scheme, then $y^I \leq \cdots \leq y^1$ so that $M(y) = \emptyset$, and hence $M(y) \cap K(f) = \emptyset$. In fact, the following lemma shows that any coalitionally ex post implementable scheme (f, y) should satisfy this condition.

[29] Since Moulin (1980), there is extensive analysis of coalitional (or group) strategy-proofness in the social choice and mechanism design literature.

Lemma 4 *Let (f, y) be a symmetric, coalitionally ex post implementable coordinating scheme. Then $M(y) \cap K(f) = \emptyset$.*

Now given a subset $K \subset N$ of networks and marginal types $y = (y^1, \ldots, y^I)$ such that $K \cap M(y) = \emptyset$, let $w(K, y)$ be defined by

$$w(K,y) = \sum_{k \in K} P\Big(\lambda^{k-1} \geq y^k, \max_{\substack{\ell > k \\ \ell \in K}} (\lambda^{\ell-1} - y^\ell) < 0\Big)\, r^k(y^k)$$

$w(K, y)$ is interpreted as the seller's expected revenue under a coalitionally implementable coordinating scheme (f, y) when the set $K(f)$ of networks formed under f equals K, and f always chooses the maximal network in K subject to individual rationality. For a monotone scheme (f, y), we have

$$R(f, y) = w(N, y) \tag{6}$$

since $K(f) = N$ when $y^1 < 1$ and $y^I > 0$.[30] The following proposition shows that for any coalitionally ex post implementable scheme (f, y), $w(K(f), y)$ gives its expected revenue $R(f, y)$.

Lemma 5 *Let (f, y) be a symmetric, coalitionally ex post implementable coordinating scheme. Then $R(f, y) = w(K(f), y)$.*

The following theorem establishes the optimality of a monotone scheme by showing that for any coalitionally ex post implementable scheme (f, y), there exists $\hat{y} = (\hat{y}^1, \ldots, \hat{y}^I)$ such that $\hat{y}^1 \leq \cdots \leq \hat{y}^I$ and $w(K(f), y) \leq w(N, \hat{y})$.

Theorem 4 *Suppose that Assumption 4 holds. Then there exists a monotone coordinating scheme that is optimal in the class of symmetric, coalitionally ex post implementable coordinating schemes.*

6 Conclusion

The sales schemes considered in the literature for network goods do not involve active coordination of the buyers' adoption decisions. In contrast, a coordinating scheme permits the seller to fully coordinate their decisions while maintaining the principle of adoption-contingent pricing. Ex post implementability required in

[30] It can be checked that (6) also holds in the degenerating cases $y^1 = 1$ or $y^I = 0$.

our analysis eliminates the multiplicity of equilibria, a central issue in the network adoption problems. We present monotonicity as a key property of the optimal ex post implementable scheme when the buyers' private types are independent. In a monotone scheme, a larger network is more affordable than a smaller network in the sense that the set of buyer types who are willing to adopt the larger network is larger than that for the smaller network, and given such pricing, assignment is efficient by choosing the maximal network subject to individual rationality.[31] Given that monotonicity is defined in terms of the private types, it has no direct implication on the actual price levels for different networks. As observed in the example in Section 4.1, however, it is not inconsistent with a lower price for a smaller network and a higher price for a larger network. Such a pricing strategy underlies the practice of introductory pricing, which provides a refund to the adopters when there are few adoptions. It remains an open question whether the optimal price of a larger network can be lower.

In this paper, we have only looked at externalities whose magnitude increases with the network size. It would be interesting to study the case of negative externalities, or more complex externalities based on graph structure.[32] Network goods are often supplied competitively as in the case of cellular phones or PC operating systems. While some aspects of such competition have been analyzed by Katz and Shapiro (1985, 1986), much remains to be understood.

Appendix

Proof of Proposition 1 (Necessity)

1. If $t_i(a) > t_i(a^k)$ for $a \in C_i(a^k) \cap B_i(s_{-i})$, then $v_i(f(s_i, s_{-i}), s_i) - t_i(f(s_i, s_{-i})) < v_i(f(s_i', s_{-i}), s_i) - t_i(f(s_i', s_{-i}))$ for s_i and s_i' such that $f(s_i, s_{-i}) = a$ and $f(s_i', s_{-i}) = a^k$, contradicting the strategy-proofness of (f, t).

2. Ex post IR for $s_i = 0$ requires that $v_i(a^1, 0) - t_i(a^1) = -t_i(a^1) \geq 0$.

[31] It is possible to see the efficiency of a monotone scheme in a different way. Suppose that the seller separates out each buyer and chooses a take-it-or-leave-it offer to him that would be optimal if he were the only buyer. The optimal monotone scheme is more efficient than such a scheme.

[32] See Sundararajan (2007) for one such formulation.

3. For s_i and s_i' such that $f(s_i, s_{-i}) = a^k$ and $f(s_i', s_{-i}) = a^{k+1}$, we have

$$\begin{aligned} v_i(a^k, s_i) - t_i(a^k) &= v_i(f(s_i, s_{-i}), s_i) - t_i(f(s_i, s_{-i})) \\ &\geq v_i(f(s_i', s_{-i}), s_i) - t_i(f(s_i', s_{-i})) \\ &= v_i(a^{k+1}, s_i) - t_i(a^{k+1}). \end{aligned}$$

Rearranging, we get

$$t_i(a^{k+1}) - t_i(a^k) \geq v_i(a^{k+1}, s_i) - v_i(a^k, s_i) \geq 0.$$

4. It suffices to show that if $f(s_i, s_{-i}) = a^k$ and $f(s_i', s_{-i}) = a^m$ for $k < m$, then $s_i < s_i'$. Since (f, t) is strategy-proof,

$$\begin{aligned} &v_i(a^m, s_i') - t_i(a^m) \\ &= v_i(f(s_i', s_{-i}), s_i') - t_i(f(s_i', s_{-i})) \\ &\geq v_i(f(s_i, s_{-i}), s_i') - t_i(f(s_i, s_{-i})) \\ &= v_i(a^k, s_i') - t_i(a^k), \end{aligned}$$

and

$$\begin{aligned} &v_i(a^k, s_i) - t_i(a^k) \\ &= v_i(f(s_i, s_{-i}), s_i) - t_i(f(s_i, s_{-i})) \\ &\geq v_i(f(s_i', s_{-i}), s_i) - t_i(f(s_i', s_{-i})) \\ &= v_i(a^m, s_i) - t_i(a^m). \end{aligned}$$

It hence follows that

$$v_i(a^m, s_i') - v_i(a^k, s_i') \geq t_i(a^m) - t_i(a^k) \geq v_i(a^m, s_i) - v_i(a^k, s_i).$$

This further implies that

$$\begin{aligned} \int_{s_i}^{s_i'} \frac{\partial v_i}{\partial s_i}(a^m, s_i)\, D' s_i &= v_i(a^m, s_i') - v_i(a^m, s_i) \\ &\geq v_i(a^k, s_i') - v_i(a^k, s_i) = \int_{s_i}^{s_i'} \frac{\partial v_i}{\partial s_i}(a^k, s_i)\, D' s_i. \end{aligned}$$

Since $\frac{\partial v_i}{\partial s_i}(a^m, \cdot) > \frac{\partial v_i}{\partial s_i}(a^k, \cdot)$ by assumption, this implies that $s_i < s_i'$.

(Sufficiency) Fix $i \in I$ and $s_{-i} \in S_{-i}$.
Strategy-proofness:

Suppose that $s_i \in [y_i^{a^{k-1}a^k},\, y_i^{a^k a^{k+1}}]$ and that $s_i' \in [y_i^{a^{\ell-1}a^\ell},\, y_i^{a^\ell a^{\ell+1}}]$ for some $k \neq \ell$. Then

$$\begin{aligned} v_i(f(s_i, s_{-i}), s_i) - t_i(f(s_i, s_{-i})) &= v_i(a^k, s_i) - t_i(a^k) \\ &\geq v_i(a^\ell, s_i) - t_i(a^\ell) \\ &= v_i(f(s_i', s_{-i}), s_i) - t_i(f(s_i', s_{-i})), \end{aligned}$$

where the inequality follows since

$$s_i \in \left[y_i^{a^{k-1}a^k},\, y_i^{a^k a^{k+1}}\right] \quad \Rightarrow \quad v_i(a^k, s_i) - t_i(a^k) = \max_{a \in B_i(s_{-i})} v_i(a, s_i) - t_i(a)$$

as would be clear from Figure 1.

Ex post IR:

Since for $s_i \in [y_i^{a^k a^{k-1}}, y_i^{a^{k+1}a^k}]$, we have

$$\begin{aligned} & v_i(a^k, s_i) - t_i(a^k) \\ & \geq v_i(a^k, y_i^{a^k a^{k-1}}) - t_i(a^k) = v_i(a^{k-1}, y_i^{a^k a^{k-1}}) - t_i(a^{k-1}) \\ & \geq v_i(a^{k-1}, y_i^{a^{k-1}a^{k-2}}) - t_i(a^{k-1}) = v_i(a^{k-2}, y_i^{a^{k-1}a^{k-2}}) - t_i(a^{k-2}) \\ & \geq \cdots \\ & \geq -t_i(a^1) \geq 0. \end{aligned}$$

Proof of Theorem 1 We begin with the following lemma.

Lemma 6 *Suppose that (f, t) is an optimal ex post implementable coordinating scheme against two buyers under Assumption 2. Then*

1) *There exist no $0 \leq \alpha_1 < \beta_1 \leq 1$ such that $f(s) = 0$ for every $s \in (\alpha_1, \beta_1) \times (y_2^{01}, 1]$.*

2) *There exist no $0 \leq \alpha_2 < \beta_2 \leq 1$ such that $f(s) = 0$ for every $s \in (y_1^{10}, 1] \times (\alpha_2, \beta_2)$.*

3) *L_{11} is a rectangle with a non-empty interior such that $(1,1) \in L_{11}$ and $(0,0) \notin L_{11}$.*

Proof.

1. Suppose that there exist such α_1 and β_1 and denote $D = (\alpha_1, \beta_1) \times (y_2^{01}, 1]$. We will show that (f, t) is suboptimal. If $y_2^{01} = 0$ or 1, let $(\hat{f}, \hat{t})$ be such that $\hat{y}_1 = y_1$, $(\hat{y}_2^{01}, \hat{y}_2^{11}) = (\frac{1}{2}, y_2^{11})$, and

$$\hat{f}(s) = \begin{cases} 01 & \text{if } s \in (\alpha_1, \beta_1) \times (\frac{1}{2}, 1], \\ f(s) & \text{otherwise.} \end{cases}$$

Then $(\hat{f}, \hat{t})$ is ex post implementable and raises a strictly positive expected revenue $P(s \in (\alpha_1, \beta_1) \times (\frac{1}{2}, 1])\, v_2(01, \frac{1}{2})$ from D. When $y_2^{01} \in (0, 1)$, let $(\hat{f}, \hat{t})$ be such that $\hat{y} = y$ and

$$\hat{f}(s) = \begin{cases} 01 & \text{if } s \in D, \\ f(s) & \text{otherwise.} \end{cases}$$

Again, $(\hat{f}, \hat{t})$ is ex post implementable and raises a strictly positive expected revenue $P(s \in D)\, v_2(01, y_2^{01})$ from D. In both cases, $R(\hat{f}, \hat{t}) > R(f, t)$.

3. If $L_{11} \neq \emptyset$, then it contains $(1, 1)$ by Assumption 2 and Proposition 1. Suppose that int $L_{11} = \emptyset$. The optimality of (f, t) would then imply that $(1, 1) \in L_{10} \cup L_{01}$. Assume without loss of generality that $(1, 1) \in L_{10}$. We will show that (f, t) is dominated by an alternative scheme $(\hat{f}, \hat{t})$ defined as follows:

$$\hat{f}(s) = \begin{cases} 11 & \text{if } s \in [y_1^{10}, 1] \times [0, 1], \\ f(s) & \text{otherwise.} \end{cases} \qquad (\hat{y}_1^{10}, \hat{y}_1^{11}) = (y_1^{10}, y_1^{10}), \quad (\hat{y}_2^{01}, \hat{y}_2^{11}) = (y_2^{01}, 0)$$

Then $(\hat{f}, \hat{t})$ is ex post implementable. Furthermore, the expected revenue under $(\hat{f}, \hat{t})$ from $[y_1^{10}, 1] \times [0, 1]$ equals

$$P(L_{11})\, v_1(11, y_1^{10}).$$

This is strictly greater than the expected revenue under (f, t) from the same set since the latter is bounded above by

$$P(L_{11})\, v_1(10, y_1^{10}),$$

and $v_1(11, y_1^{10}) > v_1(10, y_1^{10})$ by Assumption 2. The expected revenue under $(\hat{f}, \hat{t})$ and that under (f, t) are the same elsewhere. We hence conclude that $R(f, t) < R(\hat{f}, \hat{t})$.

