OPTIMAL PROVISION OF
MULTIPLE EXCLUDABLE PUBLIC GOODS

By

Hanming Fang and Peter Norman

October 2003
Revised April 2006

COWLES FOUNDATION DISCUSSION PAPER NO. 1441R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Optimal Provision of Multiple Excludable Public Goods*

Hanming Fang† Peter Norman‡

First Version: October 2002
This Version: April 2006

Abstract

This paper studies the optimal provision mechanism for multiple excludable public goods when agents' valuations are private information. For a parametric class of problems with binary valuations, we characterize the optimal mechanism, and show that it involves bundling. Bundling alleviates the free riding problem in large economies in two ways: first, it can increase the asymptotic provision probability of socially efficient public goods from zero to one; second, it decreases the extent of use exclusions.

Keywords: Public Goods Provision; Bundling; Exclusion

JEL Classification Number: H41.

*This paper was previously titled "An Efficiency Rationale for Bundling of Public Goods". We thank Mark Armstrong, Ted Bergstrom, Martin Hellwig, Larry Samuelson, Stephen Morris and seminar participants at many seminars and conferences for comments and helpful discussions. The usual disclaimer applies.

†Department of Economics, Yale University, P.O. Box 208264, New Haven, CT 06520-8264. Email: `hanming.fang@yale.edu`

‡Department of Economics, University of British Columbia, 909-1873 East Mall, Vancouver, BC V6T1Z1. Email: `penorman@interchange.ubc.ca`

1 Introduction

This paper studies the optimal provision mechanism for multiple excludable public goods, and shows that bundling is an important feature of the constrained efficient mechanism.

"Bundling" refers to the practice to sell several goods together as a package deal, as opposed to providing each good separately. This is a common selling strategy, and there is a quite large literature on the topic, which almost exclusively focuses on private goods.[1] If the technology exhibits constant or decreasing returns to scale, bundling of private goods can be rationalized as a scheme to improve the extraction of surplus from consumers by a seller with market power. However, regardless of whether preferences are private information or not, the normative recommendation in such an environment, simple marginal cost pricing, does not involve bundling.

To motivate the importance of a better understanding of bundling of *non-rival* goods, we note that many goods that are provided in bundles are close to fully non-rival. A striking example is access to electronic libraries, for which the typical contractual arrangement is a site license that allows access to every issue of every journal in the electronic library. Another obvious example is cable TV. Technologically, it would be feasible to allow customers to choose whatever channels they are willing to pay for without constraints. In practice, the basic pricing scheme consists of a limited number of available packages. While some programming is available on a per channel or pay-per-view basis, the bundled channels are simply not available in any other way than through their respective bundles. Other examples include computer software and digital music files. All these examples have in common that the pros and cons of bundling *for the consumer* have been frequently debated by the media, legal scholars, and in the courtroom. Still, there are virtually no attempts in the academic economics literature to build a normative benchmark that explicitly considers the public nature of these goods.[2]

We consider a model with m excludable public goods: the goods are fully non-rival, but consumers can be excluded from usage. Each consumer is characterized by a valuation for each good, and the willingness to pay for a subset of goods is the sum of the individual good valuations. The cost of provision for each good is independent of which other goods are provided. Under these separability assumptions, the first best benchmark is to provide good j if and only if the sum of valuations for good j exceeds the provision cost and exclude no consumer from usage. Under perfect information there is thus no role for either bundling or use exclusions.

In this paper we assume that preferences are private information to the individuals. Consumers must therefore be given incentives to truthfully reveal their willingness to pay. We also assume that consumers may decide not to participate, and that the provision mechanism must be self-financing. Finally, we assume that the preference parameters are stochastically independent across individuals. Together, these assumptions make it impossible to implement the (non-bundling)

[1] An exception is Bakos and Brynjolfsson [6].

[2] However, see Bergstrom and Bergstrom [7] who discuss site licenses for academic journals.

perfect information social optimum.[3]

The first part of the paper considers a relatively general setup, and characterizes the form of optimal provision mechanisms in symmetric environments. We then exploit these results for a special case where we obtain an exact characterization of the constrained efficient mechanism. This special case is when there are two public goods, valuations for each good are binary, and the goods are symmetric both with respects to costs of provision and joint valuation distributions. While this is obviously a very special case, the results are suggestive, and the methodology may be useful for more general (symmetric) multidimensional screening problems.

There is an element of bundling in the constrained efficient mechanism whenever valuations are not too positively correlated. The best intuition for this probably comes from considering results in McAfee et al [18] together with the more recent literature on efficient provision of (a single-dimensional) excludable public good. We know from McAfee et al [18] that introducing the bundling instrument increases the profits for a monopolist that is restricted to fixed-price mechanisms. By results in Hellwig [14] and Norman [21] we also know that, in the case with a single good, the constrained welfare problem has a Lagrangian characterization. This problem may be interpreted as maximizing a weighted average of social welfare and profits, where the relative weights come from the Lagrange multiplier on a "zero profit constraint." Given these links between constrained efficiency and a standard monopoly problem it seems highly plausible that the logic in McAfee et al [18] should carry over to our problem.

Concretely, bundling works as follows in the optimal mechanism. All agents get access to any good for which he or she has a high valuation. If valuations of the two goods are not too positively correlated, a "mixed type" is always more likely to get access to his or her low-valuation good than is an agent with low valuations for both goods. In some cases, this differential treatment leads to a drastic improvement compared to the best that can be achieved without bundling. For many parametrizations, the probability of provision tends to zero if bundling is not used, whereas bundling makes it possible to provide with probability one. Moreover, even when the goods may be provided without bundling, the proportion of agents that get access to the goods is higher under the optimal bundle mechanism.

It is important to note that, while the existing literature on bundling (of private goods) focuses on how bundling relaxes the informational constraints and improves sellers' *revenue,* we derive a *constrained efficient* mechanism that involves bundling in the public good setting. In the models used in the existing literature, the profit maximizing bundling mechanism is dominated by marginal cost pricing in terms of social efficiency, and could be trivially implemented.

The remainder of the paper is structured as follows. Section 2 presents the model and some

[3] All these restrictions are essential. Removing either the voluntary participation or the self-financing constraint makes it possible to construct pivot-mechanisms that implement the first-best. If we allow correlation in valuations, a version of the analysis in Cremer and McLean [10] can be used to implement the efficient outcome.

characterization results. Section 3 introduces the special case when valuations are binary and demonstrates by example that a (pure) bundling mechanism may improve efficiency. Section 4 characterizes the optimal mechanism for this special case, and compares our characterization with existing results in the literature, and Section 5 concludes. Proofs are collected in the Appendices.

2 The Model

This section lays out a fairly general model (Section 2.1). The set of randomized direct mechanisms is represented in a somewhat nonstandard, but useful, way (Section 2.2), before setting up the mechanism design problem (Section 2.3). We then gradually show, sometimes with additional restrictions on the environment, that it is without loss of generality to consider a smaller and more tractable class of simple, anonymous and symmetric mechanisms (Sections 2.4 and 2.5). The main results of this section are Propositions 1 and 2, which are used in later Sections to reduce the dimensionality of the design problem.

2.1 The Environment

There are m *excludable* public goods, labeled by $j \in \mathcal{J} = \{1, ..., m\}$ and n consumers, indexed by $i \in \mathcal{I} = \{1, ..., n\}$. Each public good is indivisible, and the cost of providing good j, denoted $C^j(n)$, is independent of which of the other goods are provided. Since n is the size of the economy and *not* the number of users, all goods are fully non-rival. The rationale for indexing cost by n is to be able to analyze large economies without making the public goods a free lunch in the limit. We therefore allow for the existence of $c^j > 0$ such that $\lim_{n\to\infty} C^j(n)/n = c^j > 0$. There is no need to give this assumption any economic interpretation. It is best viewed as a way to ensure that the provision problem remains "significant" with a large number of agents.

Consumer i is described by a vector $\theta_i = \left(\theta_i^1, ..., \theta_i^m\right) \in \Theta \subset R^m$, where θ_i^j is interpreted as i's valuation for good j. Agent i has preferences represented by the utility function,

$$\sum_{j\in\mathcal{J}} \mathtt{I}_i^j \theta_i^j - t_i, \tag{1}$$

where $\mathtt{I}_i^j$ is a dummy variable taking value 1 when i consumes good j and 0 otherwise, and t_i is the quantity of the numeraire good transferred from i to the mechanism designer. Preferences over lotteries are of expected utility form. One could imagine more general utility functions than (1), but the linear formulation (also used by Adams and Yellen [1], McAfee et al [18], and Manelli and Vincent [17]) has the advantage of ruling out bundling due to complementarities in preferences.

The preference vector θ_i is private information to the agent, and preferences are independently and identically distributed *across agents*. We denote by F the joint cumulative distribution over θ_i. Valuations *across goods* may be correlated for the individual agent. For brevity of notation, we let $\theta \equiv (\theta_1, ..., \theta_n) \in \Theta^n$, which will be referred to as a *type profile*. In the usual fashion, we let $\theta_{-i} = (\theta_1, .., \theta_{i-1}, \theta_{i+1}, ...\theta_n)$ and, with some abuse of notation, we write $\mathbf{F}(\theta) \equiv \Pi_{i\in\mathcal{I}} F(\theta_i)$ and $\mathbf{F}(\theta_{-i}) \equiv \Pi_{k\in\mathcal{I}\setminus i} F(\theta_k)$ as the joint distribution of θ and θ_{-i} respectively.

2.2 Randomized Direct Mechanisms

In general, the outcome of any mechanism must determine: (1). Which goods, if any, should be provided; (2). Who are to be given access to the goods that are provided; and (3). How to share the costs. The set of feasible *pure* outcomes is thus

$$A = \underbrace{\{0,1\}^m}_{\substack{\text{provision/no provision} \\ \text{for each goods } j}} \times \underbrace{\{0,1\}^{m\times n}}_{\substack{\text{inclusion/no inclusion} \\ \text{for each agent } i \text{ and good } j}} \times \underbrace{\mathrm{R}^n}_{\text{“taxes”}}. \tag{2}$$

By the revelation principle, we restrict attention to direct mechanisms for which truth-telling is a Bayesian Nash equilibrium. A pure direct mechanism is a map from Θ^n to A. We represent a randomized mechanism in analogy with the representation of mixed strategies in Aumann [5]. That is, let $\Xi \equiv [0,1]$, and think of $\vartheta \in \Xi$ as the outcome of a fictitious lottery, where, without loss of generality, ϑ is uniformly distributed and independent of θ. A *random direct mechanism* is then a measurable mapping $\mathcal{G} : \Theta^n \times \Xi \to A$. A conceptual advantage of this representation is that it allows for a useful decomposition.That is, we may write $\mathcal{G}$ as a $(2m+1)$-tuple, $\mathcal{G} = \left(\left\{\zeta^j\right\}_{j\in\mathcal{J}}, \left\{\omega^j\right\}_{j\in\mathcal{J}}, \tau\right)$ where,

$$\begin{aligned} \text{Provision Rule:} \quad & \zeta^j && : \Theta^n \times \Xi \to \{0,1\} \\ \text{Inclusion Rule:} \quad & \omega^j && : \Theta^n \times \Xi \to \{0,1\}^n \\ \text{Cost-sharing Rule:} \quad & \tau && : \Theta^n \to \mathrm{R}^n. \end{aligned} \tag{3}$$

We refer to ζ^j as the *provision rule* for good j, and interpret $\mathrm{E}_\Xi \zeta^j(\theta, \vartheta)$ as the probability of provision given announcements θ. The rule $\omega^j = \left(\omega_1^j, ..., \omega_n^j\right)$ is the *inclusion rule* for good j, and $\mathrm{E}_\Xi \omega_i^j(\theta, \vartheta)$ is interpreted as the probability that agent i gets access to good j when announcements are θ, conditional on good j being provided. Finally, $\tau = (\tau_1, ..., \tau_n)$ is the *cost-sharing rule,* where $\tau_i(\theta)$ is the transfer from agent i to the mechanism designer given announced valuations θ. In principle, transfers could also be random, but the pure cost-sharing rule in (3) is without loss of generality due to risk neutrality.

2.3 The Design Problem

Let E_{-i} denote the expectation operator with respect to (θ_{-i}, ϑ). A mechanism is *incentive compatible* if truth-telling is a Bayesian Nash equilibrium in the revelation game induced by $\mathcal{G}$,

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)\theta_i^j - \tau_i(\theta)\right] \geq \mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j(\widehat{\theta}_i,\theta_{-i},\vartheta)\omega_i^j(\widehat{\theta}_i,\theta_{-i},\vartheta)\theta_i^j - \tau_i(\widehat{\theta}_i,\theta_{-i})\right],$$
$$\forall i \in \mathcal{I}, \theta \in \Theta^n, \widehat{\theta}_i \in \Theta. \quad \text{(IC)}$$

We also require that the project be self-financing. For simplicity, this is imposed as an *ex ante balanced-budget constraint:*[4]

$$\mathrm{E}\left(\sum_{i\in\mathcal{I}} \tau_i(\theta) - \sum_{j\in\mathcal{J}} \zeta^j(\theta,\vartheta)\, C^j(n)\right) \geq 0. \quad \text{(BB)}$$

Finally, we require that a voluntary participation, or *individual rationality,* condition is satisfied. Agents are assumed to know their own type, but not the realized types of the other agents, when deciding on whether to participate. Individual rationality is thus imposed at the interim stage as,

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)\theta_i^j - \tau_i(\theta)\right] \geq 0, \qquad \forall i \in \mathcal{I}, \theta_i \in \Theta. \quad \text{(IR)}$$

A mechanism is *incentive feasible* if it satisfies (IC), (BB) and (IR). Utility is transferable, implying that constrained efficient allocations may be characterized by solving a utilitarian planning problem, where a fictitious social planner seeks to maximize total surplus in the economy, subject to the constraints (IC), (BB) and (IR). A mechanism is thus *constrained efficient* if it maximizes

$$\sum_{j\in\mathcal{J}} \mathrm{E}\zeta^j(\theta,\vartheta)\left[\sum_{i\in\mathcal{I}} \omega_i^j(\theta,\vartheta)\theta_i^j - C^j(n)\right], \quad (4)$$

over all incentive feasible mechanisms.[5]

It is ex post efficient to provide good j if and only if $\sum_{i\in\mathcal{I}} \theta_i^j \geq C^j(n)$, and to never exclude any agent from usage, which is the same rule as the first best rule for a single public good. This is implementable if and only if a non-excludable public good can be efficiently provided under (IC), (BB) and (IR). But, this is only possible in trivial cases (Mailath and Postlewaite [16]). Our setup is thus a second best problem.

[4] The ex ante constraint (BB) is literally relevant only when the designer can access fair insurance market against budget deficits. However, adapting standard arguments (see Mailath and Postlewaite [16] and Cramton et al [9]), one can show that any allocation implementable with transfers satisfying (BB) is also implementable with a transfer rule that satisfies the ex post balanced-budget constraint (i.e. feasibility for every realization of θ). The idea is simply that, since agents are risk-neutral, the insurance against budget deficits can be provided by the agents.

[5] All these constraints are noncontroversial if the design problem is interpreted as a private bargaining agreement. If the goods are government provided, the participation constraints (IR) may seem questionable. One defense in this context is that the participation constraint is a reduced form of an environment where agents may vote with their feet. Another defense is to view this as a reduced form for inequality aversion of the planner. See Hellwig [13].

2.4 Simple Anonymous Mechanisms

To simplify the analysis, we first exploit the symmetry and linearity of the constraints and the objective function. This allows us to reduce the dimensionality of the problem:

Definition 1 *A mechanism is called a simple mechanism if it can be expressed as* $(2m+1)$*-tuple* $g = \left(\left\{\rho^j\right\}_{j\in\mathcal{J}}, \left\{\eta^j\right\}_{j\in\mathcal{J}}, t\right)$ *where for each* $j \in \mathcal{J}$,

$$\begin{aligned} \textit{Provision Rule:} \quad & \rho^j : \Theta^n \to [0,1] \\ \textit{Inclusion Rule:} \quad & \eta^j : \Theta \to [0,1] \\ \textit{Cost-sharing Rule:} \quad & t : \Theta \to \mathrm{R}, \end{aligned} \tag{5}$$

ρ^j *is the provision rule for good* j, η^j *is the inclusion rule for good* j *(same for all agents), and* t *is the transfer rule (also same for all agents).*

There are a number of simplifications in (5) relative to (3). First, inclusion and transfer rules are the same for all agents; second, conditional on θ, the provision probability $\rho^j(\theta)$ is stochastically independent from all other provision probabilities, and all inclusion probabilities; third, the inclusion and transfer rules for any agent i are independent of the realization of θ_{-i}; and fourth, all agents are treated symmetrically in terms of the transfer and inclusion rules.

Symmetry in inclusion and transfer rules is built into the notion of a simple mechanism, but (5) allows provision rules to treat agents asymmetrically. We therefore need a definition to express what it means for the index of the agent to be irrelevant:

Definition 2 *A simple mechanism is called anonymous if* $\rho^j(\theta) = \rho^j(\theta')$ *for every* $j \in \mathcal{J}$, *every* $\theta \in \Theta^n$, *and every* $\theta' \in \Theta^n$ *that can be obtained from* θ *by permuting the indices of the agents.*

We now show that focusing on simple anonymous mechanisms is without loss of generality:

Proposition 1 *For any incentive feasible mechanism* $\mathcal{G}$ *of the form (3), there exists an anonymous simple incentive feasible mechanism* g *of the form (5) that generates the same social surplus.*

Consequently, the remainder of this paper only considers simple anonymous mechanisms. The idea is roughly that risk neutral agents care only about the perceived probability of consuming each good and the expected transfer. Therefore, there is nothing to gain from conditioning transfers and inclusion probabilities on θ_{-i}, or by making inclusion and provision rules conditionally dependent. Mechanisms of the form (5) are therefore sufficient. Moreover, given an incentive feasible mechanism, permuting the roles of the agents leaves the surplus unchanged and all constraints satisfied. An anonymous incentive feasible mechanism that generates the same surplus as the initial mechanism can therefore be obtained by averaging over the $n!$ permuted mechanisms.[6]

[6] The actual proof is a bit more complex than simply randomizing with equal probabilities over the $n!$ permutations. The reason is that inclusion and provision probabilities are potentially correlated.