Next, we show that L_{11} is a rectangle. If $y_1^{10} \geq y_1^{11}$, then $L_{11} = [y_1^{11}, 1] \times [y_2^{11}, 1]$ or $L_{11} = [y_1^{11}, 1] \times [y_2^{11,01}, 1]$. It is also a rectangle if $y_2^{01} \geq y_2^{11}$. Suppose then that $y_1^{11,10} > y_1^{11}$ and $y_2^{11,01} > y_2^{11}$. L_{11} may fail to be a rectangle only if $L_{11} = [y_1^{11}, 1] \times [y_2^{11}, 1] \setminus [y_1^{11}, y_1^{11,10}] \times [y_2^{11}, y_2^{11,01}]$. However, if $f(s) = 10$ for $s \in [y_1^{11}, y_1^{11,10}] \times [y_2^{11}, y_2^{11,01}]$, f is not ex post IC since for $s_1 \in (y_1^{11}, y_1^{11,10})$, $L_2(10, s_1) = [y_2^{11}, y_2^{11,01}]$ and $L_2(11, s_1) = [y_2^{11,01}, 1]$ and violates Proposition 1. Likewise, $f(s) \neq 01$, 00 for $s \in [y_1^{11}, y_1^{11,10}] \times [y_2^{11}, y_2^{11,01}]$. Therefore, L_{11} is a rectangle in all cases. Finally, $(0, 0) \notin L_{11}$ since otherwise, $L_{11} = [0, 1]^2$ and the expected revenue under (f, t) would equal zero. ■

We now return to the proof of the theorem. We have the following four cases to consider depending on the relative orderings between y_1^{11} and y_1^{10}, and between y_2^{11} and y_2^{01}.[33]

Case 1) $y_1^{11} \leq y_1^{10}$ and $y_2^{11} \leq y_2^{01}$. For $s \ll (y_1^{11}, y_2^{11})$, $f(s) = 00$ by ex post IR. For $s \in [0, y_1^{10}) \times [0, y_2^{11})$, $f(s) = 00$ by ex post IR. It then follows from Lemma 6(1) that $y_1^{10} < 1$ and that $f(s) = 10$ for $s \in (y_1^{10}, 1] \times [0, y_2^{11})$. The symmetric argument shows that $y_1^{01} < 1$, $f(s) = 00$ for $s \in [0, y_1^{11}) \times (y_2^{11}, y_2^{01})$, and $f(s) = 01$ for $s \in [0, y_1^{11}) \times (y_2^{01}, 1]$. By Lemma 6(3), it must be the case that $f(s) = 11$ for $s \in (y_1^{10}, 1] \times (y_2^{01}, 1]$. By ex post IC, we must then have $f(s) = 11$ for $s \in (y_1^{11}, y_1^{10}] \times (y_2^{01}, 1]$ and $s \in (y_1^{10}, 1] \times (y_2^{11}, y_2^{01})$. Since L_{11} is a rectangle, we can then conclude that $L_{11} = [y_1^{11}, 1] \times [y_2^{11}, 1]$. This yields (A0).

Case 2) $y_1^{10} < y_1^{11} < y_1^{11,10}$ and $y_2^{01} < y_2^{11} < y_2^{11,01}$. For $s \ll (y_1^{10}, y_2^{01})$, $f(s) = 00$ by ex post IR. For $s \in (y_1^{10}, 1] \times (0, y_2^{01})$, $f(s) \in \{00, 10\}$ by ex post IR, and hence $f(s) = 10$ by Proposition 1 and Lemma 6(1). The symmetric argument shows that $f(s) = 01$ for $s \in [0, y_1^{10}) \times (y_2^{01}, 1]$. We now proceed by separately considering possible configurations of L_{11}. Since L_{11} is a rectangle containing $(1, 1)$ by Lemma 6(3), there are four possible cases as follows:

1) $L_{11} = [y_1^{11,10}, 1] \times [y_2^{11,01}, 1]$. Proposition 1 shows that $f(s) = 10$ for $s \in (y_1^{10}, y_1^{11,10}) \times (y_2^{11,01}, 1]$ and that $f(s) = 01$ for $s \in (y_1^{11,10}, 1] \times (y_2^{01}, y_2^{11,01})$. Proposition 1 further implies that $f(s) = 00$ for $s \in (y_1^{10}, y_1^{11,10}) \times (y_2^{01}, y_2^{11,01})$. This configuration, called (D), is depicted in Figure 7.

   Now consider configurations (B1) and (B2) which have the same t as (D) above. We show that (D) is dominated by (B1) if $t_1(10) \geq t_2(01)$ and dominated by (B2) if $t_1(10) \leq t_2(01)$. To see this, note that the expected revenue under (D)

[33] Since L_{11} has a non-empty interior, it must be the case that y_1^{11}, $y_2^{11} < 1$. This shows that when $y_1^{10} < y_1^{11}$, $y_1^{11} < y_1^{11,10}$, and that when $y_2^{01} < y_2^{11}$, $y_2^{11} < y_2^{11,01}$.

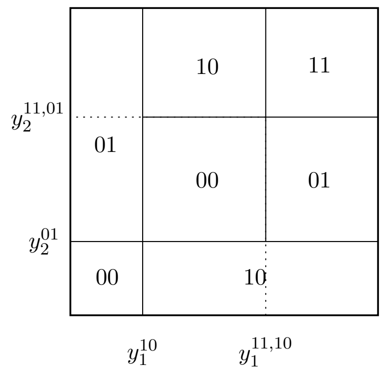

Figure 7: Configuration (D)

minus that under (B1) is written as

$$\begin{aligned} R^F - R^B &= P([y_1^{10}, y_1^{11,10}] \times [y_2^{01}, y_2^{11,01}])\,\{-t_1(10)\} \\ &+ P([y_1^{11,10}, 1] \times [y_2^{01}, y_2^{11}])\,\{t_2(01) - t_1(10)\} \\ &+ P([y_1^{11,10}, 1] \times [y_2^{11}, y_2^{11,01}])\,\{t_2(01) - t_1(11) - t_2(11)\}. \end{aligned}$$

Since $y_1^{11} > y_1^{10} \geq 0$ implies $t_1(11) = v_1(11, y_1^{11}) > v_1(10, y_1^{10}) = t_1(10) \geq 0$, this difference is strictly negative if $t_1(10) \geq t_2(01)$. Likewise, the expected revenue under (D) minus that under (B2) is written as

$$\begin{aligned} R^F - R^C &= P([y_1^{10}, y_1^{11,10}] \times [y_2^{01}, y_2^{11,01}])\,\{-t_2(01)\} \\ &+ P([y_1^{10}, y_1^{11}] \times [y_2^{11,01}, 1])\,\{t_1(10) - t_2(01)\} \\ &+ P([y_1^{11}, y_1^{11,10}] \times [y_2^{11,01}, 1])\,\{t_1(10) - t_1(11) - t_2(11)\}. \end{aligned}$$

Since $y_2^{11} > y_2^{01} \geq 0$ implies $t_2(11) = v_2(11, y_2^{11}) > v_2(01, y_2^{01}) = t_2(01) \geq 0$, the difference is strictly negative if $t_2(01) \leq t_1(10)$. Hence, (D) is never optimal.

2) $L_{11} = [y_1^{11,10}, 1] \times [y_2^{11}, 1]$. By Proposition 1, $f(s) = 10$ for $s \in (y_1^{10}, y_1^{11,10}) \times (y_2^{11}, 1]$. Furthermore, Lemma 6(1) shows that $f(s) = 10$ for $s \in (y_1^{10}, 1] \times (y_2^{01}, y_2^{11})$. This yields (B1).

3) $L_{11} = [y_1^{11}, 1] \times [y_2^{11,01}, 1]$. A similar reasoning as above shows that $f(s) = 01$ for $(y_1^{10}, 1] \times (y_2^{01}, 1] \setminus L_{11}$. This yields (B2).

4) $L_{11} = [y_1^{11}, 1] \times [y_2^{11}, 1]$. In this case, we have two possibilities:

   (a) $f(s) = 10$ for $s \in (y_1^{10}, 1] \times (y_2^{01}, y_2^{11})$ and $f(s) = 00$ for $s \in (y_1^{10}, y_1^{11}) \times (y_2^{11}, 1]$. This yields (C1).

(b) $f(s) = 01$ for $s \in (y_1^{10}, 1] \times (y_2^{01}, y_2^{11})$ and $f(s) = 00$ for $s \in (y_1^{10}, y_1^{11}) \times (y_2^{11}, 1]$. This yields (C2).

Case 3) $y_1^{10} < y_1^{11} < y_1^{11,10}$ and $y_2^{11} \leq y_2^{01}$.

By ex post IR and Lemma 6(1), $f(s) = 00$ for $s \in [0, y_1^{10}) \times [0, y_2^{01})$, $f(s) = 01$ for $s \in [0, y_1^{10}) \times (y_2^{01}, 1]$, and $f(s) = 10$ for $s \in [y_1^{10}, 1] \times [0, y_2^{11})$. By Lemma 6(3), L_{11} can be either (i) $[y_1^{11,10}, 1] \times [y_2^{11}, 1]$ or (ii) $s \in [y_1^{11}, 1] \times [y_2^{11}, 1]$. In case (i), it must be the case that $f(s) = 10$ for $s \in (y_1^{10}, y_1^{11,10}) \times (y_2^{11}, 1]$. Hence we obtain configuration (B1). In case (ii), $f(s) = 00$ for $s \in (y_1^{10}, y_1^{11}) \times (y_2^{11}, y_2^{01})$ by Proposition 1. Proposition 1 also implies that $f(s) \in \{01, 00\}$ for $s \in (y_1^{10}, y_1^{11}) \times (y_2^{01}, 1]$. However, we must have $f(s) = 01$ by Lemma 6(1). This yields (A2).

Case 4) $y_1^{11} \leq y_1^{10}$ and $y_2^{01} < y_2^{11} < y_2^{11,01}$.

The reasoning similar to that of Case 3 above yields (A1) and (B2).

. Since $y_1^{11} < 1$ by assumption, we then have $y_1^{10,11} > y_1^{11}$. Fix s such that $s_1 \in (y_1^{11}, y_1^{10,11})$ and $s_2 = y_2^{11}$. If $f(s) = 11$, then ex post implementability of (f, y) implies that $f(\hat{s}) = 00$ or 01 for $\hat{s}$ such that $\hat{s}_1 \in (y_1^{10}, y_1^{11})$ and $\hat{s}_2 = y_2^{11}$ by Proposition 1. In this case, However, this implies that (f, y) is not constrained efficient since it would require $f(s) = 11$. Intuitively, for type s_1 of buyer 1 that is just above y_1^{11}, the price of network 11 is too high giving him an incentive to underreport his type to achieve network 10.