2.5 Symmetric Treatment of the Goods

Our next result, on which we rely heavily in Sections 3 and 4, identifies conditions under which it is without loss of generality to treat goods symmetrically. Obviously, the underlying environment must be symmetric, and we formalize this by assuming that $\theta_i = \left(\theta_i^1, \cdots, \theta_i^m\right)$ is an *exchangeable* random variable, that is $F(\theta_i) = F\left(\theta_i'\right)$ whenever θ_i' is a permutation of θ_i; and that there exists $C(n)$ such that $C^j(n) = C(n)$ for all j.

Given valuation profile θ and a one-to-one permutation mapping $P : \mathcal{J} \to \mathcal{J}$ of the set of goods, let θ_i^P denote the permutation of agent i's type by changing the role of the goods in accordance to P : that is, $\theta_i^P = \left(\theta_i^{P^{-1}(1)}, \theta_i^{P^{-1}(2)} ..., \theta_i^{P^{-1}(m)}\right)$, where P^{-1} denote the inverse of P. For simplicity, write $\theta^P \equiv \left(\theta_1^P, ..., \theta_n^P\right)$ as the valuation profile obtained when the role of the goods is changed in accordance to P for every $i \in \mathcal{I}$.

Definition 3 *Mechanism g is symmetric if for every θ and every permutation $P : \mathcal{J} \to \mathcal{J}$:*

1. $\rho^{P^{-1}(j)}\left(\theta^P\right) = \rho^j(\theta)$ *for every* $j \in \mathcal{J}$;
2. $\eta^{P^{-1}(j)}\left(\theta_i^P\right) = \eta^j(\theta_i)$ *for every* $j \in \mathcal{J}$;
3. $t\left(\theta_i^P\right) = t(\theta_i)$.

In defining a symmetric mechanism, the *same permutation of goods must be applied for all agents*. As an example, suppose that there are two agents and two goods, and that the valuation for each good is either h or l. In this case $\Theta = \{(h,h), (h,l), (l,h), (l,l)\}$. Consider the type profile $\theta = (\theta_1, \theta_2) = ((h,l), (l,h)) \in \Theta^2$. Applying the only non-identity permutation of the goods, i.e., $P(1) = 2$ and $P(2) = 1$, to all agents generates a type profile $\theta^P = \left(\theta_1^P, \theta_2^P\right) = ((l,h), (h,l))$. Definition 3 requires that the allocations for type profile $((l,h), (h,l))$ is the same as the allocation for $((h,l), (l,h))$ with goods relabeled, and that transfers are unchanged.[7] The result is:

Proposition 2 *Suppose that θ_i is an exchangeable random variable and that there exists $C(n)$ such that $C^j(n) = C(n)$ for all $j \in \mathcal{J}$. Then, for any simple anonymous incentive feasible mechanism g, there exists a simple anonymous and symmetric incentive feasible mechanism that generates the same surplus as g.*

The idea is similar to that of Proposition 1, except that it is the role of the goods that are permuted. Consider the case with two goods, and suppose that the two goods are treated asymmetrically. Reversing the role of the goods, an alternative mechanism that generates the same

[7] If we were to apply different permutations for the two agents, e.g., applying the identity permutation for agent 1 and the non-identiy permutation for agent 2, then we would obtain a profile $((h,l), (h,l))$, which is a qualitatively different from either $((h,l), (l,h))$ or $((l,h), (h,l))$. In the profile $((h,l), (h,l))$, both agents have low valuations for good 2 and high valuations for good 1, whereas, in the profiles $((h,l), (l,h))$ or $((l,h), (h,l))$, one and only one agent has high valuation for both goods.

surplus is obtained. Averaging over the original and the reversed mechanism creates a symmetric mechanism where surplus is unchanged.[8] Incentive feasibility of the new mechanism follows from incentive feasibility of the original mechanism. Proposition 2 generalizes this procedure by permuting the goods ($m!$ possibilities) and creating a symmetric mechanism by averaging over these.

3 The Model with Binary Valuations

Assume that there are two public goods, and that the valuation for good j can either be "high" ($\theta_i^j = h$) or "low" ($\theta_i^j = l$). For notational brevity we henceforth write the typespace for an individual as $\Theta = \{hh, hl, lh, ll\}$, and assume that θ_i is independently drawn from an identical joint distribution over Θ according to probability distribution $\mu = (\alpha_{hh}, \alpha_{hl}, \alpha_{lh}, \alpha_{ll}) \in \Delta^4$. To apply Proposition 2, we assume that the goods are symmetric, so that $\alpha_{hl} = \alpha_{lh} \equiv \alpha_m$ and $C^1(n) = C^2(n) = cn$.[9] For future reference, let $\alpha \equiv \alpha_{hh} + \alpha_m$ be the marginal probability that an agent's valuation for a public good is h. To keep the problem non-trivial, we also assume that $l < c < h$.

Appealing to Propositions 1 and 2, we only consider simple anonymous mechanisms that treat the two public goods symmetrically. For each $\theta \in \Theta^n \equiv \{hh, hl, lh, ll\}^n$, let $x \equiv (x_{hh}, x_{hl}, x_{lh}, x_{ll})$ denote the number of agents announcing different types, and let

$$\mathcal{X}_n = \left\{ x \in \{0, ..., n\}^4 : x_{hh} + x_{hl} + x_{lh} + x_{ll} = n \right\}. \tag{6}$$

be the set of possible values of x in an economy with n agents. Anonymity means that the provision rule depends only on the *number of agents* who announce different valuation combinations. With some notational abuse, it is thus without loss of generality to consider mechanisms of the form

$$\mathcal{M} = \left(\rho^1, \rho^2, \eta, t\right) \tag{7}$$

where $\rho^j : \mathcal{X}_n \to [0,1]$ for $j = 1, 2$, $\eta \equiv (\eta_{hh}, \eta_m^h, \eta_m^l, \eta_{ll}) \in [0,1]^4$ and $t \equiv (t_{hh}, t_m, t_{ll}) \in \mathrm{R}^3$. That is, Proposition 1 states that inclusion probabilities without loss can be independent of types of other agents. Because of symmetric treatment of the goods (Proposition 2), it is sufficient with a single inclusion probability η_{hh}, which is the probability of access to both goods for type hh. For the same reason it is enough with a single inclusion probability for type ll. For types hl and lh, η_m^h is the probability for access to the high valuation good, and η_m^l is the probability for access to the low valuation good, where again the symmetric treatment comes from Proposition 2. The argument for the cost-sharing follows the same lines, only a bit simpler.

Symmetric treatment of the goods also has implications on how ρ^1 relates to ρ^2 which are used in Section 4.1 to simplify the incentive constraints.

[8] Provision probabilities and taxes are given by straightforward averaging, but since inclusion and provision probabilities may be correlated the procedure is somewhat more involved for the inclusion rules.

[9] Keeping the per capita costs constant simplifies notation, but is not necessary.

3.1 Optimal Separate Provision Mechanisms

As a benchmark, this section derives the asymptotic provision probabilities of the two public goods when the provision problem for each public good is considered in isolation. Proposition 1 applies also to the case with a single good, which for the binary case means that the provision rule may be taken to depend only on the number of agents who announce a high valuation. To emphasize that the solution depends on the size of the economy, we denote a separate provision mechanism for good j in an economy of size n as a triple $(\rho_n^j, \eta_n^j, t_n^j)$, where $\rho_n^j : \{1, ..., n\} \to [0,1]$ and $\rho_n^j(\kappa)$ denotes the probability of provision if κ agents announce a high valuation for good j; $\eta_n^j \in [0,1]$ is the inclusion probability for type l and $t_n^j = (t_n^j(h), t_n^j(l))$ are the transfers.[10]

To find the best provision mechanism where goods are provided separately is formally the same problem as finding the best provision mechanism when there is only a single good. Maximizing social surplus subject to the single-good analogues of (IC), (BB) and (IR) in Section 2.3 one obtains the following characterization of the constrained optimal separate provision mechanism:

Proposition 3 *Consider a sequence of economies of size* $\{n\}_{n=1}^{\infty}$. *Then,*
(1) if $\alpha h < c$, $\lim_{n\to\infty} \mathrm{E}\rho_n^j(\kappa) = 0$ *for any sequence of feasible separate provision mechanisms* $\left\{\rho_n^j, \eta_n^j, t_n^j\right\}$;
(2) if $\alpha h > c$, $\lim_{n\to\infty} \mathrm{E}\rho_n^{*j}(\kappa) = 1$ *for any sequence of constrained optimal separate provision mechanisms* $\left\{\rho_n^{*j}, \eta_n^{*j}, t_n^{*j}\right\}$. *Moreover, any sequence of constrained optimal mechanisms satisfies*

$$\lim_{n\to\infty} \eta_n^{*j} = \frac{\alpha h - c}{\alpha h - l}, \quad \lim_{n\to\infty} t_n^{*j}(l) = \frac{\alpha h - c}{\alpha h - l} l, \text{ and } \lim_{n\to\infty} t_n^{*j}(h) = \left[1 - \frac{\alpha h - c}{\alpha h - l}\right] h + \frac{\alpha h - c}{\alpha h - l} l.$$

The result is a two-type analogue to Propositions 2 and 3 in Norman [21], and we only provide a heuristic explanation.[11] The key idea is that the incentive constraint for the high type

$$\mathrm{E}\left[\rho_n^{*j}(\kappa) \,|\theta_i^j = h\right] h - t_n^{*j}(h) \geq \mathrm{E}\left[\rho_n^{*j}(\kappa) \,|\theta_i^j = l\right] \eta_n^{*j} h - t_n^{*j}(l), \tag{8}$$

may be replaced by

$$\mathrm{E}\rho_n^{*j}(\kappa)\, h - t_n^{*j}(h) \geq \mathrm{E}\rho_n^{*j}(\kappa)\, \eta_n^{*j} h - t_n^{*j}(l), \tag{9}$$

since the probability that agent i is pivotal for the provision decision is negligible in a large economy. Moreover, the participation constraint for the low type binds, implying that $t_n^{*j}(l) = \mathrm{E}\rho_n^{*j}(\kappa)\, \eta_n^{*j} l$. Because (8) binds in the optimal mechanism, budget balance requires that, approximately,

$$\begin{aligned} \mathrm{E}\rho_n^{*j}(\kappa)\, c &= \alpha t_n^{*j}(h) + (1-\alpha)\, t_n^{*j}(l) \approx \alpha \left[t_n^{*j}(l) + \mathrm{E}\rho_n^{*j}(\kappa)\, h \left(1 - \eta_n^{*j}\right)\right] + (1-\alpha)\, t_n^{*j}(l) \\ &= t_n^{*j}(l) + \alpha \mathrm{E}\rho_n^{*j}(\kappa)\, h \left(1 - \eta_n^{*j}\right) = \mathrm{E}\rho_n^{*j}(\kappa) \left[\eta_n^{*j} l + \alpha h \left(1 - \eta_n^{*j}\right)\right]. \end{aligned} \tag{10}$$

[10] Exclusions of type h agents are also feasible, but never occur in an optimal mechanism, since excluding type h tightens the downwards incentive constraint for h.

[11] Details available on request from the authors.

Hence, $\eta_n^{*j} \approx (\alpha h - c) / (\alpha h - l)$ follows from (10). Inspecting (10), it follows that $\lim_{n\to\infty} \mathrm{E}\rho_n^{*j}(\kappa) = 0$ if $\alpha h < c$ (since the bracketed expression is maximized at either l or αh and $l < c$ by assumption). Otherwise the budget balance constraint must be violated for large n. On the other hand, if $\alpha h > c$, it is feasible to provide for sure (for any n) with the transfers specified in Proposition 3, and inclusion probability $\eta_\infty^* \equiv (\alpha h - c) / (\alpha h - l)$. Conditional on this inclusion probability, the ex post efficient rule is to provide public good j whenever $\frac{\kappa}{n} h + \frac{n-\kappa}{n} \eta_\infty^* l \geq c$. An application of Chebyshev's inequality guarantees that

$$\text{plim}\left(\frac{\kappa}{n} h + \frac{n-\kappa}{n} \eta_\infty^* l\right) = \alpha h + (1-\alpha) \frac{\alpha h - c}{\alpha h - l} l > \alpha h > c.$$

Thus, the ex post efficient provision rule conditional on the given inclusion probability converges towards "always provide." Hence $\lim_{n\to\infty} \mathrm{E}\rho_n^{*j}(\kappa) = 1$ in the optimal mechanism. The limits for the transfers can then be obtained by substituting $\lim_{n\to\infty} \mathrm{E}\rho_n^{*j}(\kappa) = 1$ back into the incentive and participation constraints.

The optimal separate provision mechanism characterized in Proposition 3 is bounded away from first best efficiency. First of all, the asymptotic provision probability is zero when $\alpha h < c$ while efficiency requires provision whenever $\alpha h + (1-\alpha) l > c$. Moreover, when $\alpha h > c$, there is still a distortion due to positive probability of exclusion of low valuation agents.

3.2 Efficiency Gains From Bundling

Before deriving the constrained optimal mechanism, we consider an example that shows that bundling can lead to provision for sure, even though the best separate provision mechanism has an asymptotic provision probability equal to zero.

Suppose that $\alpha h + (1-\alpha) l > c$, so that provision is desirable in a large economy with a probability near one. Consider mechanism

$$\begin{aligned} t_{hh} &= t_m = h + l, \text{ and } t_{ll} = 0 \\ \eta_{hh} &= \eta_m^h = \eta_m^l = 1, \text{ and } \eta_{ll} = 0 \\ \rho^1(x) &= \rho^2(x) = 1 \text{ for all } x \in \mathcal{X}_n. \end{aligned} \tag{11}$$

That is, agents of type hh, hl and lh are taxed the willingness to pay of a mixed type and consume both goods for sure. Type-ll pays nothing, but is excluded from usage from both goods. All incentive and participation constraints are trivially satisfied by mechanism (11). The only question is thus whether the feasibility constraint (BB) is satisfied, that is, if

$$\Pr[\{hh, hl, lh\}](h + l) = (\alpha_{hh} + 2\alpha_m)(h + l) \geq 2c, \tag{12}$$

holds. It is easy to show that:

Claim 1 *For any $c > 0$, and $(\alpha_{hh}, \alpha_m, \alpha_m, \alpha_{ll}) \in \Delta^4$ such that $\alpha_m > \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$, there exist pairs (h, l) with $h > c > l$ such that (12) is satisfied, and at the same time $(\alpha_{hh} + \alpha_m) h < c$.*

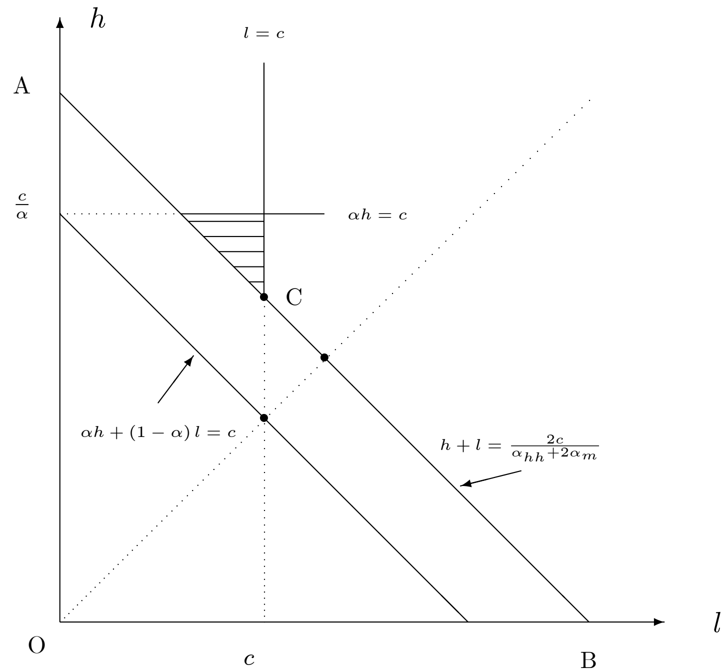

Figure 1: The Bundling Mechanism Outperforms Optimal Non-bundling Mechanism in the Shaded Region.

The set of the values of h and l for which (11) outperforms the best separate provision mechanism are depicted as the shaded region in Figure 1. It is easy to see that the shaded region will be non-empty as long as the line $\alpha h = c$ is above point C in Figure 1. Point C has a coordinate of $(h, l) = (2c/(\alpha_{hh} + 2\alpha_m) - c, l)$. Thus $\alpha h - c$ evaluated at point C is

$$\begin{aligned} \alpha h - c &= c\left[\frac{2(\alpha_{hh} + \alpha_m)}{\alpha_{hh} + 2\alpha_m} - (\alpha_{hh} + \alpha_m) - 1\right] = c\left[\frac{\alpha_{hh}}{\alpha_{hh} + 2\alpha_m} - (\alpha_{hh} + \alpha_m)\right] \\ &= c\left(\frac{\alpha_{hh}\alpha_{ll}}{1 - \alpha_{ll}} - \alpha_m\right) < 0, \end{aligned}$$

where the inequality follows from the assumption that $\alpha_m > \alpha_{hh}\alpha_{ll}/(1 - \alpha_{ll})$. The inequality, which requires that the valuations for the two goods are not too positively correlated, is needed because the revenue effect from bundling depends on a trade-off between price and the number of agents willing to pay the price. If the valuations are strongly positively correlated, there are too few mixed types for the increased sales to make up for the reduction in price.

The expected utility in the best separate provision mechanism approaches zero for all agents when provisions go to zero, whereas type-hh enjoys utility level $h - l > 0$ under mechanism (11). The proposed bundling mechanism therefore improves efficiency. Hence, we've shown that a feasible bundling mechanism can improve the efficiency over the optimal separate provision mechanisms. This is akin in spirit to McAfee et al [18] who used local analysis showing that bundling improves the profits for a monopolist. Also notice that if valuations are independent, $\alpha_{hh} = \alpha^2, \alpha_{ll} = (1 - \alpha)^2$, and $\alpha_m = \alpha(1-\alpha)$, and the inequality in Claim 1 is satisfied. Thus, just like in McAfee et al [18], the argument applies also to the case when valuations across goods are stochastically independent.

4 The Constrained Optimal Mechanism

In this section, we proceed a step further and characterize the *constrained optimal* mechanism for the binary version of the model described in the previous section.