Proof of Theorem 2 We first examine the optimality of configuration (B1), which requires $y_1^{10} < y_1^{11} < y_1^{11,10} < 1$. Since $y_1^{11,10}$ is uniquely determined as a function of $y_1 = (y_1^{10}, y_1^{11})$ in this case, we can use the pair of variables $(y_1^{10}, y_1^{10,11})$ instead of y_1 to express the seller's expected revenue.

$$
\begin{aligned}
& R^B(y_1^{11,10}, y_1^{10}, y_2^{11}, y_2^{01}) \\
&= \{1 - G_2(y_2^{11})\}\{1 - G_1(y_1^{11,10})\} \\
&\times \left\{ v_1(11, y_1^{11,10}) - v_1(10, y_1^{11,10}) + v_1(10, y_1^{10}) + v_2(11, y_2^{11}) \right\} \\
&+ \left[1 - G_1(y_1^{10}) - \{1 - G_2(y_2^{11})\}\{1 - G_1(y_1^{11,10})\} \right] v_1(10, y_1^{10}) \\
&+ G_1(y_1^{10})\{1 - G_2(y_2^{01})\} v_2(01, y_2^{01}) \\
&= \{1 - G_2(y_2^{11})\} \left\{ r_1(11, y_1^{11,10}) - r_1(10, y_1^{11,10}) + \{1 - G_1(y_1^{11,10})\} v_2(11, y_2^{11}) \right\} \\
&+ r_1(10, y_1^{10}) + G_1(y_1^{10}) r_2(01, y_2^{01}).
\end{aligned}
$$

Differentiation of R^B with respect to y_1^{10} yields:

$$\frac{\partial R^B}{\partial y_1^{11,10}}(y_1^{11,10}, y_1^{10}, y_2^{11}, y_2^{01}) = \frac{\partial r_1}{\partial s_1}(10, y_1^{10}) + g_1(y_1^{10})\, r_2(01, y_2^{01}).$$

If $y_1^{10} < \bar{z}_1^{10}$, then $\frac{\partial r_1}{\partial s_1}(10, y_1^{10}) > 0$ by Assumption 3 and hence the above partial derivative is strictly positive. It follows that the optimal y_1^{10} must satisfy $y_1^{10} \geq \bar{z}_1^{10}$. Next, differentiation of R^B with respect to $y_1^{11,10}$ yields:

$$\begin{aligned}
&\frac{\partial R^B}{\partial y_1^{11,10}}(y_1^{11,10}, y_1^{10}, y_2^{11}, y_2^{01}) \\
&= \{1 - G_2(y_2^{11})\} \left\{ \frac{\partial r_1}{\partial s_1}(11, y_1^{11,10}) - \frac{\partial r_1}{\partial s_1}(10, y_1^{11,10}) - g_1(y_1^{11,10})\, v_2(11, y_2^{11}) \right\}.
\end{aligned}$$

Since $y_1^{11,10} > y_1^{10} \geq \bar{z}_1^{10}$, $\frac{\partial r_1}{\partial s_1}(11, y_1^{11,10}) < \frac{\partial r_1}{\partial s_1}(10, y_1^{11,10})$ by Assumption 3. It follows that

$$\frac{\partial R^B}{\partial y_1^{11,10}}(y_1^{11,10}, y_1^{10}, y_2^{11}, y_2^{01}) < 0 \quad \text{for } y_1^{11,10} > y_1^{10},$$

suggesting that (B1) cannot be optimal. The symmetric discussion shows that (B2) is also suboptimal. Consider next configuration (C1) which requires $y_1^{10} < y_1^{11} < 1$. The expected revenue can be written as:

$$\begin{aligned}
&R^D(y_1^{11}, y_1^{10}, y_2^{11}, y_2^{01}) \\
&= \{1 - G_2(y_2^{11})\}\{1 - G_1(y_1^{11})\} \left\{ v_1(11, y_1^{11}) + v_2(11, y_2^{11}) \right\} \\
&+ \{1 - G_1(y_1^{10})\}\, G_2(y_2^{11})\, v_1(10, y_1^{10}) \\
&+ G_1(y_1^{10})\{1 - G_2(y_2^{01})\}\, v_2(01, y_2^{01}) \\
&= \{1 - G_2(y_2^{11})\}\, r_1(11, y_1^{11}) + \{1 - G_1(y_1^{11})\}\, r_2(11, y_2^{11}) \\
&+ G_2(y_2^{11})\, r_1(10, y_1^{10}) + G_1(y_1^{10})\, r_2(01, y_2^{01}).
\end{aligned} \tag{7}$$

Differentiation of R^D with respect to y_1^{10} yields

$$\frac{\partial R^D}{\partial y_1^{10}}(y_1^{11}, y_1^{10}, y_2^{11}, y_2^{01}) = G_2(y_2^{11})\, \frac{\partial r_1}{\partial s_1}(10, y_1^{10}).$$

Hence, the optimal y_1^{10} should equal $\bar{z}_1^{10}$. Differentiation of R^D with respect to y_1^{11} on the other hand yields

$$\frac{\partial R^D}{\partial y_1^{11}}(y_1^{11}, y_1^{10}, y_2^{11}, y_2^{01}) = \{1 - G_2(y_2^{11})\}\, \frac{\partial r_1}{\partial s_1}(11, y_1^{11}) - g_1(y_1^{11})\, r_2(11, y_2^{11}).$$

Since $y_1^{11} > y_1^{10} = \bar{z}_1^{10}$, $\frac{\partial r_1}{\partial s_1}(11, y_1^{11}) < \frac{\partial r_1}{\partial s_1}(10, y_1^{11}) < 0$ by Assumption 3. Therefore, the derivative is strictly negative and (C1) cannot be optimal. That (C2) cannot be optimal is shown by a symmetric argument. We are then left with configurations in (A), which require either $y_1^{11} \leq y_b$ or $y_2^{11} \leq y_2^{01}$. The expected revenue under each one of (A) has the same expression as that under (B2) in (7). It then follows from the discussion there that the optimal values satisfy $y_1^{10} = \bar{z}_1^{10}$, $y_2^{01} = \bar{z}_2^{01}$, $y_1^{11} \leq \bar{z}_1^{10}$ and $y_2^{11} \leq \bar{z}_2^{10}$. The optimal scheme is hence (A0), which is monotone.

Formula for $r^K(y^K)$: We can verify that the seller's expected revenue $r^K(y^K)$ from offering a menu $K = \{k_1, \ldots, k_m\}$ for $k_1 < \cdots < k_m$ equals

$$\begin{aligned}
&r^K(y^K) \\
&= \max\left\{0, G(\min\{y^{k_1k_2}, \ldots, y^{k_1k_m}\}) - G(y^{k_1})\right\} v^{k_1}(y^{k_1}) \\
&+ \sum_{n=2}^{m-1} \max\left\{0, G(\min\{y^{k_nk_{n+1}}, \ldots, y^{k_nk_m}\})\right. \\
&\left. - G(\max\{y^{k_n}, y^{k_1k_n}, \ldots, y^{k_{n-1}k_n}\})\right\} v^{k_n}(y^{k_n}) \\
&+ \max\left\{0, 1 - G(\max\{y^{k_m}, y^{k_1k_m}, \ldots, y^{k_{m-1}k_m}\})\right\} v^{k_m}(y^{k_m}).
\end{aligned}$$

When $y^{k_\ell k_n} < y^{k_m k_n}$ for every $\ell < m$ and n, we can express $r^K(y^K)$ as

$$r^K(y^K) = \sum_{n=2}^{m} \{r^{k_n}(y^{k_{n-1}k_n}) - r^{k_{n-1}}(y^{k_{n-1}k_n})\} + r^{k_1}(y^{k_1}). \tag{8}$$

Proof of Lemma 2 (i) For each $n \in N$, r^n is strictly log-concave with the (unique) maximizer $\bar{z}^n$ which satisfies $1 > \bar{z}^1 \geq \cdots \geq \bar{z}^I > 0$.

We first note that for $m < n$,

$$\frac{(v^n)'(\cdot)}{v^n(\cdot)} \leq \frac{(v^m)'(\cdot)}{v^m(\cdot)}. \tag{9}$$

This readily follows from Assumption 4, which implies that $\left(\frac{v^n(s)}{v^m(s)}\right)' \leq 0$ when $m < n$. Note now that

$$\begin{aligned}
(r^n)'(s) &= -g(s)\, v^n(s) + \{1 - G(s)\}\, (v^n)'(s) \\
&= \{1 - G(s)\}\, v^n(s) \left\{-\frac{g(s)}{1 - G(s)} + \frac{(v^n)'(s)}{v^n(s)}\right\} \\
&= r^n(s) \left\{-\frac{g(s)}{1 - G(s)} + \frac{(v^n)'(s)}{v^n(s)}\right\}.
\end{aligned}$$

Since $\frac{g(s)}{1-G(s)}$ is strictly increasing and $\frac{(v^n)'(s)}{v^n(s)}$ is strictly decreasing, $\frac{(r^n)'(\cdot)}{r^n(\cdot)}$ is strictly decreasing, implying that r^n is strictly log-concave. Hence, the maximizer $\bar{z}^n$ of r^n is unique and satisfies $\bar{z}^n \in (0,1)$ as $r^n(0) = r^n(1) = 0$. For $m < n$, $\bar{z}^m$ and $\bar{z}^n$ satisfy

$$\frac{(v^m)'(\bar{z}^m)}{v^m(\bar{z}^m)} = \frac{g(\bar{z}^m)}{1 - G(\bar{z}^m)} \quad \text{and} \quad \frac{(v^n)'(\bar{z}^n)}{v^n(\bar{z}^n)} = \frac{g(\bar{z}^n)}{1 - G(\bar{z}^n)}.$$

If $\bar{z}^m < \bar{z}^n$, then

$$\frac{g(\bar{z}^m)}{1 - G(\bar{z}^m)} < \frac{g(\bar{z}^n)}{1 - G(\bar{z}^n)},$$

and hence

$$\frac{(v^m)'(\bar{z}^m)}{v^m(\bar{z}^m)} < \frac{(v^n)'(\bar{z}^n)}{v^n(\bar{z}^n)},$$

which contradicts (9).

(ii) If $m < n$, $s < s'$, $s'' = \varphi^{MN}(s, s')$, and $(r^n)'(s'') \geq (r^m)'(s'')$, then $r^n(s') > r^{mn}(s, s')$.

We first verify

$$\frac{(r^m)'(s)}{(v^m)'(s)} > \frac{(r^n)'(s')}{(v^n)'(s')} \text{for } m < n \text{ and } s < s'. \tag{10}$$

The inequality is equivalent to

$$\{1 - G(s')\}\left[1 - \frac{g(s')}{1 - G(s')}\frac{v^n(s')}{(v^n)'(s')}\right] < \{1 - G(s)\}\left[1 - \frac{g(s)}{1 - G(s)}\frac{v^m(s)}{(v^m)'(s)}\right].$$

Since $s < s'$, this holds if $\frac{g(\cdot)}{1-G(\cdot)}$ is (strictly) increasing, and $\frac{(v^n)'(s')}{v^n(s')} \leq \frac{(v^m)'(s)}{v^m(s)}$. By the log-concavity of v^m, the latter inequality holds if $\frac{(v^n)'(s')}{v^n(s')} \leq \frac{(v^m)'(s')}{v^m(s')}$, which is true by (9).

We now show that $(r^n)'(s) \geq (r^m)'(s)$ implies that $(r^n)'(s), (r^m)'(s) \geq 0$. Note that $(r^n)'(s) \geq (r^m)'(s)$ is equivalent to

$$\frac{(v^n)'(s) - (v^m)'(s)}{v^n(s) - v^m(s)} \geq \frac{g(s)}{1 - G(s)}. \tag{11}$$

and that $(r^m)'(s) \geq 0$ is equivalent to

$$\frac{(v^m)'(s)}{v^m(s)} \geq \frac{g(s)}{1 - G(s)}. \tag{12}$$

Furthermore, since $\frac{(v^n)'(\cdot)}{v^n(\cdot)} \leq \frac{(v^m)'(\cdot)}{v^m(\cdot)}$ by (9),

$$\frac{(v^m)'(s)}{v^m(s)} \geq \frac{(v^n)'(s) - (v^m)'(s)}{v^n(s) - v^m(s)}. \tag{13}$$

(12) then follows from (13) and (11). This also implies $(r^n)'(s) \geq (r^m)'(s) \geq 0$.

Now, since $(r^n)'(s'')$, $(r^m)'(s'') \geq 0$, we have $(r^m)'(s) > 0$ and $(r^n)'(s') > 0$ for any s, $s' < s''$ by the strict log-concavity of r^m and r^n. It hence follows from (10) that for any such s and s',

$$\frac{(v^m)'(s)}{(v^n)'(s')} < \frac{(r^m)'(s)}{(r^n)'(s')}. \tag{14}$$

Now fix s'' such that $r^n(s'') > r^m(s'')$, and consider the following functions of $s \in [0, s'']$:

$$s' = (v^n)^{-1}\left(v^m(s) + v^n(s'') - v^m(s'')\right),$$

and

$$s' = (r^n)^{-1}\left(r^m(s) + r^n(s'') - r^m(s'')\right).$$

Both functions are differentiable over the domain, and the graph of the former lies above that of the latter since both of them go through (s'', s'') and have a single crossing point because of (14), which shows that the latter has a steeper slope than the former at any point of intersection between the two. Hence, for any $s < s''$, we have

$$(v^n)^{-1}\left(v^m(s) + v^n(s'') - v^m(s'')\right) > (r^n)^{-1}\left(r^m(s) + r^n(s'') - r^m(s'')\right).$$

In other words, whenever $v^n(s') = v^m(s) + v^n(s'') - v^m(s'')$, $r^n(s') > r^m(s) + r^n(s'') - r^m(s'')$. Equivalently, we have $r^n(s') > r^m(s) + r^n(s'') - r^m(s'')$ when $s'' = \varphi^{MN}(s, s')$, and $s < s'$. The desired conclusion then follows since by (8),

$$r^{MN}(s, s') = r^n(s'') - r^m(s'') + r^m(s).$$

Lemma 7 *Suppose that $v^k(s_i) = \rho^k \underline{v}^k(s_i)$ for every $k \in K$, where $0 < \rho^1 \leq \cdots \leq \rho^I$ and $\underline{v}^k : [0,1] \to \boldsymbol{R}_+$ satisfies Assumption 4 for every $k \in N$. Then there exists $\varepsilon > 0$ such that the following hold when $\max_{2 \leq k \leq I} \frac{\rho^{k-1}}{\rho^k} < \varepsilon$.*

1) For any $m < n$ and y^m, $y^n \in (0,1)$, if $r^n(y^n) \leq r^m(y^m)$ and $y^n < y^m$, then

$$\frac{(r^n)'(y^n)}{r^n(y^n)}\{G(y^m) - G(y^n)\} > g(y^n).$$

2) Let $\bar{\mu}^k = \sup_{s_i} \frac{v^{k-1}(s_i)}{v^k(s_i)} = \frac{v^{k-1}(1)}{v^k(1)}$ and $\underline{\mu}^k = \inf_{s_i} \frac{v^{k-1}(s_i)}{v^k(s_i)} = \lim_{s_i \to 0} \frac{v^{k-1}(s_i)}{v^k(s_i)}$ for $k \in N$. Then

$$\bar{\mu}^n < \frac{1 + (n-2)\prod_{k=2}^{n-1} \underline{\mu}^k}{n-1}$$

for $n = 2, \ldots, I-1$.

Proof of Lemma 7 Since the density g is continuous and strictly positive over $[0,1]$, we have

$$\frac{G(y^m)-G(y^n)}{y^m-y^n} \geq \beta\, g(y^n),$$

where $\beta = \frac{\min_{s_i}\, g(s_i)}{\max_{s_i}\, g(s_i)} > 0$. Hence, the condition in Lemma 7(1) holds if

$$\beta(y^m-y^n)\frac{(r^n)'(y^n)}{r^n(y^n)} > 1. \tag{15}$$

Since $r^m(y^m) \geq r^n(y^n)$ and $y^m > y^n$ imply $v^m(y^m) > v^n(y^n)$, dividing through by $v^n(y^m)$, we obtain

$$\frac{\rho^m \underline{v}^m(y^m)}{\rho^n \underline{v}^n(y^n)} = \frac{v^m(y^m)}{v^n(y^m)} > \frac{v^n(y^n)}{v^n(y^m)} = \frac{\underline{v}^n(y^n)}{\underline{v}^n(y^m)} \geq \frac{\min_{s_i}\,(\underline{v}^n)'(s_i)}{\max_{s_i}\,(\underline{v}^n)'(s_i)}\,\frac{y^n}{y^m}, \tag{16}$$

where the last inequality holds because $\frac{\underline{v}^n(y)}{y} = \frac{\underline{v}^n(y)-\underline{v}^n(0)}{y-0}$ is the average slope of $\underline{v}^n$ over $[0,y]$, and hence

$$\frac{\underline{v}^n(y^n)}{y^n}\,\frac{1}{\min_{s_i}\,(\underline{v}^n)'(s_i)} \geq 1 \geq \frac{\underline{v}^n(y^m)}{y^m}\,\frac{1}{\max_{s_i}\,(\underline{v}^n)'(s_i)}.$$

It follows from (16) that

$$\frac{y^m}{y^n} > \frac{\min_{s_i}\,(\underline{v}^n)'(s_i)}{\max_{s_i}\,(\underline{v}^n)'(s_i)}\,\frac{\rho^n}{\rho^m}\,\frac{\underline{v}^n(y^m)}{\underline{v}^m(y^m)} \geq \frac{\min_{s_i}\,(\underline{v}^n)'(s_i)}{\max_{s_i}\,(\underline{v}^n)'(s_i)}\,\frac{\rho^n}{\rho^m}.$$

Hence, (15) is implied by

$$\beta\left(\frac{\min_{s_i}\,(\underline{v}^n)'(s_i)}{\max_{s_i}\,(\underline{v}^n)'(s_i)}\,\frac{\rho^n}{\rho^m}-1\right)\left(y^n\frac{(r^n)'(y^n)}{r^n(y^n)}\right) > 1.$$

Suppose now that $\frac{\rho^m}{\rho^n} \to 0$. Then the quantity in the first brackets $\to \infty$. As for the second brackets, note that $y^n \to 0$ since $\frac{\underline{v}^n(y^n)}{\underline{v}^m(y^m)} < \frac{\rho^m}{\rho^n}$. Note also that

$$s_i\frac{(r^n)'(s_i)}{r^n(s_i)} = s_i\frac{(\underline{v}^n)'(s_i)}{\underline{v}^n(s_i)} - s_i\frac{g(s_i)}{1-G(s_i)} \to 1,$$

because $\lim_{s_i\to 0}\frac{s_i}{\underline{v}^n(s_i)} = \lim_{s_i\to 0}\frac{1}{(\underline{v}^n)'(s_i)} = \frac{1}{(\underline{v}^n)'(0)}$ by L'Hospital's rule. Hence, (15) holds when $\frac{\rho^m}{\rho^n}$ is sufficiently small.

Next, the condition in Lemma 7(2) holds if we take $\varepsilon < \frac{1}{n-1}$: $\frac{\rho^{k-1}}{\rho^k} < \varepsilon$ would then imply

$$\bar{\mu}^k = \sup_{s_i}\frac{v^{k-1}(s_i)}{v^k(s_i)} = \frac{\rho^{k-1}}{\rho^k}\sup_{s_i}\frac{\underline{v}^{k-1}(s_i)}{\underline{v}^k(s_i)} < \varepsilon < \frac{1}{n-1} < \frac{1+(n-2)\pi_{k=2}^{n-1}\,\underline{\mu}^k}{n-1}.$$

Proof of Theorem 3 The proof of the theorem begins with Lemmas 8, 9 and 10.

Lemma 8 *Under Assumption 4, if* $z \in \text{argmax}_{y \in S}\, w(y)$*, then there exists a permutation* $\pi_1, \ldots, \pi_I$ *of* $1, \ldots, I$ *such that for each* $n = 1, \ldots, I$*,*

$$r^{\pi_n}(z^{\pi_n}) = \max_{\emptyset \neq L \subset \Pi_n} r^L(z^L),$$

where $\Pi_I = N$ *and* $\Pi_n = N \setminus \{\pi_{n+1}, \ldots, \pi_I\}$*. In particular,* $r^{\pi_I}(z^{\pi_I}) \geq \cdots \geq r^{\pi_1}(z^{\pi_1})$*.*

Proof.
Step 1. $r^{\pi_I}(z^{\pi_I}) = \max_{\emptyset \neq J \subset I} r^J(z^J)$ for some $\pi_I \in I$.

Suppose to the contrary that there exists K such that $K \subset I$, $|K| \geq 2$, and

$$r^K(z^K) = \max_{\emptyset \neq J \subset I} r^J(z^J). \tag{17}$$

If there exists more than one such set that satisfies (17), choose any one with the smallest cardinality $|K|$. Write $K = \{\kappa_1, \ldots, \kappa_n\}$ for some $2 \leq n \leq I$ and $\kappa_1 < \cdots < \kappa_n$. Now consider y such that

$$0 < y^{\kappa_1} < \cdots < y^{\kappa_n} < 1, \text{ and } y^{\kappa_{n-1}} < y^{\kappa_n} < y^{\kappa_{n-1}\kappa_n} < 1. \tag{18}$$

Given that K has the smallest cardinality, z must satisfy (18): Otherwise, there is redundancy in K and we can find a strictly smaller set $\hat{K} \subset K$ such that $r^{\hat{K}}(z^{\hat{K}}) = r^K(z^K)$. Write $y^{\kappa_1\kappa_2}, \ldots, y^{\kappa_{n-1}\kappa_n}$ as functions of $y^{\kappa_1}, \ldots, y^{\kappa_n}$ as follows:

$$y^{\kappa_1\kappa_2} = y^{\kappa_1\kappa_2}(y^{\kappa_1}, y^{\kappa_2}), \ldots, y^{\kappa_{n-1}\kappa_n} = y^{\kappa_{n-1}\kappa_n}(y^{\kappa_{n-1}}, y^{\kappa_n}).$$

Let

$$z^{\kappa_1\kappa_2} = y^{\kappa_1\kappa_2}(z^{\kappa_1}, z^{\kappa_2}), \ldots, z^{\kappa_{n-1}\kappa_n} = y^{\kappa_{n-1}\kappa_n}(z^{\kappa_{n-1}}, z^{\kappa_n}).$$

$$z^1 = \varphi^{\kappa_1\kappa_2}(z^{\kappa_1}, z^{\kappa_2}), \ldots, \zeta^{n-1} = \varphi^{\kappa_{n-1}\kappa_n}(z^{\kappa_{n-1}}, z^{\kappa_n}).$$

By our choice of K, we must have $0 < \zeta^1 < \zeta^2 < \cdots < \zeta^{n-1} < 1$, $z^{\kappa_1} < z^{\kappa_2} < \zeta^1$, $z^{\kappa_2} < z^{\kappa_3} < \zeta^2$,...,$z^{\kappa_{n-1}} < z^{\kappa_n} < \zeta^{n-1}$.

Now define $\hat{w}$ by

$$\hat{w}(y) = \sum_{\substack{1 \in J \subset N \\ \kappa_n \notin J}} Q^J(y) \max_{\emptyset \neq L \subset J} r^L(y^L) + r^K(y^K) \sum_{\substack{1 \in J \subset N \\ K \subset J}} Q^J(y).$$

Note that $S_{-i}(J,z)=\emptyset$ for any $J\subset N$ such that $K\not\subset J$ and $\kappa_n\in J$: Suppose to the contrary that $s_{-i}\in S_{-i}(J,z)$ for such a J. Since $K\not\subset J$, there exists $k\in K$ such that $\lambda^{k-1}<z^k$. However, since $\lambda^{\kappa_n-1}\leq\lambda^{k-1}<z^k\leq z^{\kappa_n}$ for any such s_{-i}, we must have $\lambda^{\kappa_n-1}<z^{\kappa_n}$, contradicting the assumption that $\kappa_n\in J$. Hence, $\sum_{\substack{1\in J\subset N\\ \kappa_n\in J,K\not\subset J}} Q^J(z)=0$ so that

$$\hat{w}(z)=\sum_{\substack{1\in J\subset N\\ K\not\subset J}} Q^J(z)\max_{\emptyset\neq L\subset J} r^L(z^L)+r^K(z^K)\sum_{\substack{1\in J\subset N\\ K\subset J}} Q^J(z).$$

This suggests that $\hat{w}(y)\leq w(y)$ for any y, and $\hat{w}(z)=w(z)$ by our hypothesis. From the definition of $\hat{w}$, we have

$$\frac{\partial\hat{w}}{\partial y^{\kappa_n}}(z)=\sum_{\substack{1\in J\subset N\\ \kappa_n\notin J}}\frac{\partial Q^J}{\partial y^{\kappa_n}}(z)\max_{\emptyset\neq L\subset J} r^L(z^L)$$
$$+r^K(z^K)\sum_{\substack{1\in J\subset N\\ K\subset J}}\frac{\partial Q^J}{\partial y^{\kappa_n}}(z)+\frac{\partial r^K}{\partial y^{\kappa_n}}(z^K)\sum_{\substack{1\in J\subset N\\ K\subset J}} Q^J(z).$$

Using

$$\sum_{\substack{1\in J\subset N\\ \kappa_n\notin J}}\frac{\partial Q^J}{\partial y^{\kappa_n}}(z)=-\sum_{\substack{1\in J\subset N\\ K\subset J}}\frac{\partial Q^J}{\partial y^{\kappa_n}}(z),$$

we observe that the FOC $\frac{\partial\hat{w}}{\partial y^{\kappa_n}}(z)=0$ is given by

$$\sum_{\substack{1\in J\subset N\\ \kappa_n\notin J}}\frac{\partial Q^J}{\partial y^{\kappa_n}}(z^J)\left\{\max_{\emptyset\neq L\subset J} r^L(z^L)-r^K(z^K)\right\}+\frac{\partial r^K}{\partial y^{\kappa_n}}(z^K)\sum_{\substack{1\in J\subset N\\ K\subset J}} Q^J(z)=0.$$

Note that the bracketed term is negative and that $\sum_{\substack{1\in J\subset N\\ K\subset J}} Q^J(z)>0$ since $z^{\kappa_1}<\cdots<z^{\kappa_n}<1$. It follows that this equation holds only if

$$\frac{\partial r^K}{\partial y^{\kappa_n}}(z^K)\geq 0.$$

Recall from (8) that

$$r^K(z^K)=r^{\kappa_1}(z^{\kappa_1})+\sum_{\ell=2}^{n}\left\{r^{\kappa_\ell}(z^{\kappa_{\ell-1}\kappa_\ell})-r^{\kappa_{\ell-1}}(z^{\kappa_{\ell-1}\kappa_\ell})\right\}.$$