4.1 The Constraints

There are 12 incentive constraints, but the payoff for type hh or ll from pretending to be hl is the same as the payoff from pretending to be lh. Symmetrically, the payoff for type hl is the same as the payoff for type lh regardless of whether the true type hh or ll is announced (announcing the other mixed type affects the payoff). Due to the symmetry, 5 constraints can thus be immediately discarded. Using that $\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}) = \rho^2(x_{hh}, x_{lh}, x_{hl}, x_{ll})$ (Proposition 2) we express the 7 remaining incentive compatibility constraints purely in the provision rule for good 1 as,

$$\begin{aligned}
\eta_{hh}\mathrm{E}\left[\rho^1(x)|hh\right]2h - t_{hh} &\geq \eta_m^h\mathrm{E}\left[\rho^1(x)|hl\right]h + \eta_m^l\mathrm{E}\left[\rho^1(x)|lh\right]h - t_m &(13)\\
\eta_{hh}\mathrm{E}\left[\rho^1(x)|hh\right]2h - t_{hh} &\geq \eta_{ll}\mathrm{E}\left[\rho^1(x)|ll\right]2h - t_{ll} &(14)\\
\eta_m^h\mathrm{E}\left[\rho^1(x)|hl\right]h + \eta_m^l\mathrm{E}\left[\rho^1(x)|lh\right]l - t_m &\geq \eta_{hh}\mathrm{E}\left[\rho^1(x)|hh\right](h+l) - t_{hh} &(15)\\
\eta_m^h\mathrm{E}\left[\rho^1(x)|hl\right]h + \eta_m^l\mathrm{E}\left[\rho^1(x)|lh\right]l &\geq \eta_m^h\mathrm{E}\left[\rho^1(x)|hl\right]l + \eta_m^l\mathrm{E}\left[\rho^1(x)|lh\right]h &(16)\\
\eta_m^h\mathrm{E}\left[\rho^1(x)|hl\right]h + \eta_m^l\mathrm{E}\left[\rho^1(x)|lh\right]l - t_m &\geq \eta_{ll}\mathrm{E}\left[\rho^1(x)|ll\right](h+l) - t_{ll} &(17)\\
\eta_{ll}\mathrm{E}\left[\rho^1(x)|ll\right]2l - t_{ll} &\geq \eta_m^h\mathrm{E}\left[\rho^1(x)|hl\right]l + \eta_m^l\mathrm{E}\left[\rho^1(x)|lh\right]l - t_m &(18)\\
\eta_{ll}\mathrm{E}\left[\rho^1(x)|ll\right]2l - t_{ll} &\geq \eta_{hh}\mathrm{E}\left[\rho^1(x)|hh\right]2l - t_{hh}. &(19)
\end{aligned}$$

Since all other types can always mimic type ll, the only relevant participation constraint is

$$\eta_{ll}\mathrm{E}\left[\rho^1(x)|ll\right]2l - t_{ll} \geq 0. \tag{20}$$

Finally, the feasibility constraint (BB) can be simplified due to the simple transfer schemes and the constant per capita costs. Again appealing to the symmetric treatment of the goods (Proposition 2) we can express (BB) in per capita form as

$$\alpha_{hh}t_{hh} + 2\alpha_m t_m + \alpha_{ll}t_{ll} - 2c\mathrm{E}\rho^1(x) \geq 0. \tag{21}$$

4.2 Relaxed Problem: The Main Case

The most involved part of the design problem is the provision rule. This is difficult because $\rho^j(x)$ is weighted by the ex ante probability that x occurs in the objective function, while the relevant probabilities in the constraints are conditional probabilities. To be able to link the unconditional and conditional probabilities we need to be explicit about the (multinomial) probability distribution of x. Given n agents, we denote the probability of outcome $x \in \mathcal{X}_n$ by $\mathbf{a}_n(x)$, which follows a multinomial with parameters $(n, \alpha_{hh}, \alpha_m, \alpha_m, \alpha_{ll})$.[12]

Due to the symmetry, types are naturally ordered as hh being the "highest type", hl and lh being "middle types" and ll being the "lowest type". Appealing to intuition from single-dimensional problems, we conjecture that only adjacent downwards incentive constraints are relevant and will therefore ignore the upwards constraints, (15),(18) and (19), as well as (16), the constraints between type hl and lh, and the downward constraint between hh and ll, (14).[13] Expressing the two remaining constraints ((13) and (17)) explicitly in terms of the multinomial probabilities, we get

$$\begin{aligned} &\eta_{hh} \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\, \rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll})\, 2h - t_{hh} && (22)\\ \geq\ &\eta_m^h \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\, \rho^1(x_{hh}, x_{hl}+1, x_{lh}, x_{ll}) h + \eta_m^l \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\, \underbrace{\rho^1(x_{hh}, x_{hl}, x_{lh}+1, x_{ll})}_{=\rho^2(x_{hh}, x_{hl}+1, x_{lh}, x_{ll})} h - t_m \\ &\eta_m^h \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\, \rho^1(x_{hh}, x_{hl}+1, x_{lh}, x_{ll}) h + \eta_m^l \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\, \underbrace{\rho^1(x_{hh}, x_{hl}, x_{lh}+1, x_{ll})}_{=\rho^2(x_{hh}, x_{hl}+1, x_{lh}, x_{ll})} l - t_m \\ \geq\ &\eta_{ll} \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\, \rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\,(h+l) - t_{ll}, && (23) \end{aligned}$$

where (22) states that type-hh agents do not have incentives to mimic type hl; and (23) states that mixed type agents do not have incentives to mis-report as type ll. The participation constraint (20) and the feasibility constraint (21) written explicitly in terms of the multinomial distribution are

$$\eta_{ll} \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\, \rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1) 2l - t_{ll} \geq 0 \qquad (24)$$

$$\alpha_{hh} t_{hh} + 2\alpha_m t_m + \alpha_{ll} t_{ll} - 2c \sum_{x \in \mathcal{X}_n} \mathbf{a}_n(x)\, \rho^1(x) \geq 0. \qquad (25)$$

Again appealing to Proposition 2, good 2 can be eliminated from the objective function. Expressing social surplus in per capita form, the relaxed programming problem is:[14]

$$\max_{\{\rho^1, \eta, t\}} 2 \sum_{x \in \mathcal{X}_n} \mathbf{a}_n(x)\, \rho^1(x) \left[\frac{\left(\eta_{hh} x_{hh} + \eta_m^h x_{hl}\right) h + \left(\eta_m^l x_{lh} + \eta_{ll} x_{ll}\right) l}{n} - c \right] \qquad (26)$$

s.t. (22),(23), (24), and (25),

$$\eta \in [0,1]^4, \rho^1(x) \geq 0, 1 - \rho^1(x) \geq 0 \text{ for each } x \in \mathcal{X}_n, \qquad (27)$$

Lemma 1 *There exists at least one optimal solution to (26).*

[12] Explicitly, $\mathbf{a}_n(x) = \mathbf{a}_n(x_{hh}, x_{hl}, x_{lh}, x_{ll}) = \frac{n!}{x_{hh}! x_{hl}! x_{lh}! x_{ll}!} \alpha_{hh}^{x_{hh}} \alpha_m^{x_{hl}} \alpha_m^{x_{lh}} \alpha_{ll}^{x_{ll}}$.

[13] In Section 4.6 we will check the conditions for when this procedure is valid. It turns out that the only potential problem is ignoring the downward constraint between hh and ll.

[14] The multiplicative constant 2 in the objective function is redundant, but it aids interpretations by keeping the units in the objective function and the constraints comparable.

The proof is standard by first compactifying the constraint set and then applying Weierstrass Theorem. Slater's condition for constraint qualification holds, so the Kuhn-Tucker conditions are necessary for an optimum. Since a solution to (26) exists, these conditions characterize an optimal mechanism, provided that the constraints that have been ignored are satisfied.

4.3 Linking the Multipliers

Taxes enter linearly into all constraints and are not constrained by boundaries. It is therefore convenient to begin the analysis from the first order conditions with respect to $t = (t_{hh}, t_m, t_{ll})$,

$$\begin{array}{cc} (\text{w.r.t. } t_{hh}) & -\lambda_{hh} + \Lambda\alpha_{hh} = 0 \\ (\text{w.r.t. } t_m) & \lambda_{hh} - \lambda_m + 2\Lambda\alpha_m = 0 \\ (\text{w.r.t. } t_{ll}) & \lambda_m - \lambda_{ll} + \Lambda\alpha_{ll} = 0 \end{array} . \tag{28}$$

Here, λ_{hh} and λ_m are the multipliers associated with (incentive compatibility) constraints (22) and (23), λ_{ll} is the multiplier for the (participation) constraint (24), and Λ is the multiplier for the (resource) constraint (25). It is immediate from (28) that:

Lemma 2 *In any solution to (26), the multipliers* $(\lambda_{hh}, \lambda_m, \lambda_{ll}, \Lambda)$ *satisfy:* $\lambda_{hh} = \alpha_{hh}\Lambda, \lambda_m = (\alpha_{hh} + 2\alpha_m)\Lambda$, *and* $\lambda_{ll} = \Lambda$.

In all its simplicity, Lemma 2 is a key step in the solution of (26). Its role is similar to the characterization of incentive feasibility in terms of a single integral constraint in single-dimensional mechanism design problem (i.e., the approach in Myerson [19] and others). In multidimensional problems, it is impossible to collapse all constraints into a single constraint. Instead, Lemma 2 allows us to *indirectly* relate all optimality conditions to a single constraint.

4.4 Inclusion Rules

We now characterize the optimal inclusion rules η. To ease the statement of the result, we define two linear functions $G : [0,1] \to \mathrm{R}$ and $H : [0,1] \to \mathrm{R}$ as

$$\begin{aligned} G(\Phi) &\equiv (1-\Phi)\,2l + \Phi\left(\frac{\alpha_{hh} + 2\alpha_m}{\alpha_m} l - \frac{\alpha_{hh}}{\alpha_m} h\right), \\ H(\Phi) &\equiv (1-\Phi)\,2l + \Phi\left[\frac{2}{\alpha_{ll}} l - \frac{\alpha_{hh} + 2\alpha_m}{\alpha_{ll}}(h+l)\right], \end{aligned} \tag{29}$$

which allow us to express the optimal inclusion rules in terms of the multiplier on (25) as:

Lemma 3 *Let* $\mathcal{M} = (\rho^1, \rho^2, \eta, t)$ *be a symmetric solution to (26) and let* $\Phi = \Lambda/(1+\Lambda)$, *where* Λ *is the associated multiplier on the resource constraint (25). Also, suppose that* $E\big[\rho^j(x)\,|\theta_i\big] > 0$ *for all* $\theta_i \in \Theta$ *and* $j = 1, 2$. *Then,*

$$\textit{(i)}\ \eta_{hh} = \eta_m^h = 1;\ \textit{(ii)}\ \eta_m^l = \begin{cases} 1 & \textit{if } G(\Phi) > 0 \\ y \in [0,1] & \textit{if } G(\Phi) = 0 \\ 0 & \textit{if } G(\Phi) < 0; \end{cases} ;\ \textit{(iii)}\ \eta_{ll} = \begin{cases} 1 & \textit{if } H(\Phi) > 0 \\ y \in [0,1] & \textit{if } H(\Phi) = 0 \\ 0 & \textit{if } H(\Phi) < 0. \end{cases}$$

Multiplying both sides of $G(\Phi)$ by α_m, we see that $G(\Phi) \geq 0$ if and only if

$$\Phi \overbrace{\left[(\alpha_{hh} + 2\alpha_m) l - \alpha_{hh} h\right]}^{\text{Term 1}} + (1 - \Phi) \overbrace{2\alpha_m l}^{\text{Term 2}} \ \geq\ 0, \tag{30}$$

where $\Phi = \Lambda / (1 + \Lambda) \in [0, 1]$. To understand Term 1 in expression (30), consider two candidate inclusion rules. The first candidate is $\eta_m^l = \eta_{ll} = 0$, and $\eta_{hh} = \eta_m^h = 1$. That is, an agent is given access to good j if and only if her announced valuation for good j is h. Since high valuation agents are willing to pay h for access to a good, the expected revenue from such an inclusion rule is at most $2(\alpha_{hh} + \alpha_m) h$ from each agent. The second candidate inclusion rule is $\eta_m^l = \eta_{hh} = \eta_m^h = 1$ and $\eta_{ll} = 0$. That is, an agent is given access to both goods as long as one of her announced valuation is high. Under this inclusion rule, all agent types except ll could be charged $h + l$ for access to both goods. This results in an expected revenue per agent of $(\alpha_{hh} + 2\alpha_m)(h + l)$. The change in revenue from increasing η_m^l from 0 to 1 is thus

$$(\alpha_{hh} + 2\alpha_m)(h + l) - 2(\alpha_{hh} + \alpha_m) h = (\alpha_{hh} + 2\alpha_m) l - \alpha_{hh} h,$$

which is Term 1 in (30). Term 1 thus captures the expected gain or loss in revenue from increasing η_m^l from 0 to 1. Term 2 in expression (30), $2\alpha_m l$, on the other hand, is simply the increase in the expected per capita social surplus from increasing η_m^l from 0 to 1. In sum, this means that $G(\Phi)$ is a weighted average of the optimality conditions for an unconstrained social planner and a profit maximizing provider, where the weight on Term 1 – the effect on revenue – is higher when the shadow price of revenue, namely, Λ, is higher. If $(\alpha_{hh} + 2\alpha_m) l - \alpha_{hh} h$ is positive, the mixed types will get access to both goods for sure, whereas it otherwise depends on the shadow price on the resource constraint.

Analogously, $H(\Phi) \geq 0$ if and only if

$$\Phi \left[2l - (\alpha_{hh} + 2\alpha_m)(h + l)\right] + (1 - \Phi) 2\alpha_{ll} l \geq 0.$$

The term $2l - (\alpha_{hh} + 2\alpha_m)(h + l)$ is the revenue effect (which could be positive or negative) when η_{ll} is increased from 0 to 1; and the term $2\alpha_{ll} l$ reflects the gain in social surplus from such a change. Thus, $H(\Phi)$ is again a weighted average of the optimality conditions for an unconstrained social planner and a profit maximizing provider. If $2l - (\alpha_{hh} + 2\alpha_m)(h + l) > 0$, then $H(\Phi) > 0$ for sure and $\eta_{ll} = 1$ is optimal. Otherwise, the access decision for type ll will depend on the shadow price on the resource constraint.

4.5 Provision Rules

To discuss the optimal provision rules $\left\{\rho^j(x)\right\}_{j=1,2}$, it is convenient to first define

$$\begin{aligned} Q^1\left(\frac{x}{n},\Phi\right) &\equiv \frac{x_{hh}}{n}h+\frac{x_{hl}}{n}h+\frac{x_{lh}}{n}\frac{\max\{0,G(\Phi)\}}{2}+\frac{x_{ll}}{n}\frac{\max\{0,H(\Phi)\}}{2}-c. \\ Q^2\left(\frac{x}{n},\Phi\right) &\equiv \frac{x_{hh}}{n}h+\frac{x_{lh}}{n}h+\frac{x_{hl}}{n}\frac{\max\{0,G(\Phi)\}}{2}+\frac{x_{ll}}{n}\max\frac{\{0,H(\Phi)\}}{2}-c. \end{aligned} \tag{31}$$

These functions can also be interpreted as maximizing a weighted average of surplus and revenue. To see this, first consider $\Phi=0$, in which case [see definitions in (29)] $G(0)=H(0)=2l$. The value of $Q^j(x/n,0)$ is thus simply the social surplus generated if good j is provided and nobody is excluded. Similarly, as discussed in the previous section, $G(1)$ is the gain or loss in revenue if mixed types are allowed to consume their low valuation good.[15] The constrained optimal provision rule can be described in terms of Q^1 and Q^2 as:

Lemma 4 *Let $\mathcal{M}$ be an optimal solution to (26) and $\Phi=\Lambda/(1+\Lambda)$ where Λ is the multiplier associated with the constraint (25) at the optimal solution. Then, (1) $\rho^j(x)=1$ whenever $Q^j(x/n,\Phi)>0$; and (2) $\rho^j(x)=0$ whenever $Q^j(x/n,\Phi)<0$.*

To summarize, we have characterized the optimal inclusion and provision rules for any given value of the Lagrange multiplier Λ associated with the feasibility constraint. Such characterization provides some partial information regarding the asymptotic provision probability in the optimal mechanism with bundling. For example, the above characterization tells us that $\alpha h>c$ is a sufficient but not necessary condition for the provision probability to converge to one.[16] To see this, note that $\frac{x}{n}$ converges in probability to $\mu=(\alpha_{hh},\alpha_m,\alpha_m,\alpha_{ll})$. By continuity of Q^j we therefore have that $Q^j\left(\frac{x}{n},\Phi_n\right)$ converges in probability to

$$Q^j(\mu,\Phi_n)=(\alpha_{hh}+\alpha_{hl})h+\alpha_m\frac{\max\{0,G(\Phi_n)\}}{2}+\alpha_{ll}\frac{\max\{0,H(\Phi_n)\}}{2}-c, \tag{32}$$

which is strictly positive if $(\alpha_{hh}+\alpha_{hl})h=\alpha h>c$, which by Lemma 4 guarantees provision with probability 1 in the limit.

4.6 Checking the Remaining Constraints

While the value of the multiplier Φ is still unknown, we now know enough about the solution to the relaxed program (26) to check when the constraints we have ignored are indeed satisfied:

[15] The same is true about $H(\Phi)$, but given the non-triviality assumptions on the problem, giving access to type ll always reduces revenue.

[16] Recall that the bundling mechanism in Section 3.2 achieves provision with probability one for some cases despite $\alpha h<c$. In contrast, in the model without bundling, $\alpha h>c$ is necessary and sufficient for asymptotic provision with probability one.

Lemma 5 *Suppose that constraints (13) and (17), or equivalently constraints (22) and (23) in the relaxed program (26), bind. Moreover suppose that* $\eta^h_m \mathrm{E}\left[\rho^1(x)|hl\right] \geq \eta^l_m \mathrm{E}\left[\rho^1(x)|lh\right] \geq \eta_{ll} \mathrm{E}\left[\rho^1(x)|ll\right]$ *and* $\eta_{hh} \mathrm{E}\left[\rho^1(x)|hh\right] \geq \eta^l_m \mathrm{E}\left[\rho^1(x)|lh\right]$. *Then, all the remaining incentive compatibility constraints are satisfied.*

Algebra on the difference between $G(\Phi)$ and $H(\Phi)$ shows that it is positive if and only if the inequality in Claim 1 is satisfied. That is,

Lemma 6 *If* $\alpha_m \geq \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$, *then* $G(\Phi) \geq H(\Phi)$ *for any* $\Phi \geq 0$. *The inequality is strict whenever* $\alpha_m > \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$ *and* $\Phi > 0$. *Conversely,* $\alpha_m < \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$ *and* $\Phi > 0$, *then* $G(\Phi) < H(\Phi)$.

Letting $x \in \mathcal{X}_{n-1}$ denote a realized numbers of agents of the four different types other than i (which follows a multinomial with parameters $(n-1, \alpha_{hh}, \alpha_m, \alpha_m, \alpha_{ll})$) we have that, if $\alpha_m \geq \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$,

$$\begin{aligned} Q^1\left(\frac{x_{hh}+1}{n}, \frac{x_{hl}}{n}, \frac{x_{lh}}{n}, \frac{x_{ll}}{n}, \Phi\right) &= Q^1\left(\frac{x_{hh}}{n}, \frac{x_{hl}+1}{n}, \frac{x_{lh}}{n}, \frac{x_{ll}}{n}, \Phi\right) \\ /G(\Phi) \leq 2l/ \quad &> Q^1\left(\frac{x_{hh}}{n}, \frac{x_{hl}}{n}, \frac{x_{lh}+1}{n}, \frac{x_{ll}}{n}, \Phi\right) \\ /G(\Phi) \geq H(\Phi)/ \quad &\geq Q^1\left(\frac{x_{hh}}{n}, \frac{x_{hl}}{n}, \frac{x_{lh}}{n}, \frac{x_{ll}+1}{n}, \Phi\right), \end{aligned}$$

for any $x \in \mathcal{X}_{n-1}$. It follows that, if $\alpha_m \geq \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$,

$$\mathrm{E}\left[\rho^1(x)|hh\right] = \mathrm{E}\left[\rho^1(x)|hl\right] \geq \mathrm{E}\left[\rho^1(x)|lh\right] \geq \mathrm{E}\left[\rho^1(x)|ll\right]. \tag{33}$$

Moreover, Lemmas 3 and 6 imply:

Lemma 7 *In the solution to the relaxed problem (26),* $\eta_{hh} = \eta^h_m = 1$ *and*

1. *if* $\alpha_m > \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$, *then exactly one of the following is true:* $\eta^l_m = \eta_{ll} = 0$; $0 < \eta^l_m \leq 1$ *and* $\eta_{ll} = 0$; $\eta^l_m = 1$ *and* $0 \leq \eta_{ll} < 1$; *or* $\eta^l_m = \eta_{ll} = 1$.

2. *if* $\alpha_m = \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$, *then* $\eta^l_m = \eta_{ll}$.

3. *if* $\alpha_m < \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$, *then exactly one of the following is true:* $\eta^l_m = \eta_{ll} = 0$; $0 < \eta_{ll} \leq 1$ *and* $\eta^l_m = 0$; $\eta_{ll} = 1$ *and* $0 \leq \eta^l_m < 1$; *or* $\eta^l_m = \eta_{ll} = 1$.

Together with (33) and the fact that the incentive constraints (22) and (23) in the relaxed program (26) bind implies that Lemma 5 applies in the case when $\alpha_m \geq \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$.[17] We conclude:

[17] One way to see that the constraints (22) and (23) in the relaxed program (26) must bind is that, otherwise, the solution is ex post optimal, a contradiction to Proposition 3.

Lemma 8 *If* $\alpha_m \geq \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$, *then any solution to (26) satisfies all disregarded incentive constraints.*

The problem when $\alpha_m < \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$ is that if (13) and (17) bind and $\eta_m^l \mathrm{E}\left[\rho^1(x)|lh\right] < \eta_{ll}\mathrm{E}\left[\rho^1(x)|ll\right]$, then the solution to the relaxed problem violates the constraint (14) that type-hh does not have incentive to mis-report as type-ll. From Lemma 6, we know that if $\alpha_m < \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$, then $H(\Phi) > G(\Phi)$, which implies that η_{ll} may be higher than η_m^l. In a large economy with many agents, $\mathrm{E}\left[\rho^1(x)|lh\right]$ is approximately equal to $\mathrm{E}\left[\rho^1(x)|ll\right]$ (see Lemma A1 in Appendix A). Therefore, as long as the provision probability is bounded away from zero, the constraint (14) will be violated. The intuition for this is that, if the incentive constraint between hh and ll is ignored and there are few mixed types (when α_m is small), it creates an incentive to put the ll types first in line for the low valuation good. But, then it will be more appealing for type hh to mimic ll than to mimic the mixed type. We will therefore treat this case separately (see Section 4.7).

4.7 The Optimal Mechanism in a Large Economy

In this section, we provide a full characterization of the asymptotic properties of a sequence of optimal mechanisms. As a first step we will check when it is even possible to provide a non-negligible level of the public goods in a large economy. This is easier than characterizing the fully optimal mechanism first and then checking what the asymptotic provision probability is in the optimal mechanism. The result below thus saves the trouble of characterizing the optimal mechanism for parameter configurations when it is asymptotically impossible to provide the public good at all. We denote a mechanism in economy of size n as $\mathcal{M}_n = \left(\rho_n^1, \rho_n^2, \eta(n), t(n)\right)$, where $\eta(n) = \left(\eta_{hh}(n), \eta_m^h(n), \eta_m^l(n), \eta_{ll}(n)\right)$ and $t(n) = \left(t_{hh}(n), t_m(n), t_{ll}(n)\right)$. The result is:

Lemma 9 *Suppose that* $\max\{2\alpha h, (\alpha_{hh}+2\alpha_m)(h+l)\} < 2c$ *and let* $\{\mathcal{M}_n\}_{n=1}^{\infty}$ *be any sequence of incentive feasible mechanisms. Then,* $\lim_{n\to\infty} \mathrm{E}\rho_n^j(x) = 0$ *for* $j = 1, 2$.

The argument relies on that the probability that an agent is pivotal for the provision decision converges to zero as the number of agents goes out of bounds. The question of whether the feasibility constraint can be satisfied at a non-zero asymptotic provision probability can therefore be analyzed as if the goods are provided for sure. Making use of the downwards incentive constraints and the participation constraint, one shows that the maximal revenue from each type converge towards the revenue that could be collected if the goods are provided for sure. This is a monopoly pricing problem, and the revenue maximizing selling strategy is either to only sell to only the high types and collect $2\alpha h$, give both goods to everyone except type ll and collect $(\alpha_{hh}+2\alpha_m)(h+l)$, or to give access to everyone and collect $2l$. By assumption, neither $2\alpha h$, $(\alpha_{hh}+2\alpha_m)(h+l)$, or $2l$ is enough.[18]

[18] Recall the nontriviality assumption $l < c < h$. If $l \geq c$, first best, always providing and never excluding anyone, is implementable by charging $2c$ from each agent.

In the main case, when the valuations are not too positive correlated, the asymptotic properties of an optimal mechanism can be summarized as follows:

Proposition 4 *Assume that* $\alpha_m > \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$. *Let* $\{\mathcal{M}_n\}_{n=1}^{\infty}$ *be a sequence of optimal mechanism. Then, the following holds:*

1. *if* $\max\{2\alpha h, (\alpha_{hh}+2\alpha_m)(h+l)\} > 2c$, *then* $\lim_{n\to\infty} E\rho_n^j(x) \to 1$ *for* $j = 1,2$;
2. *if* $\max\{2\alpha h, (\alpha_{hh}+2\alpha_m)(h+l)\} < 2c$, *then* $\lim_{n\to\infty} \mathrm{E}\rho_n^j(x) \to 0$ *for* $j = 1,2$;
3. *if* $(\alpha_{hh}+2\alpha_m)(h+l) > 2c$, *then there exists* $N < \infty$ *such that* $\eta_m^h(n) = \eta_m^l(n) = 1$ *for every* $n \geq N$, *and*
$$\lim_{n\to\infty} \eta_{ll}(n) = \eta_{ll}^* = \frac{(\alpha_{hh}+2\alpha_m)(h+l)-2c}{(\alpha_{hh}+2\alpha_m)(h+l)-2l} \in (0,1).$$
4. *If* $2\alpha h > 2c > (\alpha_{hh}+2\alpha_m)(h+l)$,*then there exists* $N < \infty$ *such that* $\eta_{ll}(n) = 0$ *for all* $n \geq N$ *and*
$$\lim_{n\to\infty} \eta_m^l(n) = \eta_m^{l*} = \frac{2\alpha h - 2c}{2\alpha h - (\alpha_{hh}+2\alpha_m)(h+l)} \in (0,1).$$

In this case, allowing for bundling leads to strict gains in terms of economic efficiency. As we already know from the example in Section 3.2, there are cases when it is asymptotically infeasible to provide the public goods at all in the absence of bundling, but where bundling leads to almost sure implementation. In addition, the optimal bundling mechanism leads to strict efficiency gains relative the non-bundling regime because the probability of inclusion for low-valuation agents is increased, even in cases when the goods can be provided without bundling. To see this, suppose that $\alpha h > c$ so that both public goods will be asymptotically provided with probability one with or without bundling. From Proposition 3, we know that under the best separate provision mechanism, the probability for access to a low valuation agent is $(\alpha h - c)/(\alpha h - l)$. In contrast, Proposition 4 implies that the ex ante probability for access conditional on a low valuation for the case where $2c > (\alpha_{hh}+2\alpha_m)(h+l)$ is

$$\underbrace{\frac{\alpha_m}{\alpha_m+\alpha_{ll}}}_{\substack{\text{prob of mixed type}\\ \text{given low valuation}}} \underbrace{\frac{2\alpha h - 2c}{2\alpha h - (\alpha_{hh}+2\alpha_m)(h+l)}}_{\eta_m^{l*}}. \tag{34}$$

Some algebra shows that (34) is larger than $(\alpha h - c)/(\alpha h - l)$ whenever $\alpha_m > \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$, which is precisely the condition under which Proposition 4 is applicable. Fewer consumers are thus excluded in the optimal bundling mechanism. A similar calculation applies to the case where $2c < (\alpha_{hh}+2\alpha_m)(h+l)$.

Next, we consider the case when the valuations are sufficiently strongly correlated so that Proposition 4 does not apply. In this case it turns out that there are three binding incentive compatibility constraints at the optimum: the two constraints in program (26) and the downward incentive constraint between hh and ll. The asymptotic characterization for this case is given by the following proposition.

Proposition 5 *Assume that* $\alpha_m \leq \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$. *Let* $\{\mathcal{M}_n\}_{n=1}^{\infty}$ *be a sequence of optimal mechanism. Then, the following holds:*

1. *if* $\alpha h < c$, *then* $\lim_{n\to\infty} \mathrm{E}\rho_n^j(x) \to 0$ *for* $j = 1, 2$;

2. *if* $\alpha h > c$, *then* $\lim_{n\to\infty} \mathrm{E}\rho_n^j(x) \to 1$ *for* $j = 1, 2$, $\eta_{hh}(n) = \eta_m^h(n) = 1$ *and* $\eta_m^l(n) = \eta_{ll}(n)$ *for all* n, *where*
$$\lim_{n\to\infty} \eta_{ll}(n) = \frac{\alpha h - c}{\alpha h - l} \in (0, 1).$$

Hence, in this case the solution is asymptotically identical to the solution for the surplus maximization problem when bundling is not allowed (i.e. Proposition 3).

To understand why the bundling option does not change anything in this case, recall that asymptotic provision or non-provision is related to whether the maximal revenue for a monopolistic provider of the goods – if provided – exceeds the costs.[19] The revenue maximizing selling strategy for a monopolist, if both public goods are provided, is either to sell goods separately at price h (receiving an expected revenue of $2\alpha h$), or sell the goods as a bundle at price $h + l$ (receiving an expected revenue of $(\alpha_{hh} + 2\alpha_m)(h + l)$), or to charge l for each good (receiving an expected revenue of $2l$). It is still possible that the revenue maximizing selling strategy is to charge $h + l$ for the bundle in the current case where $\alpha_m \leq \alpha_{hh}\alpha_{ll}/(1 - \alpha_{ll})$. However, whenever $\alpha h < c$, we can show using the non-triviality assumption that $l < c$, that

$$\begin{aligned}
(\alpha_{hh} + 2\alpha_m)(h + l) &< (\alpha_{hh} + 2\alpha_m)(h + c) \\
&< (\alpha_{hh} + 2\alpha_m)\left(\frac{1}{\alpha} + 1\right)c \\
&= \left[2 + \frac{\alpha_m(1 - \alpha_{ll}) - \alpha_{hh}\alpha_{ll}}{\alpha_{hh} + \alpha_m}\right]c,
\end{aligned}$$

which is less than $2c$ when $\alpha_m \leq \alpha_{hh}\alpha_{ll}/(1 - \alpha_{ll})$. Thus, if it is impossible to balance the budget when the goods are provided separately, and if $\alpha_m \leq \alpha_{hh}\alpha_{ll}/(1 - \alpha_{ll})$, it it also impossible to balance the budget if the goods are bundled. Finally, if $(\alpha_{hh} + 2\alpha_m)(h + l) > 2\alpha h > 2c$, the public goods can be provided with probability one asymptotically both under the separate provision mechanisms (described in part 2 of Proposition 5), or under the bundling mechanism (described in part 3 of Proposition 4). However, the calculations immediately following Proposition 4 show that if $(\alpha_{hh} + 2\alpha_m)(h + l) < 2c$, the social surplus is actually smaller using the bundling mechanism.[20]

[19] If both goods are provided for sure, it does not matter whether the goods are private or public. The relevant revenue maximization problem that provides the condition for asymptotic provision versus non-provision is thus the same as in Armstrong and Rochet [4] and Bolton and Dewatripont [8].

[20] Bolton and Dewatripont [8] consider a setup that only differs from ours in that goods are rival and that they study revenue maximization. They conclude that (pure) bundling is optimal under a condition slightly different from

Bundling╲Exclusion	No Exclusion	Exclusion
No Bundling	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite [16])	$\mathrm{E}\rho_n^{j*} \to 0$, if $\alpha h < c$ $\mathrm{E}\rho_n^{j*} \to 1$, if $\alpha h > c$ (Norman [21])
Bundling Allowed	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite [16])	$\mathrm{E}\rho_n^{j*} \to 0$, if $\max\{2\alpha h, (\alpha_{hh}+2\alpha_m)(h+l)\} < 2c$; $\mathrm{E}\rho_n^{j*} \to 1$, if $\max\{2\alpha h, (\alpha_{hh}+2\alpha_m)(h+l)\} > 2c$ (This Paper)

Table 1: The Asymptotic Provision Probability under Different Bundling and Exclusion Possibilities.

Table 1 summarizes the asymptotic provision probabilities under different bundling and exclusion possibilities, and contrasts the results proved in this paper with those in the literature.[21]

5 Discussion

The anti-trust legislation on bundling is somewhat vague both in the US and elsewhere; bundling may be or may not be considered an illegal anti-competitive practice depending on the details of the case. Our simple model shows that bundling by a benevolent provider in general is desirable. Obviously, a for-profit monopoly provider may use the bundling instrument in a way that reduces the welfare of the consumers even when there is bundling in the solution to the welfare optimization problem. However, in the cases when the constrained optimum in the absence of bundling leads to asymptotic non-provision and bundling leads to provision, it follows immediately that requiring a for-profit monopolist to unbundle the goods would force the firm out of business and unambiguously make the consumers worse off. Indeed, this line of reasoning was an important part of the motivation in the decision by the Office of Fair Trading [22] in the UK on alleged anti-competitive mixed bundling by the British Sky Broadcasting Limited. In general, the spirit of our analysis should carry over to "natural monopoly" environments with falling average costs, where the concern raised by the results in our paper is that, even if goods are provided by a profit maximizer, bundling may actually make some products viable that would not be available in the absence of bundling.

We also note that it does not seem crucial for the analysis that all goods are non-rival. This suggests that an extension of the model that have both public and private goods may be interesting from a public finance perspective. Public provision of private goods is typically viewed as an inefficient instrument to achieve some redistributive objective in the public finance literature. However,

our condition, but, like in our analysis, the condition rules out too strong positive correlation. The reason for the difference is that we look at constrained efficiency.

[21] Mailath and Postlewaite [16] considers a single-dimensional problem without use exclusion. However, the probabilities of provision in a multidimensional setting can be bounded from above by a single-dimensional problem, where the valuation is the maximum of the individual good valuations.

the logic of our results suggests that it is entirely possible that some private goods are included in the government bundle in order to alleviate the free-riding problem in public good provision.

References

[1] Adams, W.J. and Janet L. Yellen. "Commodity Bundling and the Burden of Monopoly." *Quarterly Journal of Economics,* **90**, 1976, 475 –98.

[2] Armstrong, Mark. "Multiproduct Non-Linear Pricing," *Econometrica*, **64**, 1996, 51-76.

[3] Armstrong, Mark. "Price Discrimination by a Many-Product Firm," *Review of Economic Studies,* **66,** 1999, 151-168.

[4] Armstrong, Mark and Jean-Charles Rochet. "Multi-dimensional screening: A user's guide," *European Economic Review*, **43** (4), 1999, 959-979.

[5] Aumann, Robert. J. "Mixed and Behavioral Strategies in Infinite Extensive Games," *Annals of Mathematical Studies,* **52**, 1964, 627-650.

[6] Bakos, Yannis and Brynjolfsson, Erik. "Bundling Information Goods: Pricing, Profits and Efficiency," *Management Science*, **45**, 1999, 1613-1630.

[7] Bergstrom, Carl T, and Theodore C. Bergstrom. "The Costs and Benefits of Library Site Licenses to Academic Journals," *Proceedings of the National Academy of Sciences in the United States of America*, **101**, 2004, 897-902.

[8] Bolton, Patrick and Mathias Dewatripont. *Contract Theory*, MIT Press, January 2005

[9] Cramton, Peter, Robert Gibbons and Paul Klemperer, "Dissolving a Partnership Efficiently." *Econometrica*, **55**, 1987, 615-632.

[10] Cremer, Jacques and Richard P. McLean. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions", *Econometrica,* **56**, 1988, 1247-57.

[11] Dana, James D. "The Organization and Scope of Agents: Regulating Multiproduct Industries," *Journal of Economic Theory,* **59,** 1993, 288-309.