The derivative of r^K is hence given by

$$\frac{\partial r^K}{\partial y^{\kappa_n}}(z^K)=\left\{(r^{\kappa_n})'(z^{\kappa_{n-1}\kappa_n})-(r^{\kappa_{n-1}})'(z^{\kappa_{n-1}\kappa_n})\right\}\frac{\partial y^{\kappa_{n-1}\kappa_n}}{\partial y^{\kappa_n}}(z^{\kappa_{n-1}},z^{\kappa_n}).$$

Since $\frac{\partial y^{\kappa_{n-1}\kappa_n}}{\partial y^{\kappa_n}}(z^{\kappa_{n-1}}, z^{\kappa_n}) > 0$, $(r^{\kappa_n})'(z^{\kappa_{n-1}\kappa_n}) - (r^{\kappa_{n-1}})'(z^{\kappa_{n-1}\kappa_n}) \geq 0$. Since $z^{\kappa_\ell - 1} < z^{\kappa_\ell} < z^{\kappa_{\ell-1}\kappa_\ell} \leq z^{\kappa_{n-1}\kappa_n}$ for each $\ell \leq n$, we have by Lemma 2

$$r^{\kappa_\ell}(z^{\kappa_\ell}) > r^{\kappa_{\ell-1}\kappa_\ell}(z^{\kappa_{\ell-1}}, z^{\kappa_\ell}) = r^{\kappa_{\ell-1}}(z^{\kappa_{\ell-1}}) + r^{\kappa_\ell}(z^{\kappa_{\ell-1}\kappa_\ell}) - r^{\kappa_{\ell-1}}(z^{\kappa_{\ell-1}\kappa_\ell}).$$

Substituting this into (6), we obtain

$$\begin{aligned}
& r^K(z^K) \\
&= r^{\kappa_1}(z^{\kappa_1}) + \sum_{\ell=2}^{n} \Big\{ r^{\kappa_\ell}(z^{\kappa_{\ell-1}\kappa_\ell}) - r^{\kappa_{\ell-1}}(z^{\kappa_{\ell-1}\kappa_\ell}) \Big\} \\
&< r^{\kappa_1}(z^{\kappa_1}) + \sum_{\ell=2}^{n} \Big\{ r^{\kappa_\ell}(z^{\kappa_\ell}) - r^{\kappa_{\ell-1}}(z^{\kappa_{\ell-1}}) \Big\} \\
&= r^{\kappa_n}(z^{\kappa_n}).
\end{aligned}$$

This however contradicts our original supposition.

Step 2.

As an induction hypothesis, suppose that for $m = \mu + 1, \ldots, I$, there exists $\pi_m \in \Pi_m$ such that

$$r^{\pi_m}(z^{\pi_m}) = \max_{\emptyset \neq L \subset \Pi_m} r^L(z^L).$$

We will show that

$$r^K(z^K) < \max_{\emptyset \neq J \subset \Pi_\mu} r^J(z^J)$$

for any K such that $K \subset \Pi_\mu$ and $|K| \geq 2$. Suppose to the contrary that $r^K(z^K) = \max_{\emptyset \neq J \subset \Pi_\mu} r^J(z^J)$ for some $K = \{\kappa_1, \ldots, \kappa_n\}$ such that $K \subset \Pi_\mu$ and $n \geq 2$. Define

$$\begin{aligned}
\hat{w}(y) = & \sum_{\substack{1 \in J \subset \Pi_\mu \\ \kappa_n \notin J}} Q^J(y) \max_{\emptyset \neq L \subset J} r^L(y^L) + r^K(y^K) \sum_{\substack{1 \in J \subset \Pi_\mu \\ K \subset J}} Q^J(y) \\
& + \sum_{m=\mu+1}^{I} r^{\pi_m}(y^{\pi_m}) \sum_{\substack{1 \in J \subset \Pi_m \\ \pi_m \in J}} Q^J(y).
\end{aligned}$$

As in Step 1, we observe that $S_{-i}(J, z) = \emptyset$ for any J such that $\kappa_n \in J$ and $K \not\subset J$.[34] Then $\hat{w}(y) \leq w(y)$ for any y and $\hat{w}(z) = w(z)$ by the induction hypothesis. Since

$$\sum_{\substack{1 \in J \subset \Pi_m \\ \pi_m \in J}} Q^J(y) = P\Big(\lambda^{\pi_m - 1} \geq y^{\pi_m}, \max_{\ell > m}(\lambda^{\pi_\ell - 1} - y^{\pi_\ell}) < 0\Big),$$

[34] The reasoning is the same as that following the definition of $\hat{w}$ in Step 1.

the third term in the definition of $\hat{w}$ is independent of y^{Π_μ}. It follows that

$$\frac{\partial \hat{w}}{\partial y^{\kappa_n}}(z) = \sum_{\substack{1\in J\subset\Pi_\mu \\ \kappa_n\notin J}} \frac{\partial Q^J}{\partial y^{\kappa_n}}(z) \max_{\emptyset\neq L\subset J} r^L(z^L) + r^K(z^K) \sum_{\substack{1\in J\subset\Pi_\mu \\ K\subset J}} \frac{\partial Q^J}{\partial y^{\kappa_n}}(z)$$
$$+ \frac{\partial r^K}{\partial y^{\kappa_n}}(z^K) \sum_{\substack{1\in J\subset\Pi_\mu \\ K\subset J}} Q^J(z).$$

Noting that $\sum_{\substack{1\in J\subset\Pi_\mu \\ \kappa_n\notin J}} \frac{\partial Q^J}{\partial y^{\kappa_n}}(z) = -\sum_{\substack{1\in J\subset\Pi_\mu \\ K\subset J}} \frac{\partial Q^J}{\partial y^{\kappa_n}}(z)$, we conclude as before that $\frac{\partial r^K}{\partial y^{\kappa_n}}(z^K) \geq 0$. Using the same logic as in Step 1, we can then derive the contradiction that $r^K(z^K) < r^{\kappa_n}(z^{\kappa_n})$. This advances the induction step and completes the proof. ■

Lemma 9 *Under Assumption 4, if* $z \in \operatorname{argmax}_{y\in S} w(y)$*, then* $z \in (0,1)^I$.

Proof. Suppose not and take the largest n for which $z^{\pi_n} = 0$ or 1, where $\pi_1, \ldots, \pi_I$ are as defined in Lemma 8. It would then follow that $r^{\pi_n}(z^{\pi_n}) = 0$ and hence that $r^{\pi_\ell}(z^{\pi_\ell}) = 0$ for every $\ell < n$ as well. Define $\hat{z}$ to be such that

$$\hat{z}^{\pi_\mu} = \begin{cases} z^{\pi_\mu} & \text{if } \mu \neq n, \\ \frac{1}{2} & \text{if } \mu = n. \end{cases}$$

We then have

$$\begin{aligned} w(\hat{z}) &\geq \sum_{\mu=1}^{I} r^{\pi_\mu}(\hat{z}^{\pi_\mu}) \sum_{\substack{1\in J\subset\Pi_\mu \\ \pi_\mu\in J}} Q^J(\hat{z}) \\ &= r^{\pi_n}(\hat{z}^{\pi_n})\, P\Big(\lambda^{\pi_n - 1} \geq \hat{z}^{\pi_n},\, \max_{\ell>n}(\lambda^{\pi_\ell - 1} - z^{\pi_\ell}) < 0\Big) \\ &+ \sum_{\mu=n+1}^{I} r^{\pi_\mu}(z^{\pi_\mu})\, P\Big(\lambda^{\pi_\mu - 1} \geq z^{\pi_\mu},\, \max_{\ell>\mu}(\lambda^{\pi_\ell - 1} - z^{\pi_\ell}) < 0\Big) \\ &> \sum_{\mu=n+1}^{I} r^{\pi_\mu}(z^{\pi_\mu})\, P\Big(\lambda^{\pi_\mu - 1} \geq z^{\pi_\mu},\, \max_{\ell>\mu}(\lambda^{\pi_\ell - 1} - z^{\pi_\ell}) < 0\Big) \\ &= w(z), \end{aligned}$$

where the inequality holds since $r^{\pi_n}(\frac{1}{2}) > 0$ and $z^{\pi_\ell} > 0$ for $\ell > n$. This is a contradiction. ■

Lemma 10 *Suppose that Assumptions 4 and the condition in 7(1) hold. If* $z \in \operatorname{argmax}_{y\in S} w(y)$, *then* $\pi_\mu = \mu$ *for* $\mu = 1, \ldots, I$, *or equivalently,*

$$r^1(z^1) \le \cdots \le r^I(z^I).$$

Proof. Note first that $\operatorname{argmax}_y w(y) \neq \emptyset$ since w is a continuous function over the compact domain S. Let $z = (z^1, \ldots, z^I) \in \operatorname{argmax}_y w(y)$ be any maximizer. We prove the claim by induction over $\mu = 1, \ldots, I$.

Step 1. $r^I(z^I) = \max_{\emptyset \neq J \subset I} r^J(z^J)$.

Given the conclusion of Lemma 8, the claim is equivalent to $\pi_I = I$, where $\pi_1, \ldots . \pi_I$ are as defined there. Suppose to the contrary that $\pi_I < I$, and take $n < I$ such that $\pi_n = I$. If we define

$$\begin{aligned}\hat{w}(y) &= \sum_{\mu=1}^{I} r^{\pi_\mu}(y^{\pi_\mu}) \sum_{\substack{1\in J\subset \Pi_\mu \\ \pi_\mu \in J}} Q^J(y) \\ &= \sum_{\mu=1}^{I} r^{\pi_\mu}(y^{\pi_\mu})\, P\Big(\lambda^{\pi_\mu - 1} \ge y^{\pi_\mu},\, \max_{\ell>\mu} (\lambda^{\pi_\ell - 1} - y^{\pi_\ell}) < 0\Big),\end{aligned}$$

then $\hat{w}(y) \le w(y)$ for any y and $\hat{w}(z) = w(z)$. Differentiating $\hat{w}$ with respect to $y^{\pi_n} = y^I$, we obtain

$$\begin{aligned}\frac{\partial \hat{w}}{\partial y^{\pi_n}}(y) &= \sum_{\mu=1}^{n-1} r^{\pi_\mu}(y^{\pi_\mu}) \sum_{\substack{1\in J\subset \Pi_\mu \\ \pi_\mu \in J}} \frac{\partial Q^J}{\partial y^{\pi_n}}(y) \\ &\quad + (r^{\pi_n})'(y^{\pi_n}) \sum_{\substack{1\in J\subset \Pi_n \\ \pi_n \in J}} Q^J(y) \\ &\quad + r^{\pi_n}(y^{\pi_n}) \sum_{\substack{1\in J\subset \Pi_n \\ \pi_n \in J}} \frac{\partial Q^J}{\partial y^{\pi_n}}(y).\end{aligned} \tag{19}$$

Since $z \in (0,1)^I$ by Lemma 9, the FOC $\frac{\partial \hat{w}}{\partial y^{\pi_n}}(z) = 0$ holds at $y = z$. Furthermore, since $r^{\pi_\mu}(z^{\pi_\mu}) \le r^{\pi_n}(z^{\pi_n})$ for every $\mu < n$, and

$$\sum_{\substack{1\in J\subset \Pi_\mu \\ \pi_\mu \in J}} \frac{\partial Q^J}{\partial y^{\pi_n}}(z) = -\sum_{\substack{1\in J\subset \Pi_n \\ \pi_n \in J}} \frac{\partial Q^J}{\partial y^{\pi_n}}(z) > 0, \tag{20}$$

the sum of the first and third terms on the right-hand side of (19) evaluated at z is ≤ 0, implying that $(r^{\pi_n})'(z^{\pi_n}) \geq 0$. This and $r^{\pi_\ell}(z^{\pi_\ell}) \geq r^{\pi_n}(z^{\pi_n})$ for $\ell > n$ together imply that $z^{\pi_\ell} > z^{\pi_n}$ for $\ell > n$. Let z^{π_m} be the smallest among $z^{\pi_{n+1}}, \ldots, z^{\pi_I}$.