[12] Fang, Hanming and Peter Norman. "An Efficiency Rationale for Bundling of Public Goods." Cowles Foundation Discussion Paper No. 1441, 2003.

[13] Hellwig, Martin, F. "A Utilitarian Approach to the Provision and Pricing of Excludable Public Goods," *manuscript*, University of Mannheim, August 2003.

[14] Hellwig, Martin, F. "Public Good Provision with Many Participants," *Review of Economic Studies,* **70**, 2003, 589-614.

[15] Jackson, Matthew O. and Hugo F. Sonnenschein. "Overcoming Incentive Constraints," *manuscript*, California Institute of Technology, December 2003.

[16] Mailath, G. and A. Postlewaite. "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies,* **57,** 1990, 351-368.

[17] Manelli, Alejandro M. and Daniel R. Vincent. "Bundling as an Optimal Mechanism by a Multiple-Good Monopolist", *manuscript*, University of Maryland, October, 2002.

[18] McAfee, R. Preston, John McMillan and Michael D. Whinston. "Multiproduct Monopoly,Commodity Bundling, and Correlation of Values." *Quarterly Journal of Economics,* **104**, 1989, 371–84.

[19] Myerson, Roger. "Optimal Auction Design," *Mathematics of Operations Research,* **6**, 1981, 58-73.

[20] Myerson, R. and M.A. Satterthwaite. "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory,* **29,** 1983, 265-281.

[21] Norman, Peter. "Efficient Mechanisms for Public Goods with Use Exclusion." *Review of Economic Studies*, **71**, 2004, 1163-1188.

[22] Office of Fair Trading, *BSkyB decision dated 17 December 2002: rejection of applications under section 47 of the Competition Act 1998 by ITV Digital in Liquidation and NTL*, July 2003, London, UK. Available at:
`http://www.oft.gov.uk/business/competition+act/decisions/bskyb.htm`

A Appendix A: Proofs of Results in Section 4.

Proof of Lemma 1.

Proof. For each $x \in \mathcal{X}_n, j = 1, 2, \theta_i \in \Theta$ we have that $\rho^j(x) \in [0,1], \eta^j_{\theta_i} \in [0,1]$. Next, we note that if $t_{ll} < 0$ and all constraints are satisfied, then a deviation where taxes are changed from t to $t' = (t_{hh}, t_m, 0)$ and where inclusion and provision rules are unchanged will satisfy all constraints in the relaxed program (26). Similarly, if all constraints hold and $t_m < -l - h$ the deviation

$$t' = (t_{hh}, t_m, -l - h, \max(0, t_{ll}))$$

satisfies all constraints (in the relaxed program). A symmetric argument restricts $t_m \geq -h - l$. Finally, if $t_{hh} < -3h - l$, then a deviation to

$$t' = (-3h - l, \max(t_m, -l - h), \max(0, t_{ll}))$$

will leave all constraints satisfied. We conclude that there is a lower bound $\underline{t} > -\infty$ such that for any mechanism where $t_{\theta_i} < \underline{t}$ for some θ_i, there exists an alternative mechanism that supports the same allocation (and therefore generates the same surplus) where $t_{\theta_i} \geq \underline{t}$. Also, if $t_{\theta_i} > \bar{t} = 2h$ for some θ_i then at least one constraint in (26) must be violated. We therefore conclude that there is no loss in generality to restrict t_{θ_i} to be a number in $\left[\underline{t}, \bar{t}\right]$. All constraints and the objective function are linear in the choice variables and therefore continuous, so we conclude that the optimization problem has a compact feasible set and a continuous objective. It is easy to check that the feasible set is non-empty, which proves the claim by appeal to the Weierstrass Theorem. ■

Proof of Lemma 3.

Notation for optimality conditions to program (26). The proofs that follow make direct use of the Kuhn-Tucker conditions to the optimization problem (26). For easy reference, Table 2 summarizes our notation for the multipliers associated with each constraint.

Constraint	Multiplier
(22) Type hh IC	λ_{hh}
(23) Type hl (lh) IC	λ_m
(24) Type ll IR	λ_{ll}
(25) Feasibility	Λ
$\eta_{hh}, \eta^h_m, \eta^l_m, \eta_{ll} \geq 0$	$\gamma_{hh}, \gamma^h_m, \gamma^l_m, \gamma_{ll}$
$1 - \eta_{hh}, 1 - \eta^h_m, 1 - \eta^l_m, 1 - \eta_{ll} \geq 0$	$\phi_{hh}, \phi^h_m, \phi^l_m, \phi_{ll}$
$\rho^1(x) \geq 0$	$\gamma(x)$
$1 - \rho^1(x) \geq 0$	$\phi(x)$

Table 2: Notation of multipliers.

Proof. **[Step 1]** The Kuhn-Tucker optimality conditions for η_{hh} are given by

$$2\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{hh}h}{n}+2\lambda_{hh}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}+1,x_{hl},x_{lh},x_{ll})h+\gamma_{hh}-\phi_{hh} = 0$$
$$\gamma_{hh}\eta_{hh}=0,\phi_{hh}(1-\eta_{hh})=0,\gamma_{hh}\geq 0,\phi_{hh}\geq 0. \tag{A1}$$

All terms except $\gamma_{hh}-\phi_{hh}$ in the first order condition are strictly positive. Hence, $\phi_{hh}>0$, which requires that $\eta_{hh}=1$ for the complementary slackness constraint to be fulfilled.

[Step 2] The first order condition with respect to η_m^h reads

$$\begin{aligned}0=2\textstyle\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,h\frac{x_{hl}}{n}-\lambda_{hh}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,h\\ +\lambda_m\textstyle\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,h+\gamma_m^h-\phi_m^h\end{aligned} \tag{A2}$$

One checks that $\mathbf{a}_n(x)=\frac{n}{x_{hl}}\alpha_m\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})$ holds for any x such that $x_{hl}\geq 1$ by using the functional form of the multinomial. Hence

$$\begin{aligned}\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,h\frac{x_{hl}}{n} &= \sum_{x\in\mathcal{X}_n:x_{hl}\geq 1}\frac{n}{x_{hl}}\alpha_m\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,\rho^1(x)\,h\frac{x_{hl}}{n}\\ &= \alpha_m h\sum_{x\in\mathcal{X}_n:x_{hl}\geq 1}\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,\rho^1(x)\\ &= \alpha_m h\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,.\end{aligned} \tag{A3}$$

Let

$$\begin{aligned}\widehat{\gamma}_m^h &= \frac{\gamma_m^h}{\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})}\\ \widehat{\phi}_m^h &= \frac{\phi_m^h}{\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})}\end{aligned}\,. \tag{A4}$$

Substituting (A3) into (A2) and using Lemma 2, we obtain the condition

$$\begin{aligned}2\alpha_m h-\lambda_{hh}h+\lambda_m h+\widehat{\gamma}_m^h-\widehat{\phi}_m^h &= 2\alpha_m h-\alpha_{hh}\Lambda h+(\alpha_{hh}+2\alpha_m)\,\Lambda h+\widehat{\gamma}_m^h-\widehat{\phi}_m^h\\ &= 2\alpha_m h+2\alpha_m\Lambda h+\widehat{\gamma}_m^h-\widehat{\phi}_m^h=0.\end{aligned}$$

The "rescaled multipliers" are well-defined, weakly positive, and equal to zero if and only if the "original multiplier" is equal to zero. Since $2\alpha_m h+2\alpha_m\Lambda h>0$, we conclude that $\widehat{\phi}_m^h>0$. Hence $\eta_m^h=1$ by the complementarity slackness condition.

[Step 3] The first order condition for η_m^l is

$$\begin{aligned}0=2\textstyle\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{lh}}{n}l-\lambda_{hh}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\rho^1(x_{hh},x_{hl},x_{lh}+1,x_{ll})h\\ +\lambda_m\textstyle\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\rho^1(x_{hh},x_{hl},x_{lh}+1,x_{ll})l+\gamma_m^l-\phi_m^l\end{aligned} \tag{A5}$$

A calculation following the same steps as (A3) shows that

$$\textstyle\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{lh}}{n}l=\alpha_m l\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl},x_{lh}+1,x_{ll}). \tag{A6}$$

Substituting (A6) into (A5) and simplifying one obtains

$$\begin{aligned} 0 &= 2\alpha_m l - \lambda_{hh} h + \lambda_m l + \widehat{\gamma}_m^l - \widehat{\phi}_m^l = 2\alpha_m l - \alpha_{hh} h \Lambda + (\alpha_{hh} + 2\alpha_m)\Lambda l + \widehat{\gamma}_m^l - \widehat{\phi}_m^l \\ &= \alpha_m (1+\Lambda)\left[(1-\Phi)\, 2l + \Phi\left(\frac{\alpha_{hh}+2\alpha_m}{\alpha_m} l - \frac{\alpha_{hh}}{\alpha_m} h\right) + \frac{\widehat{\gamma}_m^l - \widehat{\phi}_m^l}{(1+\Lambda)\,\alpha_m}\right] \\ &= \alpha_m (1+\Lambda)\left[G(\Phi) + \frac{\widehat{\gamma}_m^l - \widehat{\phi}_m^l}{(1+\Lambda)\,\alpha_m}\right] \qquad \text{(A7)} \end{aligned}$$

where $\widehat{\gamma}_m^l$ and $\widehat{\phi}_m^l$ are defined like in (A4). We conclude that $G(\Phi) > 0$ implies that $\widehat{\phi}_m^l > 0$ and, by complementary slackness, $\eta_m^l = 1$. Symmetrically, $G(\Phi) < 0$ implies that $\widehat{\gamma}_m^l > 0$ and $\eta_m^l = 0$. If $G(\Phi) = 0$, then $\widehat{\gamma}_m^l = \widehat{\phi}_m^l = 0$, imposing no restrictions on η_m^l.

[Step 4] The first order condition for η_{ll} is

$$\begin{gathered} 0 = 2\textstyle\sum_{x\in\mathcal{X}_n : x_{ll}\geq 1} \mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{ll}}{n} - \lambda_m \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\,(h+l) \\ +\lambda_{ll}\textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1) 2l + \gamma_{ll} - \phi_{ll} \end{gathered}$$

Using the multinomial identity $\mathbf{a}_n(x) = \frac{n}{x_{ll}}\alpha_{ll}\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)$ and following the same steps as in Step 3, we can rewrite the first order condition as

$$\begin{aligned} 0 &= \alpha_{ll} 2l + \Lambda\left[2l - (\alpha_{hh}+2\alpha_m)(h+l)\right] + \widehat{\gamma}_{ll} - \widehat{\phi}_{ll} \\ &= \alpha_{ll}(1+\Lambda)\left\{\frac{1}{1+\Lambda} 2l + \frac{\Lambda}{1+\Lambda}\left[\frac{2}{\alpha_{ll}} l - \frac{\alpha_{hh}+2\alpha_m}{\alpha_{ll}}(h+l)\right] + \frac{\widehat{\gamma}_{ll} - \widehat{\phi}_{ll}}{\alpha_{ll}(1+\Lambda)}\right\} \\ &= \alpha_{ll}(1+\Lambda)\left[H(\Phi) + \frac{\widehat{\gamma}_{ll} - \widehat{\phi}_{ll}}{\alpha_{ll}(1+\Lambda)}\right]. \end{aligned}$$

where $\widehat{\gamma}_{ll}$ and $\widehat{\phi}_{ll}$ are defined like in (A4). Arguing as in Step 3 completes the proof. ■

Proof of Lemma 4.

Proof. The first order condition with respect to $\rho^1(x)$ is

$$\begin{aligned} &2\mathbf{a}_n(x)\left[\tfrac{(\eta_{hh}x_{hh} + \eta_m^h x_{hl})h + (\eta_m^l x_{lh} + \eta_{ll} x_{ll})l}{n} - c\right] + \lambda_{hh}\left[2\eta_{hh}\mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll})\, h\right] \\ &-\lambda_{hh}\left[\eta_m^h \mathbf{a}_{n-1}(x_{hh}, x_{hl}-1, x_{lh}, x_{ll})\, h - \eta_m^h \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\, h\right] \\ &+\lambda_m\left[\eta_m^h \mathbf{a}_{n-1}(x_{hh}, x_{hl}-1, x_{lh}, x_{ll})\, h + \eta_m^l \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\, l\right] \qquad \text{(A8)} \\ &-\lambda_m\left[\eta_{ll}\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)(h+l)\right] + \lambda_{ll} 2\eta_{ll}^1 \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)\, l \\ &-\Lambda\mathbf{a}_n(x)\, 2c + \gamma(x) - \phi(x) = 0, \end{aligned}$$

where the convention is that $\mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll}) = 0$ if $x_{hh} = 0$, and so on. Using the following identities between multinomials,

$$\begin{array}{ll} \mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll}) = \frac{\mathbf{a}_n(x)}{\alpha_{hh}}\frac{x_{hh}}{n}, & \mathbf{a}_{n-1}(x_{hh}, x_{hl}-1, x_{lh}, x_{ll}) = \frac{\mathbf{a}_n(x)}{\alpha_m}\frac{x_{hl}}{n} \\ \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll}) = \frac{\mathbf{a}_n(x)}{\alpha_m}\frac{x_{lh}}{n}, & \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1) = \frac{\mathbf{a}_n(x)}{\alpha_{ll}}\frac{x_{ll}}{n}, \end{array} \qquad \text{(A9)}$$

exploiting the relationships between multipliers in Lemma 2, and substituting $\eta_{hh} = \eta_m^h = 1$ due to Lemma 3, we can simplify (A8) to

$$\begin{aligned}
&2\left[\tfrac{(x_{hh}+x_{hl})h+\left(\eta_m^l x_{lh}+\eta_{ll}x_{ll}\right)l}{n} - c\right](1-\Phi) \\
&+\alpha_{hh}\Phi\left[2\tfrac{1}{\alpha_{hh}}\tfrac{x_{hh}}{n}h - \tfrac{1}{\alpha_m}\tfrac{x_{hl}}{n}h - \eta_m^l\tfrac{1}{\alpha_m}\tfrac{x_{lh}}{n}h\right] \\
&+(\alpha_{hh}+2\alpha_m)\Phi\left[\tfrac{1}{\alpha_m}\tfrac{x_{hl}}{n}h + \eta_m^l\tfrac{1}{\alpha_m}\tfrac{x_{lh}}{n}l - \eta_{ll}\tfrac{1}{\alpha_{ll}}\tfrac{x_{ll}}{n}(h+l)\right] \\
&+\Phi 2\eta_{ll}^1\tfrac{1}{\alpha_{ll}}\tfrac{x_{ll}}{n}l - \Phi 2c + \tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)} = 0,
\end{aligned} \tag{A10}$$

where $\Phi = \Lambda/(1+\Lambda)$. This condition can be interpreted as a weighted average of surplus (the term multiplied by $1-\Phi$) and profit maximization (the terms multiplied by Φ). Collecting terms in (A10) and simplifying we get

$$\begin{aligned}
&2\tfrac{x_{hh}}{n}h + 2\tfrac{x_{hl}}{n}h - 2c + \tfrac{x_{lh}}{n}\eta_m^l\overbrace{\left\{(1-\Phi)2l + \Phi\left[\frac{\alpha_{hh}+2\alpha_m}{\alpha_m}l - \frac{\alpha_{hh}}{\alpha_m}h\right]\right\}}^{G(\Phi)} \\
&+\tfrac{x_{ll}}{n}\eta_{ll}\overbrace{\left\{(1-\Phi)2l + \Phi\left[\frac{2}{\alpha_{ll}}l - \frac{\alpha_{hh}+2\alpha_m}{\alpha_{ll}}(h+l)\right]\right\}}^{H(\Phi)} + \tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)} \\
&/(31)\ / = 2Q^1\left(\tfrac{x}{n},\Phi\right) + \tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)} = 0,
\end{aligned} \tag{A11}$$

where the equality uses the fact (from Lemma 3) that $\eta_m^l = 0$ if $G(\Phi) < 0$ and $\eta_{ll} = 0$ if $H(\Phi) < 0$. The result follows. ■

Proof of Lemma 5.

Proof. Define $A_{hh} = \eta_{hh}\mathrm{E}\left[\rho^1(x)\,|hh\right]$, $A_m^h = \eta_m^h\mathrm{E}\left[\rho^1(x)\,|hl\right]$, $A_m^l = \eta_m^l\mathrm{E}\left[\rho^1(x)\,|lh\right]$ and $A_{ll} = \eta_{ll}\mathrm{E}\left[\rho^1(x)\,|ll\right]$. The hypothesis that (13) and (17) bind can then be restated as

$$A_{hh}2h - t_{hh} = A_m^h h + A_m^l h - t_m \tag{A12}$$

$$A_m^h h + A_m^l l - t_m = A_{ll}(h+l) - t_{ll}. \tag{A13}$$

Constraint (14): if hh announces ll, the payoff is

$$\begin{aligned}
\underbrace{A_{ll}2h - t_{ll}}_{hh\text{ announces }ll} &= A_{ll}(h+l) - t_{ll} + A_{ll}(h-l)\,/(\text{A13})/ = A_m^h h + A_m^l l - t_m + A_{ll}(h-l) \\
&= A_m^h h + A_m^l h - t_m + \left(A_{ll} - A_m^l\right)(h-l)\Big/A_{ll} \le A_m^l\Big/ \\
&\le A_m^h h + A_m^l h - t_m\ /\ (\text{A12})/ = \underbrace{A_{hh}2h - t_{hh}}_{\text{truth-telling}}.
\end{aligned} \tag{A14}$$

Constraint (15): if lh or hl announces hh, the payoff is

$$\begin{aligned}
\underbrace{A_{hh}(h+l) - t_{hh}}_{\text{mixed type announces }hh} &= A_{hh}2h - t_{hh} + A_{hh}(l-h)\,/(\text{A12})/ = A_m^h h + A_m^l h - t_m + A_{hh}(l-h) \\
&= A_m^h h + A_m^l l - t_m + \left(A_m^l - A_{hh}\right)(h-l)\Big/A_m^l \le A_{hh}\Big/ \le \underbrace{A_m^h h + A_m^l l - t_m}_{\text{truth-telling}}.
\end{aligned}$$

Constraint (16): follows trivially since $A_m^h \geq A_m^l$.

Constraint (18): if ll announces lh or hl, the payoff is

$$\begin{aligned} \underbrace{A_m^h l + A_m^l l - t_m}_{ll \text{ announces mixed type}} &= A_m^h h + A_m^l l - t_m - (h-l) A_m^l \,/(\text{A13})/ = A_{ll}(h+l) - t_{ll} - (h-l) A_m^l \\ &= A_{ll} 2l - t_{ll} + (h-l)\left(A_{ll} - A_m^l\right) \Big/ A_{ll} \leq A_m^l \Big/ \leq \underbrace{A_{ll}2l - t_{ll}}_{\text{truth-telling}}. \end{aligned} \tag{A15}$$

Constraint (19): if ll announces hh, the payoff is

$$\begin{aligned} \underbrace{A_{hh}2l - t_{hh}}_{ll \text{ announces } hh} &= A_{hh}2h - t_{hh} - A_{hh}2(h-l) \,/(\text{A12})/ = A_m^h h + A_m^l h - t_m - A_{hh}2(h-l) \\ &= A_m^h h + A_m^l l - t_m + A_m^l (h-l) - A_{hh} 2(h-l) \,/(\text{A13})/ \\ &= A_{ll}(h+l) - t_{ll} + A_m^l (h-l) - A_{hh}2(h-l) \qquad\qquad \text{(A16)} \\ &= A_{ll}2l - t_{ll} + \left(A_m^l + A_{ll} - 2A_{hh}\right)(h-l) \leq \underbrace{A_{ll}2l - t_{ll}}_{\text{truth-telling}}. \end{aligned}$$

■

Proof of Lemma 6.

Proof. If $\Phi = 0$, then $G(\Phi) = H(\Phi)$. For $\Phi > 0$, $G(\Phi) - H(\Phi)$ has the same sign as

$$\begin{aligned} \Delta &= \frac{\alpha_{hh} + 2\alpha_m}{\alpha_m} l - \frac{\alpha_{hh}}{\alpha_m} h - \left[\frac{2}{\alpha_{ll}} l - \frac{\alpha_{hh} + 2\alpha_m}{\alpha_{ll}}(h+l)\right] \\ &= -\frac{\alpha_{hh}}{\alpha_m}(h-l) + 2l - \left[\frac{2}{\alpha_{ll}} l - \frac{(1-\alpha_{ll})}{\alpha_{ll}}(h+l)\right] \\ \Big/ \alpha_m \geq \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}} \Big/ \quad &\geq -\frac{(1-\alpha_{ll})}{\alpha_{ll}}(h-l) + 2l - \left[\frac{2}{\alpha_{ll}} l - \frac{(1-\alpha_{ll})}{\alpha_{ll}}(h+l)\right] \\ &= 2l\left(1 - \frac{1}{\alpha_{ll}} + \frac{(1-\alpha_{ll})}{\alpha_{ll}}\right) = 0, \end{aligned}$$

where the inequality on the third line is strict if $\alpha_m > \alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$. ■

Proof of Lemma 9.

The proof of Lemma 9 uses the following fact:

Lemma A1 *For every $\epsilon > 0$, there exists N such that $\left|\mathrm{E}\rho_n^1(x) - \mathrm{E}\left[\rho_n^1(x)|\theta_i\right]\right| \leq \epsilon$ for all $\theta_i \in \Theta$ and $n \geq N$.*

The proof of this result, which is a restatement of the paradox of voting, is omitted (see Fang and Norman [12] for details). The implication is that, if n is large, the perceived provision probability is almost independent of her announcement. We use this below in order to be able to separate the maximal revenue that can be raised per "unit of probability of provision" from the provision rule.

Proof of Lemma 9 (Continued):

For contradiction (taking a subsequence if necessary) suppose that $\lim_{n\to\infty} \mathrm{E}\rho_n^1(x) = \rho^* > 0$. Let $\delta \equiv 2c - \max\{2\alpha h, (\alpha_{hh} + 2\alpha_m)(h+l), 2l\} > 0$ (recall that $l < c < h$). Denote the inclusion probabilities in mechanism $\mathcal{M}_n$ by $\eta \equiv \left(\eta_{hh}(n), \eta_m^h(n), \eta_m^l(n), \eta_{ll}(n)\right)$ and let $A(n) \equiv \left(A_{hh}(n), A_m^h(n), A_m^l(n), A_{ll}(n)\right)$ be defined as in Lemma 5. Rearrange (13), (14), (17), and (24) as

$$\begin{aligned} t_{hh} &\le t_m + A_{hh}(n)\,2h - A_m^h(n)\,h - A_m^l(n)\,h \\ t_{hh} &\le t_{ll} + A_{hh}(n)\,2h - A_{ll}(n)\,2h \\ t_m &\le t_{ll} + A_m^h(n)\,h + A_m^l(n)\,l - A_{ll}(n)(h+l) \\ t_{ll} &\le A_{ll}(n)\,2l \end{aligned} \tag{A17}$$

Purely for expositional simplicity, define two functions $B : [0,1]^4 \to R$ and $T : [0,1]^3 \to R$ as follows:

$$\begin{aligned} B(z_1, z_2, z_3, z_4) &\equiv 2lz_4 + (\alpha_{hh} + 2\alpha_m)[z_2 h + z_3 l - z_4(h+l)] + \alpha_{hh}[2z_1 h - z_2 h - z_3 h], \\ T(t_{hh}, t_m, t_{ll}) &= a_{hh}t_{hh} + 2\alpha_m t_m + \alpha_{ll}t_{ll}. \end{aligned} \tag{A18}$$

Combining all but the second of the inequalities in (A17), one shows that $T(t(n)) \le B(A(n))$, and from all but the first inequality it follows that

$$\begin{aligned} T(t(n)) &\le A_{ll}(n)\,2l + 2\alpha_m\left[A_m^h(n)\,h + A_m^l(n)\,l - A_{ll}(n)(h+l)\right] + \alpha_{hh}[A_{hh}(n)\,2h - A_{ll}(n)\,2h] \\ &= B(A(n)) + \alpha_{hh}\left[A_m^l(n) - A_{ll}(n)\right](h-l) \end{aligned} \tag{A19}$$

Let $\varepsilon = \frac{\rho^*\delta}{2c} > 0$. From Lemma A1 and the hypothesis that $\lim_{n\to\infty} \mathrm{E}\rho_n^1(x) = \rho^* > 0$, we know that there exists N such that $\mathrm{E}\rho_n^1(x) > \frac{1}{2}\rho^*$, $A_{hh}(n) = \mathrm{E}\left[\rho_n^1(x)\,|hh\right]\eta_{hh}(n) \le \left[\mathrm{E}\rho_n^1(x) + \varepsilon\right]\eta_{hh}(n)$, $A_m^h(n) = \mathrm{E}\left[\rho_n^1(x)\,|hl\right]\eta_m^h(n) \le \left[\mathrm{E}\rho_n^1(x) + \varepsilon\right]\eta_m^h(n)$, $A_m^l(n) = \mathrm{E}\left[\rho_n^1(x)\,|lh\right]\eta_m^l(n) \le \left[\mathrm{E}\rho_n^1(x) + \varepsilon\right]\eta_m^l(n)$, and $A_{ll}(n) = \mathrm{E}\left[\rho_n^1(x)\,|ll\right]\eta_{ll}(n) \le \left[\mathrm{E}\rho_n^1(x) + \varepsilon\right]\eta_{ll}(n)$ for every $n \ge N$. Hence, there exists N such that

$$T(t(n)) \le \left[\mathrm{E}\rho_n^1(x) + \varepsilon\right]\min\left\{B(\eta(n)), B(\eta(n)) + \alpha_{hh}\left[\eta_m^l(n) - \eta_{ll}(n)\right](h-l)\right\} \tag{A20}$$

Define

$$V = \max_{\eta\in[0,1]^4} \min\left\{B(\eta), B(\eta) + \alpha_{hh}\left(\eta_m^l - \eta_{ll}\right)(h-l)\right\}. \tag{A21}$$

Note that $B(\eta)$ and $B(\eta) + \alpha_{hh}\left(\eta_m^l - \eta_{ll}\right)(h-l)\,\eta_{hh}$ are both increasing in η_{hh} and η_m^h, so the solution to the maximization problem in (A21) must have $\eta_{hh} = \eta_m^h = 1$. The only question is thus the inclusion probabilities η_m^l and η_{ll}:

Possibility 1: Suppose that $\eta_m^l > \eta_{ll}$ in a solution to (A21). Since B is linear in η it follows that we may assume that $\eta_{ll} = 0$ and that $\eta_m^l = 1$ solves $\max_{\eta_{ll},\eta_m^l\in[0,1]^2} B\left(1, 1, \eta_m^l, \eta_{ll}\right)$ and that $V = B(1,1,1,0) = (\alpha_{hh} + 2\alpha_m)(h+l)$

Possibility 2: Suppose $\eta_m^l = \eta_{ll}$ solves (A21). Then, $\eta_m^l = \eta_{ll} \in \arg\max_{x\in[0,1]} B(1,1,x,x)$. Evaluating, we see that $B(1,1,x,x) = x2l + 2(\alpha_{hh} + \alpha_m)\,h(1-x)$. Hence, either $\eta_{ll} = \eta_m^l = 0$ and $V = 2\alpha h$, or $\eta_{ll} = \eta_m^l = 1$ with $V = 2l$.

Possibility 3: Suppose that $\eta_m^l < \eta_{ll}$ in a solution to (A21). We can then without loss assume that $\eta_m^l = 0$ and that $\eta_{ll} = 1$ solves $\max_{\eta_{ll}} \eta_{ll} 2l + 2\alpha_m [h - \eta_{ll}(h+l)] + \alpha_{hh}[\eta_{hh} 2h - \eta_{ll} 2h]$. The associated value of the maximand is then $2l(1-\alpha_m) < 2l$, therefore this could not be a solution to the problem.

To summarize, $V = \max\{2\alpha h, (\alpha_{hh} + 2\alpha_m)(h+l), 2l\}$ which is equal to $2c - \delta$. Combining with (A20), it follows that

$$\begin{aligned} T(t(n)) - 2\mathrm{E}\rho_n^1(x)c &= a_{hh} t_{hh}(n) + 2\alpha_m t_m(n) + \alpha_{ll} t_{ll}(n) - 2\mathrm{E}\rho_n^1(x)c \\ &\leq \left[\mathrm{E}\rho_n^1(x) + \varepsilon\right](2c-\delta) - 2\mathrm{E}\rho_n^1(x)c = -\mathrm{E}\rho_n^1(x)\delta + \varepsilon(2c-\delta) \\ &< -\frac{\rho^*\delta}{2} + 2c\varepsilon = 0, \end{aligned} \quad \text{(A22)}$$

implying that the (25) is violated when n is sufficiently large. Hence, $\lim_{n\to\infty} \mathrm{E}\rho_n^1(x) = 0$. ■

Proof of Proposition 4.

The next Lemma is used in the proof of Proposition 4:

Lemma A2 *For any $\epsilon > 0$ there exists N such that* $\Pr\left(\left|Q^1\left(\frac{x}{n}, \Phi_n\right) - Q^1(\mu, \Phi_n)\right| \geq \epsilon\right) \leq \epsilon$ *for every $n \geq N$.*

Proof. We notice that $Q^1\left(\frac{x}{n}, \Phi_n\right) = \sum_{i=1}^n \frac{y_i(\theta_i;\Phi_n)}{n}$, where $y_i(\theta_i;\Phi_n)$ is given by

$$y_i(\theta_i;\Phi_n) = \begin{cases} h - c & \text{if } \theta_i \in \{hh, hl\} \\ \max\{0, G(\Phi_n)\} - c & \text{if } \theta_i = lh \\ \max\{0, H(\Phi_n)\} - c & \text{if } \theta_i = ll. \end{cases} \quad \text{(A23)}$$

Since $y_i(\theta_i;\Phi_n)$ has bounded support, there exists $\sigma^2 < \infty$ such that the variance of $Y_i(\theta_i;\Phi_n)$ is less than σ^2 for any $\Phi_n \in [0,1]$. Moreover, $\{Y(\theta_i;\Phi_n)\}_{i=1}^n$ is a sequence of i.i.d. random variables and $\mathrm{E}_{\theta_i} y_i(\theta_i;\Phi_n) = Q^1(\mu, \Phi_n)$. Hence, Chebyshev's inequality is applicable, which implies that for any $\varepsilon > 0$,

$$\begin{aligned} \Pr\left(\left|Q^1\left(\frac{x}{n}, \Phi_n\right) - Q^1(\mu, \Phi_n)\right| \geq \epsilon\right) &= \Pr\left(\left|\sum_{i=1}^n \frac{y_i(\theta_i;\Phi_n)}{n} - \mathrm{E}_{\theta_i} y_i(\theta_i;\Phi_n)\right| \geq \epsilon\right) \\ &\leq \frac{\mathrm{Var}[Y_i(\theta_i;\Phi_n)]}{n\epsilon^2} \leq \frac{\sigma^2}{n\epsilon^2}. \end{aligned} \quad \text{(A24)}$$

Hence, $\Pr\left(\left|Q^1\left(\frac{x}{n}, \Phi_n\right) - Q^1(\mu, \Phi_n)\right| \geq \epsilon\right) \leq \epsilon$ for all $n \geq N = \sigma^2/\epsilon^3 < \infty$. ■

Proof of Proposition 4 (Continued).

(Part 1) From (32), $Q^1(\mu, \Phi_n) \geq (\alpha_{hh} + \alpha_m)h - c = \alpha h - c$ for any $\Phi_n \in [0,1]$ (recall $\alpha \equiv \alpha_{hh} + \alpha_m$), hence $\lim_{n\to\infty} Q^1(\mu, \Phi_n) \geq \alpha h - c$. If $\alpha h > c$, part 1 is therefore immediate from

Lemmas 4 and A2. Suppose instead that $(\alpha_{hh} + 2\alpha_m)(h + l) > 2c \geq 2\alpha h$. Then, for any $\Phi_n \in [0, 1]$

$$\begin{aligned} Q^1(\mu, \Phi_n) &= \alpha h + \alpha_m \frac{\max\{0, G(\Phi_n)\}}{2} + \alpha_{ll} \frac{\max\{0, H(\Phi_n)\}}{2} - c \geq \alpha h - c + \frac{\alpha_m G(\Phi_n)}{2} \\ &= (1 - \Phi_n)[\alpha h + \alpha_m l - c] + \Phi_n \left[\frac{(\alpha_{hh} + 2\alpha_m)(h + l)}{2} - c \right] \\ &> \left[\frac{(\alpha_{hh} + 2\alpha_m)(h + l)}{2} - c \right] > 0, \end{aligned}$$

where the first inequality comes from the fact that $\alpha h + \alpha_m l > \frac{(\alpha_{hh}+2\alpha_m)(h+l)}{2}$. By Lemma A2 (let $\varepsilon = Q^1(\mu, \Phi_n)$ and note that $\Pr[Q^1 \leq 0] \leq \Pr[|Q^1 - EQ^1| \leq EQ^1]$), it follows that $\lim_{n\to\infty} \Pr[Q^1(\frac{x}{n}, \Phi^n) \leq 0] = 0$. Appealing to Lemma 4 completes the proof.

(Part 2) Immediate from Lemma 9.

(Part 3) Consider the sequence of mechanisms $\left\{\widehat{\mathcal{M}}_n\right\}$, where for each n

$$\begin{aligned} \widehat{\rho}_n^1(x) &= 1 \text{ for all } x \in \mathcal{X}_n \\ \widehat{\eta}_{hh}(n) &= \widehat{\eta}_m^h(n) = \widehat{\eta}_m^l(n) = 1, \widehat{\eta}_{ll}(n) = \eta_{ll}^* \equiv \frac{(\alpha + \alpha_m)(h + l) - 2c}{(\alpha + \alpha_m)(h + l) - 2l} \\ \widehat{t}_{hh}(n) &= \widehat{t}_m(n) = (h + l) - \eta_{ll}^*(h - l) \\ \widehat{t}_{ll}(n) &= \eta_{ll}^* 2l. \end{aligned}$$

It is easy to check that (22),(23), (24), and (25) are all satisfied with equality. Moreover, Lemma 5 applies, so all other incentive constraints are also satisfied. Hence, $\widehat{\mathcal{M}}_n$ is incentive feasible for any n. The associated expected per capita surplus is $s\left(\widehat{\mathcal{M}}_n\right) = 2(\alpha h + \alpha_m l + \alpha_{ll}\eta_{ll}^* - c)$. Now, suppose for contradiction that there does not exist N such that $\widehat{\eta}_m^l(n) = 1$ for all $n \geq N$. Then (taking a subsequence if necessary) $\widehat{\eta}_m^l(n) < 1$ for all n, which (Lemma 7) implies that $\eta_{ll}(n) = 0$ for all n (in the subsequence). The per capita surplus generated by the optimal mechanism $\mathcal{M}_n$ in the n^{th} economy in the sequence, denoted $s(\mathcal{M}_n)$, is then

$$s(\mathcal{M}_n) = \frac{2\mathrm{E}\rho_n^1(x)\left[(x_{hh} + x_{hl})h + \left(\eta_m^l(n)x\right)l - cn\right]}{n} \leq 2(\alpha h + \alpha_m l) - 2\mathrm{E}\rho_n^1(x)c$$

Since $\mathrm{E}\rho_n^1(x) \to 1$ as $n \to \infty$ (Part 1) it follows that for every $\varepsilon > 0$ there exists N such that $s(\mathcal{M}_n) \leq 2(\alpha h + \alpha_m l - c) + \varepsilon < s\left(\widehat{\mathcal{M}}_n\right)$ for ε sufficiently small. Hence we have arrived at a contradiction, implying that there does exist N such that $\widehat{\eta}_m^l(n) = 1$ for all $n \geq N$. It remains to be shown that $\eta_{ll}(n) \to \eta_{ll}^*$ for any sequence of optimal mechanisms. First assume that there is a subsequence such that $\eta_{ll}(n) \to \eta' < \eta_{ll}^*$. Arguing just like above one shows that for every $\varepsilon > 0$, there exists $N < \infty$ such that $s(\mathcal{M}_n) \leq 2(\alpha h + \alpha_m l + \alpha_{ll}\eta' l - c) + \varepsilon$. Hence, for $\varepsilon = \alpha_{ll}(\eta_{ll}^* - \eta')l > 0$ we have that

$$s\left(\widehat{\mathcal{M}}_n\right) - s(\mathcal{M}_n) \geq 2\alpha_{ll}\left(\eta_{ll}^* - \eta'\right)l - \varepsilon = \alpha_{ll}\left(\eta_{ll}^* - \eta'\right)l > 0,$$

again contradicting the optimality sequence $\{\mathcal{M}_n\}$ for n is sufficiently large. Finally, suppose there is a subsequence such that $\eta_{ll}(n) \to \eta' > \eta_{ll}^*$. We will now show that this implies that, for some sufficiently large n, the mechanisms in the sequence are not incentive feasible. Define

$$\varepsilon = \left[\frac{(\alpha_{hh} + 2\alpha_m)(h+l)}{2} - l \right] \frac{(\eta' - \eta_{ll}^*)}{(4h+1)c} > 0$$