For any y such that $y^{\pi_n} < y^{\pi_m} < \min_{\ell > n, \ell \neq m} y^{\pi_\ell}$, we now evaluate the probability $\sum_{\substack{1 \in J \subset \Pi_n \\ \pi_n \in J}} Q^J(y)$ appearing in (19) in two ways. First, assigning just one of $I-1$ types to the interval $[y^{\pi_n}, y^{\pi_m}]$, we see that

$$\begin{aligned}\sum_{\substack{1 \in J \subset \Pi_n \\ \pi_n \in J}} Q^J(y) &= P\Big(\lambda^{\pi_n - 1} \geq y^{\pi_n},\, \max_{\ell > n} (\lambda^{\pi_\ell - 1} - y^{\pi_\ell}) < 0\Big) \\ &= \binom{I-1}{1} \{G(y^{\pi_m}) - G(y^{\pi_n})\} \\ &\quad \times P\Big(\lambda_{I-2}^{\pi_n - 2} \geq y^{\pi_n},\, \max_{\ell > n} (\lambda_{I-2}^{\pi_\ell - 1} - y^{\pi_\ell}) < 0\Big),\end{aligned} \tag{21}$$

where λ_{I-2}^k is the kth largest value among $I-2$ types, and $\lambda_k^\ell = 0$ whenever $k < \ell$. Second, suppose we assign p of $I-1$ types to $[y^{\pi_n}, y^{\pi_m}]$ and the remaining $q = I-1-p$ types to $(y^{\pi_m}, 1]$. In this case, $q < \pi_m - 1$ must hold since $\lambda^{\pi_m - 1} < y^{\pi_m}$. Hence,

$$\begin{aligned}\sum_{\substack{1 \in J \subset \Pi_n \\ \pi_n \in J}} Q^J(y) &= P\Big(\lambda^{\pi_n - 1} \geq y^{\pi_n}, \max_{\ell > n} (\lambda^{\pi_\ell - 1} - y^{\pi_\ell}) < 0\Big) \\ &= \sum_{\substack{p+q=I-1 \\ q < \pi_m - 1}} \binom{I-1}{p} \{G(y^{\pi_m}) - G(y^{\pi_n})\}^p \\ &\qquad \times P\Big(\max_{\ell > n, \ell \neq m} (\lambda_q^{\pi_\ell - 1} - y^{\pi_\ell}) < 0,\, \lambda_q^q \geq y^{\pi_m}\Big).\end{aligned} \tag{22}$$

Differentiating (22) with respect to $y^{\pi_n} = y^I$ and rearranging, we obtain

$$\begin{aligned}&\sum_{\substack{1 \in J \subset \Pi_n \\ \pi_n \in J}} \frac{\partial Q^J}{\partial y^{\pi_n}}(y) \\ &= -(I-1)\, g(y^{\pi_n})\, P\Big(\lambda_{I-2}^{\pi_n - 2} \geq y^{\pi_n},\, \max_{\ell > n} (\lambda_{I-2}^{\pi_\ell - 1} - y^{\pi_\ell}) < 0\Big).\end{aligned} \tag{23}$$

Now substitute (21) and (23) into (19) and set $y = z$ to get

$$\begin{aligned}&\frac{\partial \hat{w}}{\partial y^{\pi_n}}(z) \\ &\geq P\Big(\lambda_{I-2}^{\pi_n - 2} \geq z^{\pi_n},\, \max_{\ell > n} (\lambda_{I-2}^{\pi_\ell - 1} - z^{\pi_\ell}) < 0\Big) \\ &\times (I-1) \Big[-g(z^{\pi_n})\, r^{\pi_n}(z^{\pi_n}) + (r^{\pi_n})'(z^{\pi_n})\, \{G(z^{\pi_m}) - G(z^{\pi_n})\}\Big] \\ &> 0.\end{aligned}$$

where the first inequality follows from the fact that the first term on the right-hand side of (19) is positive (*i.e.,* (20)), and the second from Lemma 7(1) along with the fact that $z^{\pi_n} < z^{\pi_m}$ and $r^{\pi_m}(z^{\pi_m}) \geq r^{\pi_n}(z^{\pi_n})$. We have hence derived a contradiction to the fact that z is an interior maximizer.

Step 2. For $n = 1, \ldots, I-1$, $r^n(z^n) = \max_{\ell \leq n} r^\ell(z^\ell)$.

As an induction hypothesis, suppose that the statement holds for $n+1, \ldots, I$. Define

$$\begin{aligned}
\hat{w}_n(y) &= \sum_{k=1}^{n} r^{\pi_k}(y^{\pi_k}) \sum_{\substack{1 \in J \subset \Pi_k \\ \pi_k \in J}} Q^J(y) + \sum_{k=n+1}^{I} r^k(y^k) \sum_{\substack{1 \in J \subset N \\ \max J = k}} Q^J(y) \\
&= \sum_{k=1}^{n} r^{\pi_k}(y^{\pi_k}) \sum_{\substack{1 \in J \subset \Pi_k \\ \pi_k \in J}} Q^J(y) \\
&+ \sum_{k=n+1}^{I} P\Big(\lambda^{k-1} \geq y^k, \max_{\ell > k} (\lambda^{\ell-1} - y^\ell) < 0\Big) r^k(y^k).
\end{aligned}$$

We then have $\hat{w}_n(y) \leq w(y)$ for any y, and by the induction hypothesis, $\hat{w}_n(z) = w(z)$. Hence, since z is a maximizer of w, it is a maximizer of $\hat{w}_n$ as well. Note that the second term on the right-hand side above is independent of $(y^1, \ldots, y^n)$, and the first term has the same form as $\hat{w}$ in Step 1 with the only exception that n replacing I. This implies that the same reasoning as that in step 1 proves

$$r^n(z^n) = \max_{\ell \leq n} r^\ell(z^\ell).$$

■

We now return to the proof of the theorem. We will show that any maximizer z of $w : \boldsymbol{R}^I \to \boldsymbol{R}$ satisfies $z^I \leq \cdots \leq z^1$ under Lemma 7(2). For $n = 1, \ldots, I$ and $y \in S$, define

$$R^n(y) = \sum_{\substack{1 \in J \subset I_n \\ n \in J}} Q^J(y) = P(\lambda^{n-1} \geq y^n, \max_{\ell > n} (\lambda^{\ell-1} - y^\ell) < 0).$$

For any y such that $y^I > y^{I-1}$, we have

$$\frac{\partial R^1}{\partial y^I}(\cdot) = \cdots = \frac{\partial R^{I-2}}{\partial y^I}(\cdot) = 0.$$

Furthermore, since

$$R^I(y) = \{1 - G(y^I)\}^{I-1},$$

and

$$
\begin{aligned}
&R^{I-1}(y) \\
&= P(\lambda^{I-2} \geq y^{I-1},\, \lambda^{I-1} < y^{I}) \\
&= (I-1)\{1-G(y^{I-1})\}^{I-2} G(y^{I-1}) \\
&+ \{1-G(y^{I-1})\}^{I-1} - \{1-G(y^{I})\}^{I-1},
\end{aligned}
$$

we have

$$
\frac{\partial R^{I}}{\partial y^{I}}(y) = -(I-1)\{1-G(y^{I})\}^{I-2} g(y^{I}),
$$

and

$$
\frac{\partial R^{I-1}}{\partial y^{I-1}}(y) = -(I-1)(I-2)\{1-G(y^{I-1})\}^{I-3} G(y^{I-1}) g(y^{I-1}).
$$

Suppose now that there exists a maximizer z of w such that $z^{I} > z^{I-1}$. Since $z \in (0,1)^{I}$ by Lemma 9, z satisfies the FOC's:

$$
\begin{aligned}
\frac{\partial w}{\partial y^{I}}(z) = &\frac{\partial R^{I-1}}{\partial y^{I}}(z)\, r^{I-1}(z^{I-1}) \\
&+ \frac{\partial R^{I}}{\partial y^{I}}(z)\, r^{I}(z^{I}) + R^{I}(z)\, (r^{I})'(z^{I}) = 0,
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial w}{\partial y^{I-1}}(z) = &\sum_{n=1}^{I-1} \frac{\partial R^{n}}{\partial y^{I-1}}(z)\, r^{n}(z^{n}) \\
&+ R^{I-1}(z)\, (r^{I-1})'(z^{I-1}) = 0.
\end{aligned}
$$

Noting that $\frac{\partial R^{I-1}}{\partial y^{I}}(z) = -\frac{\partial R^{I-1}}{\partial y^{I}}(z) > 0$, and $\sum_{n=1}^{I-2} \frac{\partial R^{n}}{\partial y^{I-1}}(z) = -\frac{\partial R^{I-1}}{\partial y^{I-1}}(z) > 0$, we obtain

$$
\begin{aligned}
\frac{(r^{I})'(z^{I})}{r^{I}(z^{I})} &= -\frac{1}{R^{I}(z)} \frac{\partial R^{I}}{\partial y^{I}}(z) \left\{1 - \frac{r^{I-1}(z^{I-1})}{r^{I}(z^{I})}\right\}, \\
&= (I-1)\, \frac{g(z^{I})}{1-G(z^{I})} \left\{1 - \frac{r^{I-1}(z^{I-1})}{r^{I}(z^{I})}\right\},
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
\frac{(r^{I-1})'(z^{I-1})}{r^{I-1}(z^{I-1})} &= \frac{1}{R^{I-1}(z)} \left\{-\frac{\partial R^{I-1}}{\partial y^{I-1}}(z) - \sum_{n=1}^{I-2} \frac{\partial R^{n}}{\partial y^{I-1}}(z)\, \frac{r^{n}(z^{n})}{r^{I-1}(z^{I-1})}\right\} \\
&\leq -\frac{1}{R^{I-1}(z)} \frac{\partial R^{I-1}}{\partial y^{I-1}}(z) \left\{1 - \frac{r^{1}(z^{1})}{r^{I-1}(z^{I-1})}\right\}.
\end{aligned}
\tag{25}
$$

Let $\hat{R}^{I-1}(y) = R^{I-1}(\hat{y})$, where $\hat{y}^k = y^k$ for $k \neq I$ and $\hat{y}^I = y^{I-1}$. Then $\hat{R}^{I-1}(y) \leq R^{I-1}(y)$, and

$$-\frac{1}{\hat{R}^{I-1}(z)} \frac{\partial R^{I-1}}{\partial y^{I-1}}(z) = (I-2)\, \frac{g(z^{I-1})}{1-G(z^{I-1})}.$$

Since $z^{I-1} < z^1 = \bar{z}^1$ and $(r^1)'(y^1) > 0$ for $y^1 < \bar{z}^1$, we have

$$\frac{r^1(z^1)}{r^{I-1}(z^{I-1})} > \frac{r^1(z^{I-1})}{r^{I-1}(z^{I-1})} = \frac{v^1(z^{I-1})}{v^{I-1}(z^{I-1})} \geq \prod_{k=2}^{I-1} \underline{\mu}^k.$$

Hence, (25) implies

$$\frac{(r^{I-1})'(z^{I-1})}{r^{I-1}(z^{I-1})} \leq (I-2)\, \frac{g(z^{I-1})}{1-G(z^{I-1})} \left\{ 1 - \prod_{k=2}^{I-1} \underline{\mu}^k \right\}. \tag{26}$$

Since $(r^I)'(z^I) > 0$, $z^{I-1} < z^I$ and $(r^{I-1})'(y) > 0$ whenever $(r^I)'(y) > 0$, we also have

$$\frac{r^{I-1}(z^{I-1})}{r^I(z^I)} < \frac{r^{I-1}(z^I)}{r^I(z^I)} = \frac{v^{I-1}(z^I)}{v^I(z^I)} \leq \bar{\mu}^I.$$

substituting this into (24), we get

$$\frac{(r^I)'(z^I)}{r^I(z^I)} > (I-1)\, \frac{g(z^I)}{1-G(z^I)}\, (1-\bar{\mu}^I). \tag{27}$$

Combining (26) and (27) together, we see that

$$\begin{aligned}
&(I-1)\, \frac{g(z^I)}{1-G(z^I)}\, (1-\bar{\mu}^I) \\
&< \frac{(r^I)'(z^I)}{r^I(z^I)} \leq \frac{(r^I)'(z^{I-1})}{r^I(z^{I-1})} \leq \frac{(r^{I-1})'(z^{I-1})}{r^{I-1}(z^{I-1})} \\
&\leq (I-2)\, \frac{g(z^{I-1})}{1-G(z^{I-1})} \left\{ 1 - \prod_{k=2}^{I-1} \underline{\mu}^k \right\},
\end{aligned}$$

where the inequalities in the second line hold because $\frac{(r^I)'}{r^I}$ is decreasing, $z^I > z^{I-1}$. and $\frac{(r^I)'(s_i)}{r^I(s_i)} \leq \frac{(r^{I-1})'(s_i)}{r^{I-1}(s_i)}$ for any s_i. Furthermore, given the increasing hazard rate, we must have

$$(I-1)\,(1-\bar{\mu}^I) < (I-2) \left\{ 1 - \prod_{k=2}^{I-1} \underline{\mu}^k \right\}.$$

This, however, contradicts Lemma 7(2).