Applying (A20), it follows that there exists N_1 such that the per capita revenue in mechanism $\mathcal{M}_n$ satisfies

$$T(t(n)) \leq \left[\mathrm{E}\rho_n^1(x) + \varepsilon 2h \right] B(1,1,1,\eta_{ll}(n)) \leq (1 + \varepsilon 2h) B(1,1,1,\eta_{ll}(n))$$

for all $n \geq N$, where, after some algebra

$$B(1,1,1,\eta_{ll}(n)) \equiv \eta_{ll}(n) 2l + (1 - \eta_{ll}(n))(\alpha_{hh} + 2\alpha_m)(h+l).$$

Moreover, since $l < c < \frac{(\alpha_{hh}+2\alpha_m)(h+l)}{2}$,we also have that $B(1,1,1,\eta_{ll}(n))$ is decreasing in $\eta_{ll}(n)$. Let $\delta^1 \equiv \eta' - \eta_{ll}^* > 0$ and $\delta^2 = \frac{(\alpha_{hh}+2\alpha_m)(h+l)}{2} - l > 0$. Given that $\eta_{ll}(n) \to \eta' > \eta_{ll}^*$ there exists N_2 such that $\eta_{ll}(n) \geq \eta_{ll}^* + \frac{\delta^1}{2}$, implying that

$$\begin{aligned} B(1,1,1,\eta_{ll}(n)) &\leq B\left(1,1,1,\eta_{ll}^* + \frac{\delta^1}{2}\right) \\ &= \underbrace{\eta_{ll}^* 2l + (1 - \eta_{ll}^*)(\alpha_{hh} + 2\alpha_m)(h+l)}_{=B(1,1,1,\eta_{ll}^*)=2c} + \delta^1 \Bigg(\underbrace{l - \frac{(\alpha_{hh} + 2\alpha_m)(h+l)}{2}}_{=-\delta^2} \Bigg) \end{aligned} \tag{A25}$$

and since $\mathrm{E}\rho_n^1(x) \to 1$ we have that there exists N_3 such that the expected per capita costs satisfies $2\mathrm{E}\rho_n^1(x)c \geq (2 - \varepsilon)c$. Hence, for any $n \geq \max\{N_1, N_2, N_3\}$ it follows that

$$\begin{aligned} T(t(n)) - \mathrm{E}\rho_n^1(x)c &= a_{hh}t_{hh}(n) + 2\alpha_m t_m(n) + \alpha_{ll}t_{ll}(n) - \mathrm{E}\rho_n^1(x)c \\ &\leq (1 + \varepsilon 2h)\left[2c - \delta^1\delta^2\right] - (2 - \varepsilon)c \\ &= -\delta^1\delta^2 + \varepsilon\left[(4h+1)c - 2h\delta^1\delta^2\right] \\ &< -\delta^1\delta^2 + \varepsilon\left[(4h+1)c\right] \\ &= \left[l - \frac{(\alpha_{hh} + 2\alpha_m)(h+l)}{2} \right] \frac{(\eta' - \eta_{ll}^*)}{(4h+1)c} + \varepsilon = 0. \end{aligned} \tag{A26}$$

We conclude that the feasibility constraint (25) is violated for n large enough, contradicting the hypothesis that there is a subsequence such that $\eta_{ll}(n) \to \eta' > \eta_{ll}^*$. Since there must be some convergent subsequence and since no subsequence can converge to anything else than η_{ll}^* we conclude that $\eta_{ll}(n) \to \eta_{ll}^*$ as n goes out of bounds.

(Part 4) The proof of this proceeds along the same steps as in Part 3 and is omitted. The only difference is that one in this case begins by arguing that if there is no N such that $\eta_{ll}(n) = 0$ for all $n \geq N$ then the mechanism eventually violates incentive feasibility. To establish that $\eta_m^l(n) \to \frac{2\alpha h - 2c}{2\alpha h - (\alpha_{hh}+2\alpha_m)(h+l)}$ proceeds along the same lines as in Part 3. One checks that a sequence of mechanisms $\left\{\widehat{\mathcal{M}}_n\right\}$where both goods are provided for sure and where $\widehat{\eta}_m^l(n) = \frac{2\alpha h - 2c}{2\alpha h - (\alpha_{hh}+2\alpha_m)(h+l)}$ and $\widehat{\eta}_{ll}(n) = 0$ for each n is incentive feasible. Then, one argues that any sequence of mechanisms where $\eta_m^l(n)$ is bounded away from $\frac{2\alpha h-2c}{2\alpha h-(\alpha_{hh}+2\alpha_m)(h+l)}$ from below is eventually dominated by $\widehat{\mathcal{M}}_n$, and that any sequence of mechanisms where $\eta_m^l(n)$ is bounded away from $\frac{2\alpha h-2c}{2\alpha h-(\alpha_{hh}+2\alpha_m)(h+l)}$ from above is eventually violating the feasibility constraint. ∎

<u>*Proof of Proposition 5.*</u>

To prove Proposition 5, we need to add the constraint (14) to the relaxed program. In terms of the binomial probability distribution $\mathbf{a}_{n-1}$, (14) reads,

$$2\eta_{hh}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}+1,x_{hl},x_{lh},x_{ll}+1)h-t_{hh}\geq 2\eta_{ll}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl},x_{lh},x_{ll}+1)h-t_{ll}. \tag{A27}$$

The relevant programming problem is to thus to solve,

$$\max_{\{\rho^1,\eta,t\}} 2\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\left[\frac{\left(\eta_{hh}x_{hh}+\eta_m^h x_{hl}\right)h+\left(\eta_m^l x_{lh}+\eta_{ll}x_{ll}\right)l}{n}-c\right] \tag{A28}$$

s.t. (22), (23), (24), (25) and (A27),

$$\eta\in[0,1]^4,\rho^1(x)\geq 0,1-\rho^1(x)\geq 0 \text{ for each } x\in\mathcal{X}_n. \tag{A29}$$

We let ψ_{hh} denote the multiplier on the new constraint (A27) and keep the rest of the notation same as before. The first order conditions with respect to t are now

$$\begin{array}{ll}(\text{w.r.t. } t_{hh}) & -\lambda_{hh}-\psi_{hh}+\Lambda\alpha_{hh}=0,\\ (\text{w.r.t. } t_m) & \lambda_{hh}-\lambda_m+2\Lambda\alpha_m=0,\ ,\\ (\text{w.r.t. } t_{ll}) & \lambda_m+\psi_{hh}-\lambda_{ll}+\Lambda\alpha_{ll}=0.\end{array} \tag{A30}$$

which immediately implies that:

Lemma A3 *In the solution to problem (A28), there are three possibilities:*

(i). Only constraints (22) and (23) bind, in which case $\lambda_{hh}=\Lambda\alpha_{hh},\lambda_m=\Lambda(\alpha_{hh}+2\alpha_m)$ *and* $\lambda_{ll}=\Lambda$;

(ii). Only constraints (A27) and (23) bind, in which case $\psi_{hh}=\Lambda\alpha_{hh},\lambda_m=2\Lambda\alpha_m$ *and* $\lambda_{ll}=\Lambda$;

(iii). All constraints (22), (A27) and (23) bind, in which case $\lambda_{hh}+\psi_{hh}=\Lambda\alpha_{hh},\lambda_m=2\Lambda\alpha_m+\lambda_{hh}\in[2\Lambda\alpha_m,\Lambda(\alpha_{hh}+2\alpha_m)]$, *and* $\lambda_{ll}=\Lambda$.

We now consider these possibilities in turn:

Lemma A4 *Suppose that* $\alpha_m\leq\frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$ *and* $\alpha h>c$. *Then it cannot be an optimal solution to (A28) if only constraints (22) and (23) bind.*

Proof. In this case, the solution to (A28) must solve (26). But when $\alpha_m\leq\alpha_{hh}\alpha_{ll}/(1-\alpha_{ll})$, we have from Lemma 6 that $H(\Phi)\geq G(\Phi)$ for any Φ, which by Lemma 7 implies that $A_{ll}=$

$\mathrm{E}\rho^1(x|ll)\,\eta_{ll} \geq \mathrm{E}\rho^1(x|lh)\,\eta_m^l = A_m^l$ in the solution to the relaxed problem. Moreover, by the same calculation as those when we check constraint (14) in the proof of Lemma 5, we have

$$\underbrace{A_{ll}2h - t_{ll}}_{hh \text{ announces } ll} = A_{hh}2h - t_{hh} + \left(A_{ll} - A_m^l\right)(h-l) \geq \underbrace{A_{hh}2h - t_{hh}}_{\text{truth-telling}},$$

which contradicts the hypothesis (A27) does not bind. ■

Lemma A5 *Suppose that $\alpha_m \leq \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$ and $\alpha h > c$. Then it cannot be an optimal solution to (A28) if only constraints (A27) and (23) bind.*

Proof. In this case, one can first show from the first order conditions that $\eta_{hh} = \eta_m^h = 1$. Next, the first order condition with respect to η_m^l is

$$2\sum_{x\in\mathcal{X}_n} \mathbf{a}_n(x)\,\rho^1(x)\,l\frac{x_{lh}}{n} + \lambda_m \sum_{x\in\mathcal{X}_n : x_{lh}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\rho^1(x)l + \gamma_m^l - \phi_m^l = 0.$$

Since all terms except the non-negativity multipliers are strictly positive, it must be that $\eta_m^l = 1$.

The first order condition with respect to $\rho^1(x)$ reads

$$\begin{aligned}
&2\mathbf{a}_n(x)\left[\tfrac{(\eta_{hh}x_{hh}+\eta_m^h x_{hl})h+(\eta_m^l x_{lh}+\eta_{ll}x_{ll})l}{n} - c\right] + \psi_{hh}\left[2\eta_{hh}\mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll})\,h\right]\\
&-\psi_{hh}2\eta_{ll}\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)\,h\\
&+\lambda_m\left[\eta_m^h\mathbf{a}_{n-1}(x_{hh}, x_{hl}-1, x_{lh}, x_{ll})\,h + \eta_m^l\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\,l\right]\\
&-\lambda_m\left[\eta_{ll}\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)(h+l)\right] + \lambda_{ll}2\eta_{ll}\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)\,l\\
&-\Lambda\mathbf{a}_n(x)\,2c + \gamma(x) - \phi(x) = 0.
\end{aligned} \quad , \tag{A31}$$

where, just like in (A8), the convention is that $\mathbf{a}_{n-1}(x) = 0$ if $x_{\theta_i} = 0$. Using that $\eta_{hh} = \eta_m^h = \eta_m^l = 1, \psi_{hh} = \Lambda\alpha_{hh}, \lambda_m = 2\Lambda\alpha_m$ and $\lambda_{ll} = \Lambda$ and the multinomial identities in (A9), we can simplify (A31) to

$$0 = 2\left[\frac{(x_{hh}+x_{hl})\,h + (x_{lh}+\eta_{ll}x_{ll})\,l}{n} - c\right] - \Phi\frac{x_{ll}}{n}\eta_{ll}\left[\frac{2\alpha_{hh}h + 2\alpha_m(h+l)}{\alpha_{ll}}\right] + \frac{\gamma(x)-\phi(x)}{(1+\Lambda)\,\mathbf{a}_n(x)}$$

Define

$$\widehat{Q}(x,\Phi,\eta_{ll}) = 2\left[\frac{(x_{hh}+x_{hl})\,h + (x_{lh}+\eta_{ll}x_{ll})\,l}{n} - c\right] - \Phi\frac{x_{ll}}{n}\eta_{ll}\left[\frac{2\alpha_{hh}h + 2\alpha_m(h+l)}{\alpha_{ll}}\right],$$

we may write the optimal provision rule in terms of the normalized multiplier and the (still unknown) inclusion probability for type ll as

$$\rho^1(x) = \begin{cases} 1 & \widehat{Q}(x,\Phi,\eta_{ll}) > 0 \\ 0 & \widehat{Q}(x,\Phi,\eta_{ll}) < 0 \end{cases}.$$

It is easy to check that for any $x \in \mathcal{X}_{n-1}$ and $\Phi > 0$, we have that

$$\begin{aligned}
\widehat{Q}(x_{hh}+1, x_{hl}, x_{lh}, x_{ll}, \Phi, \eta_{ll}) &= \widehat{Q}(x_{hh}, x_{hl}+1, x_{lh}, x_{ll}, \Phi, \eta_{ll})\\
&> \widehat{Q}(x_{hh}, x_{hl}, x_{lh}+1, x_{ll}, \Phi, \eta_{ll})\\
&> \widehat{Q}(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1, \Phi, \eta_{ll}),
\end{aligned}$$

implying that $\text{E}\left[\rho^1(x)\,|lh\right] \geq \text{E}\left[\rho^1(x)\,|ll\right]$. Hence, letting A_{hh}, A_m^h, A_m^h and A_m^h be defined as in Lemma 5, we have

$$\begin{aligned} 2A_{hh}h - t_{hh}\ /(\text{A27}) \text{ binds}/ &= 2A_{ll}h - t_{ll} = 2A_{ll}(h+l) - t_{ll} + A_{ll}(h-l) \\ /(23) \text{ binds}/ &= A_m^h h + A_m^l l - t_m + A_{ll}(h-l) \\ &= \underbrace{A_m^h h + A_m^l h - t_m}_{\text{payoff from announcing } hl} + \left(A_{ll} - A_m^l\right)(h-l) \end{aligned}$$

But, $\eta_{ll} \leq \eta_m^l = 1$ and $\text{E}\left[\rho^1(x)\,|ll\right] \leq \text{E}\left[\rho^1(x)\,|lh\right]$ imply that $A_{ll} \leq A_m^l$. Thus $2A_{hh}h - t_{hh} \leq A_m^h h + A_m^l h - t_m$, contradicting the hypothesis that there is slack in (22). ■

Lemma A6 *Suppose that* $\alpha_m \leq \frac{\alpha_{hh}\alpha_{ll}}{1-\alpha_{ll}}$ *and* $\alpha h > c$. *Then there exists* N *such that* $0 < \eta_{ll}(n) < 1$ *and* $0 < \eta_m^l(n) < 1$ *in the solution to (A28)*

Proof. **(Sketch)** Lemmas A4 and A5 imply that (22), (A27) and (23) all bind, which can only happen when $A_{ll} = A_m^l$. Hence, if $\eta_{ll}(n) = 0$, then $\eta_m^l(n) = 0$ and vice versa. But if $\eta_{ll}(n) = \eta_m^l(n) = 0$, one can construct an alternative sequence of incentive feasible mechanisms with $\widehat{\eta}_{ll}(n) = \widehat{\eta}_m^l(n) = \frac{\alpha h - c}{\alpha h - l}$ for every n and provide the good with probability 1 that will eventually outperforms the assumed optimal solution. Moreover, if there is a subsequence such that $\eta_{ll}(n) = 1$ for all n, then (since the pivot probability is eventually negligible and $\text{E}\rho_n(x) \to \rho^* > 0$) $\eta_m^l(n) \to 1$ along the same subsequence and vice versa since otherwise $A_{ll}(n) = A_m^l(n)$ cannot hold for every n. Arguing as in the proof of Part 3 of Proposition 4, one can show that the mechanism must be infeasible when n is sufficiently large (the idea is that the per capita revenue is approximately $\rho^* 2l$ and the per capita cost is $\rho^* 2c$). ■

Proof of Proposition 5 (Continued):

From Lemma A6, we know that when n is sufficiently large, $0 < \eta_m^l(n) < 1$ and $0 < \eta_m^l(n) < 1$. The optimality condition for $\eta_m^l(n)$ is the same as that for program (26); and the only change for the optimality condition of $\eta_{ll}(n)$ is that a term $-\psi_{hh}\sum_{x\in\mathcal{X}_n:x_{ll}\geq 1}\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)\rho^1(x)2h$ is added to the first order condition. Hence, for $0 < \eta_m^l(n) < 1$ and $0 < \eta_m^l(n) < 1$, to satisfy the optimality conditions it must be that

$$\begin{aligned} 0 &= 2\alpha_m l - \lambda_{hh}h + \lambda_m l, \\ 0 &= 2\alpha_{ll} l - \lambda_m(h+l) - \psi_{hh}2h + \lambda_{ll}2l. \end{aligned} \tag{A32}$$

The first order condition with respect to $\rho^1(x)$ is

$$\begin{aligned}
&2\mathbf{a}_n(x)\left[\tfrac{(\eta_{hh}x_{hh}+\eta_m^h x_{hl})h+(\eta_m^l x_{lh}+\eta_{ll}x_{ll})l}{n}-c\right]+(\lambda_{hh}+\psi_{hh})\left[\eta_{hh}\mathbf{a}_{n-1}(x_{hh}-1,x_{hl},x_{lh},x_{ll})\,2h\right]\\
&-\lambda_{hh}\left[\eta_m^h\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,h-\eta_m^l\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})\,h\right]\\
&-\psi_{hh}\left[\eta_{ll}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh},x_{ll}-1)\,2h\right]\\
&+\lambda_m\left[\eta_m^h\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,h+\eta_m^l\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})\,l\right]\\
&-\lambda_m\left[\eta_{ll}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh},x_{ll}-1)\,(h+l)\right]+\lambda_{ll}2\eta_{ll}^1\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh},x_{ll}-1)\,l\\
&-\Lambda\mathbf{a}_n(x)\,2c+\gamma(x)-\phi(x)=0,
\end{aligned}\tag{A33}$$

which, by using that $\eta_{hh}=\eta_m^h=1$, may be rearranged as

$$\begin{aligned}
&\tfrac{x_{hh}}{n}\left[2h+(\lambda_{hh}+\psi_{hh})\tfrac{1}{\alpha_{hh}}2h\right]+\tfrac{x_{hl}}{n}\left[2h+(\lambda_m-\lambda_{hh})\tfrac{1}{\alpha_m}h\right]\\
&+\tfrac{x_{lh}}{n}\left[2\eta_m^l l+\tfrac{\eta_m^l}{\alpha_m}(\lambda_m l-\lambda_{hh}h)\right]+\tfrac{x_{ll}}{n}\left[2\eta_{ll}l+\tfrac{\eta_{ll}}{\alpha_{ll}}\left[\lambda_{ll}2l-\psi_{hh}2h-\lambda_m(h+l)\right]\right]\\
&-2c(1+\Lambda)+\widehat{\gamma}(x)-\widehat{\phi}(x)=0
\end{aligned}\tag{A34}$$

But,

$$\begin{aligned}
&\tfrac{\eta_m^l}{\alpha_m}(\lambda_m l-\lambda_{hh}h)=\tfrac{\eta_m^l}{\alpha_m}\left(\underbrace{\alpha_m 2l+\lambda_m l-\lambda_{hh}h}_{=0\text{ by (A32)}}-\alpha_m 2l\right)=-2\eta_m^l l,\\
&\tfrac{\eta_{ll}}{\alpha_{ll}}\left[\lambda_{ll}2l-\psi_{hh}2h-\lambda_m(h+l)\right]=\tfrac{\eta_{ll}}{\alpha_{ll}}\left[\underbrace{2\alpha_{ll}l-\lambda_m(h+l)-\psi_{hh}2h}_{=0\text{ by (A32)}}+\lambda_{ll}2l-2\alpha_{ll}l\right]=-2\eta_{ll}l,\\
&(\lambda_{hh}+\psi_{hh})\tfrac{1}{\alpha_{hh}}2h=2\Lambda h\text{ (Lemma A3)},\\
&(\lambda_m-\lambda_{hh})\tfrac{1}{\alpha_m}=2\Lambda\text{ (Lemma A3)}.
\end{aligned}$$

Hence, the optimal provision rule is to provide if and only if

$$\left(\frac{x_{hh}}{n}+\frac{x_{hl}}{n}\right)2h(1+\Lambda)-2c(1+\Lambda)>0.$$

This provision rule for good 1 depends only on the number of high valuation agents for good 1. Hence, $\mathrm{E}\left[\rho_n^1(x)\,|lh\right]=\mathrm{E}\left[\rho_n^1(x)\,|ll\right]$, which implies that $\eta_{ll}(n)=\eta_m^l(n)$ and thus $A_{ll}(n)=A_m^l(n)$. We conclude that the solution to the problem must coincide with the solution for the problem where goods 1 and 2 are treated separately. The asymptotic provision and inclusion probabilities can thus be taken from Proposition 3, which completes the proof. ■

B Appendix B: Proofs for Results in Section 2.