As an induction hypothesis, suppose that $z^I \leq \cdots \leq z^n$. Suppose that $z^n > z^{n-1}$. For any y such that $y^I \leq \cdots \leq y^n$ and $y^n > y^{n-1}$, we have $\frac{\partial R^1}{\partial y^n}(y) = \cdots = \frac{\partial R^{n-1}}{\partial y^n}(y) = 0$, and

$$\begin{aligned} & R^n(y) \\ & = P(\lambda^{n-1} \geq y^n, \lambda^n < y^{n+1}, \dots, \lambda^{I-1} < y^I) \\ & = \binom{I-1}{n-1} \{1 - G(y^n)\}^{n-1} P(\lambda^1_{I-n} < y^{n+1}, \dots, \lambda^{I-n}_{I-n} < y^I), \end{aligned}$$

where λ^k_{I-n} is the kth largest value among $I-n$ types. Hence,

$$\begin{aligned} & \frac{\partial R^n}{\partial y^n}(y) \\ & = -g(y^n)(n-1)\{1-G(y^n)\}^{n-2} \binom{I-1}{n-1} P(\lambda^1_{I-n} < y^{n+1}, \dots, \lambda^{I-n}_{I-n} < y^I), \end{aligned}$$

and

$$\frac{1}{Q^n(y)} \frac{\partial Q^n}{\partial y^n}(y) = -(n-1)\frac{g(y^n)}{1-G(y^n)}.$$

The first-order condition $\frac{\partial w}{\partial y^n}(z) = 0$ then yields

$$\begin{aligned} \frac{(r^n)'(z^n)}{r^n(z^n)} &= -\frac{1}{Q^n}\frac{\partial Q^n}{\partial y^n}\left\{1 - \frac{r^{n-1}(z^{n-1})}{r^n(z^n)}\right\} \\ &= (n-1)\frac{g(z^n)}{1-G(z^n)}\left\{1 - \frac{r^{n-1}(z^{n-1})}{r^n(z^n)}\right\}. \end{aligned} \tag{28}$$

On the other hand, suppose $y^{n+1} \leq y^{n-1} < y^n$. Other cases can be treated in a similar manner.

$$\begin{aligned} & R^{n-1}(y) \\ & = P(\lambda^{n-2} \geq y^{n-1}, \lambda^{n-1} < y^n, \lambda^n < y^{n+1}, \dots, \lambda^{I-1} < y^I) \\ & = \binom{I-1}{n-2} \{1 - G(y^{n-1})\}^{n-2} \\ & \times \Big[(I-n+1)\{G(y^{n-1}) - G(y^{n+1})\} P(\lambda^1_{I-n} < y^{n+1}, \dots, \lambda^{I-n}_{I-n} < y^I) \\ & + P(\lambda^1_{I-n+1} < y^{n+1}, \dots, \lambda^{I-n+1}_{I-n+1} < y^I)\Big] \\ & + \binom{I-1}{n-1} \Big[\{1-G(y^{n-1})\}^{n-1} - \{1-G(y^n)\}^{n-1}\Big] P(\lambda^1_{I-n} < y^{n+1}, \dots, \lambda^{I-n}_{I-n} < y^I). \end{aligned}$$

Differentiating R^{n-1} with respect to y^{n-1}, we obtain

$$\begin{aligned}
&\frac{\partial R^{n-1}}{\partial y^{n-1}}(y) \\
&= -g(y^{n-1}) \binom{I-1}{n-2} (n-2)\{1-G(y^{n-1})\}^{n-3} \\
&\times \Big[(I-n+1)\{G(y^{n-1}) - G(y^{n+1})\} P(\lambda^1_{I-n} < y^{n+1}, \ldots, \lambda^{I-n}_{I-n} < y^I) \\
&+ P(\lambda^1_{I-n+1} < y^{n+1}, \ldots, \lambda^{I-n+1}_{I-n+1} < y^I)\Big].
\end{aligned}$$

Let $\hat{R}^{n-1}(y) = R^{n-1}(\hat{y})$, where $\hat{y}^k = y^k$ for $k \neq n$ and $\hat{y}^n = y^{n-1}$. We have $\hat{R}^{n-1}(y) \leq R^{n-1}(y)$ and can also verify that

$$\frac{1}{\hat{R}^{n-1}(y)} \frac{\partial R^{n-1}}{\partial y^{n-1}}(y) = -(n-2) \frac{g(y^{n-1})}{1-G(y^{n-1})}.$$

The first-order condition $\frac{\partial w}{\partial y^{n-1}}(z) = 0$ then yields

$$\begin{aligned}
&\frac{(r^{n-1})'(z^{n-1})}{r^{n-1}(z^{n-1})} \\
&= -\frac{1}{R^{n-1}(z)} \left[\sum_{k=1}^{n-1} \frac{\partial R^k}{\partial y^{n-1}}(z) \frac{r^k(z^k)}{r^{n-1}(z^{n-1})}\right] \\
&\leq (n-2) \frac{g(y^n)}{1-G(y^n)} \left(1 - \frac{r^1(z^1)}{r^{n-1}(z^{n-1})}\right) \\
&< (n-2) \frac{g(y^n)}{1-G(y^n)} \left(1 - \frac{r^1(z^{n-1})}{r^{n-1}(z^{n-1})}\right) \\
&\leq (n-2) \frac{g(y^n)}{1-G(y^n)} \left(1 - \prod_{k=2}^{n-1} \underline{\mu}^k\right).
\end{aligned} \tag{29}$$

On the other hand,

$$\begin{aligned}
\frac{(r^n)'(z^n)}{r^n(z^n)} &= (n-1) \frac{g(z^n)}{1-G(z^n)} \left\{1 - \frac{r^{n-1}(z^{n-1})}{r^n(z^n)}\right\} \\
&> (n-1) \frac{g(z^n)}{1-G(z^n)} \left\{1 - \frac{r^{n-1}(z^n)}{r^n(z^n)}\right\} \\
&> (n-1) \frac{g(z^n)}{1-G(z^n)} \left\{1 - \bar{\mu}^n\right\}.
\end{aligned} \tag{30}$$

Just as in Step 1, we can combine (29) and (30) to yield a contradiction to Lemma 7(2), which is equivalent to

$$(n-1)\ \{1-\bar{\mu}^n\} \geq (n-2) \left(1 - \prod_{k=2}^{n-1} \underline{\mu}^k\right).$$

This advances the induction step and completes the proof.

Proof of Proposition 2 Take any $J \subset I$, $s = (s_J, s_{-J})$ and $\hat{s}_J$. Denote $\hat{s} = (\hat{s}_J, s_{-J})$ and $k = |f(s)|$. If $|f(\hat{s})| = m > k$, then $|\{i : s_i \geq z^k\}| = k < m$ and $|\{i : \hat{s}_i \geq z^m\}| = m$ by the definition of a monotone scheme. Hence, there exists at least one buyer $i \in J$ for whom $s_i < z^m$, $\hat{s}_i \geq z^m$ and $f_i(\hat{s}) = 1$. It follows that for this i,

$$v_i(f(\hat{s}), s_i) - t_i(f(\hat{s})) = v^m(s_i) - t^m < 0 \leq v_i(f(s), s_i) - t_i(f(s)),$$

suggesting that $\hat{s}_J$ is not a profitable deviation of J at s. If $|f(\hat{s})| = m < k$, take any $i \in J$ for whom $f_i(\hat{s}) = 1$. If there exists no such $j \in J$, then $\hat{s}$ is not a profitable deviation for J. Since $z^m \geq z^k$, we have $v^k(z^m) - v^m(z^m) \geq v^k(z^k) - v^m(z^m)$. Furthermore, since $(v^k)' \geq (v^m)'$, $v^k(s_i) - v^m(s_i) \geq v^k(z^m) - v^m(z^m)$ for any $s_i \geq z^m$. It follows that $v^k(s_i) - v^m(s_i) \geq v^k(z^k) - v^m(z^m)$ for any $s_i \geq z^m$. In other words, if $f_i(\hat{s}) = 1$ and $s_i \geq z^m$, then

$$\begin{aligned} v_i(f(\hat{s}), s_i) - t_i(f(\hat{s})) &= v^m(s_i) - t^m \\ &= v^m(s_i) - v^m(z^m) \\ &\leq v^k(s_i) - v^k(z^k) \\ &= v_i(f(s), s_i) - t_i(f(s)). \end{aligned}$$

This implies that $\hat{s}$ is not a profitable deviation for J at s.

Proof of Lemma 4 Suppose that (f, y) is such that $M(y) \cap K(f) \neq \emptyset$, and take $m \in M(y) \cap K(f)$ so that $|f(\hat{s})| = m$ for some $\hat{s}$ and $y^m < y^n$ for some $n > m$. By symmetry, we can take $\hat{s}$ such that $f(\hat{s}) = (\underbrace{1, \ldots, 1}_{m}, 0, \ldots, 0)$. Take s such that $y^m < s_1 = \cdots = s_n < y^n$ and $s_{n+1} = \cdots = s_I = 0$. Symmetry and ex post IR then imply that $f(s)$ equals either $(\underbrace{0, \ldots, 0}_{n}, \underbrace{1, \ldots, 1}_{I-n})$, $(0, \ldots, 0)$, or $(1, \ldots, 1)$. When $f(s) = (0, \ldots, 0, 1, \ldots, 1)$, $t^{I-n} = 0$ should hold by ex post IR, and when $f(s) = (1, \ldots, 1)$, then $t^I = 0$ should hold by ex post IR.

If $f(s) = (0, \ldots, 0)$. then $\hat{s}$ is a profitable deviation for the coalition $J = I$ at s: For $i = 1, \ldots, m$,

$$v_i(f(\hat{s}), s_i) - t_i(f(\hat{s})) = v^m(s_i) - t^m > 0 = v_i(f(s), s_i) - t_i(s),$$

where the inequality follows from the fact that $s_i > y^m$. Furthermore, for $i = m+1, \dots, I$,

$$v_i(f(\hat{s}), s_i) - t_i(f(\hat{s})) = 0 = v_i(f(s), s_i) - t_i(s).$$

If $f(s) = (\underbrace{0, \dots, 0}_{n}, 1, \dots, 1)$, then $t^{I-n} = 0$ as noted above and $\hat{s}$ is a profitable deviation for the coalition $J = I$ at s: For $i = 1, \dots, m$,

$$v_i(f(\hat{s}), s_i) - t_i(f(\hat{s})) = v^m(s_i) - t^m > 0 = v_i(f(s), s_i) - t_i(f(s)),$$

for $i = m+1, \dots, n$,

$$v_i(f(\hat{s}), s_i) - t_i(f(\hat{s})) = 0 = v_i(f(s), s_i) - t_i(f(s)),$$

and for $i = n+1, \dots, I$,

$$v_i(f(\hat{s}), s_i) - t_i(f(\hat{s})) = 0 = v^{I-n}(0) - t^{I-n} = v_i(f(s), s_i) - t_i(f(s)).$$

If $f(s) = (1, \dots, 1)$, then $y^I = 0$ as noted above and s is a profitable deviation for the coalition $J = I$ at $\hat{s}$: For $i = 1, \dots, m$,

$$v_i(f(s), \hat{s}_i) - t_i(f(s)) = v^I(\hat{s}_i) - t^I > v^m(\hat{s}_i) - t^m = v_i(f(\hat{s}), \hat{s}_i) - t_i(f(\hat{s})),$$

and for $i = m+1, \dots, I$,

$$v_i(f(s), \hat{s}_i) - t_i(f(s)) = v^I(\hat{s}_i) - t^I > 0 = v_i(f(\hat{s}), \hat{s}_i) - t_i(f(\hat{s})).$$

Therefore, (f, t) is not coalitionally strategy-proof.

Proof of Lemma 5 Let (f, y) be such that $K(f) = K$. Fix $k \in K$ and s_{-i} such that $\lambda^{k-1} \geq y^k$ and $\max_{\substack{\ell \in K \\ \ell > k}} (\lambda^{\ell-1} - y^\ell) < 0$. By our choice of s_{-i}, $|f(s_i, s_{-i})| \leq k$ for any s_i by ex post IR. Moreover, if $s_i < y^k$, then $s_i < y^m$ for any $m < k$ so that i is not assigned the good: $f_i(s_i, s_{-i}) = 0$. In what follows, we show that $|f(s_i, s_{-i})| = k$ whenever $s_i > y^k$. If this holds, then

$$\begin{aligned} E_{s_i}[t_i(f(s_i, s_{-i})) \mid s_{-i}] &= P(s_i < y^k)\, E_{s_i}[t_i(f(s_i, s_{-i})) \mid s_i < y^k, s_{-i}] \\ &\quad + P(s_i > y^k)\, E_{s_i}[t_i(f(s_i, s_{-i})) \mid s_i > y^k, s_{-i}] \\ &= P(s_i > y^k)\, t^k \\ &= r^k(y^k). \end{aligned}$$

This in turn implies that

$$\begin{aligned}
R(f,t) &= E_s[t_i(f(s))] \\
&= E_{s_{-i}}\left[E_{s_i}[t_i(f(s_i, s_{-i})) \mid s_{-i}]\right] \\
&= \sum_{k\in K} r^k(y^k)\, P(\lambda^{k-1} \geq y^k,\ \max_{\substack{\ell\in K \\ \ell>k}} (\lambda^{\ell-1} - y^\ell) < 0) \\
&= w(K, y).
\end{aligned}$$

Suppose that $s_i > y^k$ and denote $s = (s_i, s_{-i})$. We will derive a contradiction when $m = |f(s)| < k$. Let $J \subset I$ be such that $i \in J$, $|J| = k$, and

$$\{j : f_j(s) = 1\} \subset J \subset \{j : s_j \geq y^k\}.$$

Such a set J exists since $f_j(s) = 1$ implies that $s_j \geq y^m$ by ex post IR and $y^m \geq y^k$ by Lemma 4. Since $k \in K = K(f)$, take $\hat{s}$ such that $|f(\hat{s})| = k$ and $\{j : f_j(\hat{s}) = 1\} = J$. Such a type profile $\hat{s}$ exists by symmetry. When $s_j \geq y^k$, note that

$$v^k(s_j) - v^m(s_j) \geq v^k(y^k) - v^m(y^k) \geq v^k(y^k) - v^m(y^m),$$

where the first inequality follows from $(v^k)' > (v^m)'$, and the second from $y^k \leq y^m$. We will show that $\hat{s}$ is a profitable deviation for I at s:
For $j \in J \cap \{j : f_j(s) = 1\}$,

$$\begin{aligned}
v_j(f(\hat{s}), s_j) - t_j(f(\hat{s})) &= v^k(s_j) - v^k(y^k) \\
&\geq v^m(s_j) - v^m(y^m) \qquad (31) \\
&= v_j(f(s), s_j) - t_j(f(s)).
\end{aligned}$$

For $j \in J \cap \{j : f_j(s) = 0\}$,

$$v_j(f(\hat{s}), s_j) - t_j(f(\hat{s})) = v^k(s_j) - t^k \geq 0 = v_j(f(s), s_j) - t_j(f(s)). \qquad (32)$$

For $j \in I \setminus J$,

$$v_j(f(\hat{s}), s_j) - t_j(f(\hat{s})) = 0 = v_j(f(s), s_j) - t_j(f(s)).$$

Since $s_i > y^k$, if $f_i(s) = 1$, then (31) holds with strict inequality for $j = i$, and if $f_i(s) = 0$, then (32) holds with strict inequality for $j = i$. Hence, (f, t) is not strategy-proof.

Proof of Theorem 4 Take any coalitionally ex post implementable scheme (f, y), and denote $K = K(f)$. We will show that when $K \neq N$, $w(K, y) \leq w(N, \hat{y})$ for some $\hat{y}$ such that $\hat{y}^I \leq \cdots \leq \hat{y}^1$. Since $w(N, \cdot)$ is continuous over the compact set $\{y : y^I \leq \cdots \leq y^1\}$, it achieves a maximum at some $z = (z^1, \ldots, z^I)$ in this set. By Proposition 2, the monotone scheme with marginal types z raises the expected revenue $w(N, z)$ and is optimal.

When $y^k = 1$ for some $k \in K$, then $w(K, y) = w(K \setminus \{k\}, y)$ so that we may restrict attention to the case where $\max_{k \in K} y^k < 1$. Suppose that $K \neq N$ and fix any y such that $K \cap M(y) = \emptyset$. We have

$$w(K, y) = \sum_{\ell \in K} P(\lambda^{\ell-1} \geq y^\ell, \max_{\substack{m \in K \\ m > \ell}} (\lambda^{m-1} - y^m) < 0)\, r^\ell(y^\ell).$$

Let $n = \min N \setminus K$ and $\hat{K} = K \cup \{n\}$. If $n = 1$, then let $\hat{y}$ be such that $\hat{y}^1 = \max_{k \in K} y^k$ and $\hat{y}^k = y^k$ for $k > 1$. Then $\hat{K} \cap M(\hat{y}) = \emptyset$ and $w(\hat{K}, \hat{y})$ is given by

$$\begin{aligned}
w(\hat{K}, \hat{y}) &= \sum_{k \in \hat{K}} P(\lambda^{k-1} \geq y^k, \max_{\substack{\ell \in \hat{K} \\ \ell > k}} (\lambda^{\ell-1} - y^\ell) < 0)\, r^k(y^k) \\
&= P(\max_{\ell \in K} (\lambda^{\ell-1} - y^\ell) < 0)\, r^1(y^1) \\
&+ \sum_{k \in K} P(\lambda^{k-1} \geq y^k, \max_{\substack{\ell \in K \\ \ell > k}} (\lambda^{\ell-1} - y^\ell) < 0)\, r^k(y^k) \\
&\geq w(K, y).
\end{aligned}$$

If $n > 1$, then $n - 1 \in K$ and let $\hat{y}$ be such that

$$\hat{y}^k = \begin{cases} y^k & \text{if } k \neq n, \\ y^{n-1} & \text{if } k = n. \end{cases}$$

Since $\hat{K} \cap M(\hat{y}) = \emptyset$, $w(\hat{K}, \hat{y})$ is given by

$$\begin{aligned}
w(\hat{K}, \hat{y}) &= \sum_{\ell \in \hat{K}} P\Big(\lambda^{\ell-1} \geq \hat{y}^\ell, \max_{\substack{m \in \hat{K} \\ m > \ell}} (\lambda^{m-1} - \hat{y}^m) < 0\Big)\, r^\ell(\hat{y}^\ell) \\
&= \sum_{\substack{\ell \in K \\ \ell < n}} P\Big(\lambda^{\ell-1} \geq y^\ell, \lambda^{n-1} < y^{n-1}, \max_{\substack{m \in K \\ m > \ell}} (\lambda^{m-1} - y^m) < 0\Big)\, r^\ell(y^\ell) \\
&+ P\Big(\lambda^{n-1} \geq y^{n-1}, \max_{\substack{m \in K \\ m > n}} (\lambda^{m-1} - y^m) < 0\Big)\, r^n(y^{n-1}) \\
&+ \sum_{\substack{\ell \in K \\ \ell > n}} P\Big(\lambda^{\ell-1} \geq y^\ell, \max_{\substack{m \in K \\ m > \ell}} (\lambda^{m-1} - y^m) < 0\Big)\, r^\ell(y^\ell).
\end{aligned}$$

Noting that $\lambda^{n-2} < y^{n-1}$ implies $\lambda^{n-1} < y^{n-1}$, we can decompose the first line of the right-hand side above to rewrite $w(\hat{K}, \hat{y})$ as

$$\begin{aligned} w(\hat{K}, \hat{y}) = & \sum_{\substack{\ell \in K \\ \ell < n-1}} P\Big(\lambda^{\ell-1} \geq y^{\ell},\, \max_{\substack{m \in K \\ m > \ell}} (\lambda^{m-1} - y^m) < 0\Big)\, r^{\ell}(y^{\ell}) \\ & + P\Big(\lambda^{n-2} \geq y^{n-1},\, \lambda^{n-1} < y^{n-1},\, \max_{\substack{m \in K \\ m > n-1}} (\lambda^{m-1} - y^m) < 0\Big)\, r^{n-1}(y^{n-1}) \\ & + P\Big(\lambda^{n-1} \geq y^{n-1},\, \max_{\substack{m \in K \\ m > n}} (\lambda^{m-1} - y^m) < 0\Big)\, r^{n}(y^{n-1}) \\ & + \sum_{\substack{\ell \in K \\ \ell > n}} P\Big(\lambda^{\ell-1} \geq y^{\ell},\, \max_{\substack{m \in K \\ m > \ell}} (\lambda^{m-1} - y^m) < 0\Big)\, r^{\ell}(y^{\ell}). \end{aligned}$$

Summing the probabilities in the second and third lines above yields

$$\begin{aligned} & P\Big(\lambda^{n-2} \geq y^{n-1},\, \lambda^{n-1} < y^{n-1},\, \max_{\substack{m \in K \\ m > n-1}} (\lambda^{m-1} - y^m) < 0\Big) \\ & + P\Big(\lambda^{n-1} \geq y^{n-1},\, \max_{\substack{m \in K \\ m > n}} (\lambda^{m-1} - y^m) < 0\Big) \\ & = P\Big(\lambda^{n-2} \geq y^{n-1},\, \max_{\substack{m \in K \\ m > n-1}} (\lambda^{m-1} - y^m) < 0\Big). \end{aligned}$$

Using this and the fact that $r^{n}(y^{n-1}) \geq r^{n-1}(y^{n-1})$, we obtain $w(\hat{K}, \hat{y}) \geq w(K, y)$. Iteration of this process shows that $w(N, \hat{y}) \geq w(K, y)$ for some $\hat{y}$ such that $\hat{y}^1 \leq \cdots \leq \hat{y}^I$.

References

[1] Masaki Aoyagi (2010), “Optimal sales schemes against interdependent buyers,” American Economic Journal: Microeconomics 2(1), 150-182.

[2] Mark Armstrong (2006), “Competition in two-sided markets,” RAND Journal of Economics 37(3), 668-91.

[3] Parimal Kanti Bag, and Eyal Winter (1999), “Simple Subscription Mechanisms for Excludable Public Goods,” Journal of Economic Theory 87, 72-94.

[4] Shai Bernstein and Eyal Winter (2010), “Contracting with Heterogeneous Externalities,” working paper.

[5] Luis M. B. Cabral, David J. Salant, and Glenn A. Woroch (1999), "Monopoly pricing with network externalities," International Journal of Industrial Organization, 17, 199-214.

[6] Rajat Deb, and Laura Razzolini (1999a), "Voluntary cost sharing for an excludable public project," Mathematical Social Science, 37, 123-138.

[7] Rajat Deb, and Laura Razzolini (1999b), "Auction-like mechanisms for pricing excludable public goods," Journal of Economic Theory, 88, 340-368.

[8] Philip H. Dybvig, and Chester S. Spatt (1983), "Adoption externalities as public goods," Journal of Public Economics, 20, 231-247.

[9] Douglas Gale (1995),"Dynamic coordination games," Economic Theory, 5, 1-18.

[10] Douglas Gale (2001), "Monotone games with positive spillovers," Games and Economic Behavior, 37, 295-320.

[11] Michael L. Katz, and Carl Shapiro (1985), "Network externalities, competition, and compatibility," American Economic Review, 75, 424-440.

[12] Michael L. Katz, and Carl Shapiro (1986), "Technology adoption in the presence of network externalities," Journal of Political Economy, 94, 822-841.

[13] Herve Moulin (1980), "On strategy-proofness and single peakedness," Public Choice, 35(4), 437-55.

[14] Herve Moulin (1994), "Serial cost sharing of excludable public goods," Review of Economic Studies, 61, 305-325.

[15] Jack Ochs, and In-UK Park (2010), "Overcoming the coordination problem: dynamic formation of networks," Journal of Economic Theory, 145, 689-720.

[16] In-Uck Park (2004), "A simple inducement scheme to overcome adoption externalities," Contributions to Theoretical Economics, 4(1), Article 3.

[17] Jeffrey H. Rohlfs (1974), "A theory of interdependent demand for a communications service," Bell Journal of Economics 5 (1), 16-37.

[18] Ilya Segal (2003), "Coordination and discrimination in contracting with externalities: Divide and conquer?" Journal of Economic Theory 113, 147-327.

[19] Tadashi Sekiguchi (2009), "Pricing of durable network goods under dynamic coordination failure," working paper.

[20] Ran Shao, and Lin Zhou (2008), "Optimal allocation of an indivisible good," working paper.

[21] Tatsuhiro Shichijo, and Yuji Nakayama (2009) "A two-step subsidy scheme to overcome network externalities in a dynamic game," The B.E. Journal of Theoretical Economics, 9(1) (Topics), Article 4.

[22] Arun Sundararajan (2007) "Local network effects and complex network structure," The B.E. Journal of Theoretical Economics, 7(1) (Contributions), Article 46.