Proof of Proposition 1.

Claim B1 *For any incentive feasible mechanism $\mathcal{G}$ of the form (3), there exist an incentive feasible mechanism*

$$G = \left(\left(\rho^j, \eta_1^j, ..., \eta_n^j \right)_{j \in \mathcal{J}}, (t_i)_{i \in \mathcal{I}} \right), \tag{B35}$$

that generates the same social surplus, where $\rho^j : \Theta^n \to [0,1]$ is the provision rule for good j, $\eta_i^j : \Theta \to [0,1]$ is the inclusion rule for agent i and good j,and $t_i : \Theta \to R$ is the transfer rule for agent i.

Proof. Consider an incentive feasible mechanism $\mathcal{G}$. Pick $k \in [0,1]$ arbitrarily and define,

$$\begin{aligned}
\rho^j(\theta) &= \mathrm{E}_{\Xi}\zeta^j(\theta,\vartheta) = \int_0^1 \zeta^j(\theta,\vartheta)\, d\vartheta \qquad\qquad (\text{B36})\\
\eta_i^j(\theta_i) &= \begin{cases} \frac{\mathrm{E}_{-i}\zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)}{\mathrm{E}_{-i}\zeta^j(\theta,\vartheta)} = \frac{\int_{\Theta_{-i}^n}\int_0^1 \zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)d\vartheta d\mathbf{F}(\theta_{-i})}{\int_{\Theta_{-i}^n}\int_0^1 \zeta^j(\theta,\vartheta)d\vartheta d\mathbf{F}(\theta_{-i})} & \text{if } \int_{\Theta_{-i}^n}\int_0^1 \zeta^j(\theta,\vartheta)\, d\vartheta d\mathbf{F}(\theta_{-i}) > 0 \\ k & \text{if } \int_{\Theta_{-i}^n}\int_0^1 \zeta^j(\theta,\vartheta)\, d\vartheta d\mathbf{F}(\theta_{-i}) = 0 \end{cases}\\
t_i(\theta_i) &= \mathrm{E}_{-i}\tau(\theta) = \int_{\Theta_{-i}^n} \tau(\theta)\, d\mathbf{F}(\theta_{-i}),
\end{aligned}$$

for each $\theta \in \Theta^n, j \in \mathcal{J}$ and $i \in \mathcal{I}$. This is a mechanism of the form in (B35), and we will call it G. Use of the law of iterated expectations on $\rho^j(\theta)$ and $t_i(\theta_i)$ shows that (BB) is unaffected when switching from $\mathcal{G}$ to G. It remains to show that the surplus is unchanged, and that (IC) and (IR) continue to hold under G. The utility of agent i of type $\theta_i \in \Theta$ who announces $\hat{\theta}_i \in \Theta$ is

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right) \omega_i^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\theta_i - \tau\left(\hat{\theta}_i, \theta_{-i}\right)\right] \text{ in mechanism } \mathcal{G} \tag{B37}$$

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \rho^j\left(\hat{\theta}_i, \theta_{-i}\right) \eta_i^j\left(\hat{\theta}_i\right)\theta_i - t_i\left(\hat{\theta}_i\right)\right] \text{ in mechanism } G. \tag{B38}$$

If $\int_{\Theta_{-i}^n}\int_0^1 \zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right) d\vartheta d\mathbf{F}(\theta_{-i}) = 0$, we trivially have that the payoffs in (B37) and (B38) are identical, whereas if $\int_{\Theta_{-i}^n}\int_0^1 \zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right) d\vartheta d\mathbf{F}(\theta_{-i}) > 0$, we have that

$$\begin{aligned}
\mathrm{E}_{-i}\rho^j\left(\hat{\theta}_i, \theta_{-i}\right)\eta_i^j\left(\hat{\theta}_i\right)\theta_i &= \mathrm{E}_{-i}\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\frac{\mathrm{E}_{-i}\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)}{\mathrm{E}_{-i}\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)} \qquad (\text{B39})\\
&= \mathrm{E}_{-i}\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\theta_i.
\end{aligned}$$

Trivially, $\mathrm{E}_{-i}t_i(\theta_i) = t_i(\theta_i) = \mathrm{E}_{-i}\tau(\theta)$, which combined with (B39) implies that the payoffs in (B37) and (B38) are identical. Since the equality between (B37) and (B38) were established for any i, θ_i and $\hat{\theta}_i$, it follows that all incentive and participation constraints (IC) and (IR) hold for mechanism G given that they are satisfied in mechanism $\mathcal{G}$. Moreover, [again by (B39)]

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i\right] = \mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \omega_i^j(\theta,\vartheta)\,\zeta^j(\theta,\vartheta)\,\theta_i\right], \tag{B40}$$

so it follows by integration over Θ and summation over i that

$$\mathrm{E}\left[\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i\right]=\mathrm{E}\left[\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\zeta^j(\theta,\vartheta)\,\omega_i^j(\theta,\vartheta)\,\theta_i\right], \tag{B41}$$

By construction, we also have that $\rho^j(\theta)=\mathrm{E}_\Xi\zeta^j(\theta,\vartheta)$ for every θ. Thus $\mathrm{E}\left[\rho^j(\theta)\,C^j(n)\right]=\mathrm{E}\left[\zeta^j(\theta,\vartheta)\,C^j(n)\right]$, implying that

$$\sum_{j\in\mathcal{J}}\mathrm{E}\rho^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_i^j(\theta_i)\,\theta_i-C^j(n)\right]=\sum_{j\in\mathcal{J}}\mathrm{E}\zeta^j(\theta,\vartheta)\left[\sum_{i\in\mathcal{I}}\omega_i^j(\theta,\vartheta)\,\theta_i-C^j(n)\right]. \tag{B42}$$

Hence, $\mathcal{G}$ and G generate the same social surplus. ■

Claim B2 *For every incentive feasible mechanism of the form (B35), there exists an anonymous simple incentive feasible mechanism g of the form (5) that generates the same surplus.*

Proof. Consider an incentive feasible simple mechanism G on form (B35). For $k\in\{1,....,n!\}$, let $P_k:\mathcal{I}\to\mathcal{I}$ denote the k-th permutation of the set of agents $\mathcal{I}$. Note that $P_k^{-1}(i)$ gives the index of the agent who takes agent i's position in permutation P_k. Moreover, for any given $\theta\in\Theta^n$, let $\theta^{P_K}=\left(\theta_{P_k^{-1}(1)},...,\theta_{P_k^{-1}(n)}\right)\in\Theta^n$ denote the corresponding k-th permutation of θ.[22] For each $k\in\{1,...,n!\}$, let $G_k=\left(\left(\rho_k^j,\eta_{k1}^j,...,\eta_{kn}^j\right)_{j=1,2},t_{k1},...,t_{kn}\right)$ be given by

$$\begin{aligned}\rho_k^j(\theta)&=\rho^j\left(\theta^{P_k}\right)\quad\forall\,\theta\in\Theta^n,\ j\in\mathcal{J},\\ \eta_{ki}^j(\theta_i)&=\eta_{P_k^{-1}(i)}^j(\theta_i)\quad\forall\,\theta_i\in\Theta,\ j\in\mathcal{J},i\in\mathcal{I},\\ t_{ki}(\theta_i)&=t_{P_k^{-1}(i)}(\theta_i)\quad\forall\,\theta_i\in\Theta,i\in\mathcal{I},\end{aligned} \tag{B43}$$

and let $g=\left(\left(\widetilde{\rho}^j,\widetilde{\eta}_1^j,...,\widetilde{\eta}_n^j\right)_{j=1,2},\widetilde{t}_1,...,\widetilde{t}_n\right)$ be given by

$$\begin{aligned}&\widetilde{\rho}^j(\theta)=\tfrac{1}{n!}\textstyle\sum_{k=1}^{n!}\rho_k^j(\theta)\quad\forall\,\theta\in\Theta^n,\ j\in\mathcal{J}\\ &\widetilde{\eta}_i^j(\theta_i)=\frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]}\quad\forall\,\theta_i\in\Theta,i\in\mathcal{I},\ j\in\mathcal{J}\\ &\widetilde{t}_i(\theta_i)=\tfrac{1}{n!}\textstyle\sum_{k=1}^{n!}t_{ki}(\theta_i)\quad\forall\,\theta_i\in\Theta,i\in\mathcal{I}.\end{aligned} \tag{B44}$$

We now note that: (1) for each $j\in\mathcal{J}$, $\widetilde{\rho}^j(\theta)=\widetilde{\rho}^j(\theta')$ if θ' is a permutation of θ. This is immediate since the sets $\left\{\rho_k^j(\theta)\right\}_{k=1}^{n!}=\left\{\rho^j(P_k(\theta))\right\}_{k=1}^{n!}$ and $\left\{\rho_k^j(\theta')\right\}_{k=1}^{n!}=\left\{\rho^j(P_k(\theta'))\right\}_{k=1}^{n!}$ are the same; (2) for $j\in\mathcal{J}$ and each pair $i,i'\in\mathcal{I},\widetilde{\eta}_i^j(\cdot)=\widetilde{\eta}_{i'}^j(\cdot)$. That is, the inclusion rules are the same for all agents. To see this, consider agent i and i', and suppose that $\theta_i=\theta_{i'}$. We then have that $\left\{\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)\right\}_{k=1}^{n!}$ and $\left\{\mathrm{E}_{-i'}\left[\rho_k^j(\theta)\right]\eta_{ki'}^j(\theta_{i'})\right\}_{k=1}^{n!}$ are identical and that $\mathrm{E}_{-i}\left[\widetilde{\rho}^j(\theta)\right]=$

[22] To illustrate, suppose $n=3,m=2,\theta=(\theta_1,\theta_2,\theta_3)=((1,2),(3,2),(2,1))$. Consider, for example, purmutation k given by $P_k(1)=2,P_k(2)=1,P_k(3)=3$. Then $P_k^{-1}(1)=2,P_k^{-1}(2)=1,P_k^{-1}(3)=3$ and $\theta^{P_k}=\left(\theta_{P_k^{-1}(1)},\theta_{P_k^{-1}(2)},\theta_{P_k^{-1}(3)}\right)=(\theta_2,\theta_1,\theta_3)=((3,2),(1,2),(2,1))$.

$\mathrm{E}_{-i'}\left[\widetilde{\rho}^j(\theta)\right]$; and (3) for each pair $i, i' \in \mathcal{I}, \widetilde{t}_i(\cdot) = \widetilde{t}_{i'}(\cdot)$, which is obvious since the sets $\{t_{ki}(\theta_i)\}_{k=1}^{n!}$ and $\left\{t_{ki}\left(\theta_i'\right)\right\}_{k=1}^{n!}$ are identical. Together, (1), (2) and (3) establishes that g is anonymous and simple.

Now we show that g is incentive feasible and generates the same expected surplus as G. First, since G and G_k are identical except for the permutation of the agents, we have, for $k = 1, ..., n!$,

$$\sum_{j\in\mathcal{J}} \mathrm{E}\left\{\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}} \eta_{ki}^j(\theta_i)\,\theta_i^j - C^j(n)\right]\right\} = \sum_{j\in\mathcal{J}} \mathrm{E}\left\{\rho^j(\theta)\left[\sum_{i\in\mathcal{I}} \eta_i^j(\theta_i)\,\theta_i^j - C^j(n)\right]\right\}. \qquad \text{(B45)}$$

Hence,

$$\begin{aligned}
\sum_{j\in\mathcal{J}} \mathrm{E}\left\{\widetilde{\rho}^j(\theta)\left[\sum_{i\in\mathcal{I}} \widetilde{\eta}^j(\theta_i)\,\theta_i^j - C^j(n)\right]\right\} &= \sum_{j\in\mathcal{J}} \mathrm{E}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}} \frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\rho_k^j(\theta)\,\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\rho_k^j(\theta)}\theta_i^j - C^j(n)\right]\right\} \\
= \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} \mathrm{E}_{\theta_i}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\mathrm{E}_{-i}\rho_k^j(\theta)\,\eta_{ki}^j(\theta_i)\,\theta_i^j\right\} &- \mathrm{E}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]C^j(n) \\
= \frac{1}{n!}\sum_{k=1}^{n!}\sum_{j\in\mathcal{J}} \mathrm{E}\left\{\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}} \eta_{ki}^j(\theta_i)\,\theta_i^j - C^j(n)\right]\right\} &= \sum_{j\in\mathcal{J}} \mathrm{E}\left\{\rho^j(\theta)\left[\sum_{i\in\mathcal{I}} \eta_i^j(\theta_i)\,\theta_i^j - C^j(n)\right]\right\}, \qquad \text{(B46)}
\end{aligned}$$

where the last equality follows from (B45). Hence the surplus generated by g is identical to that by original mechanism G. To show that g is incentive feasible we first note that $\mathrm{E}\rho_k^j(\theta) = \mathrm{E}\rho^j(\theta)$ and $\mathrm{E}\sum_{i\in\mathcal{I}} t_{ki}(\theta_i) = \mathrm{E}\sum_{i\in\mathcal{I}} t_i(\theta_i)$ for all k, since the agents' valuations are drawn from identical distributions and G_k and G only differ in the index of the agents. Thus

$$\begin{aligned}
\mathrm{E}\textstyle\sum_{i\in\mathcal{I}}\widetilde{t}_i(\theta_i) - \sum_{j\in\mathcal{J}}\mathrm{E}\widetilde{\rho}^j(\theta)\,C^j(n) &= \mathrm{E}\textstyle\sum_{i\in\mathcal{I}}\frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\theta_i) - \sum_{j\in\mathcal{J}}\mathrm{E}\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\,C^j(n) \\
&= \mathrm{E}\textstyle\sum_{i\in\mathcal{I}} t_i(\theta_i) - \sum_{j\in\mathcal{J}}\mathrm{E}\rho^j(\theta)\,C^j(n)\ , \qquad \text{(B47)}
\end{aligned}$$

so g satisfies (BB) if G does. Second, (IC) holds for any permuted mechanism, that is,

$$\mathrm{E}_{-i}\textstyle\sum_{j\in\mathcal{J}}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\,\theta_i^j - t_{ki}(\theta_i) \geq \mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j - t_{ki}(\widehat{\theta}_i,\theta_{-i}) \qquad \text{(B48)}$$

for all $i \in \mathcal{I}$, and $\theta_i, \widehat{\theta}_i \in \Theta$. Hence,

$$\begin{aligned}
\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\widetilde{\rho}^j(\theta)\widetilde{\eta}^j(\theta_i)\,\theta_i^j - \widetilde{t}(\theta_i) &= \mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]\frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]}\theta_i^j - \frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\theta_i) \\
= \frac{1}{n!}\sum_{k=1}^{n!}\left[\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\rho_k^j(\theta)\,\eta_{ki}^j(\theta_i)\,\theta_i^j - t_{ki}(\theta_i)\right] &\geq \frac{1}{n!}\sum_{k=1}^{n!}\left[\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j - t_{ki}(\widehat{\theta}_i,\theta_{-i})\right] \\
= \mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j - \frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\widehat{\theta}_i,\theta_{-i}) &= \sum_{j\in\mathcal{J}}\mathrm{E}_{-i}\widetilde{\rho}^j(\widehat{\theta}_i,\theta_{-i})\widetilde{\eta}_i^j\left(\widehat{\theta}_i\right)\theta_i^j - \widetilde{t}(\widehat{\theta}_i), \qquad \text{(B49)}
\end{aligned}$$

where the inequality follows from (B48). Hence g is satisfies (IC). Finally, g also satisfies the (IR) because (see the second line in (B49)) all the permuted mechanisms satisfy participation constraints. Proposition 1 follows by combining Claims B1 and B2. ■

Proof of Proposition 2.

Notation: This proof requires us to be explicit about the coordinates of the vector θ when permuting $\mathcal{J}$. We therefore need some extra notation for this proof (only). We write $\theta_i^{-j} = \left(\theta_i^1, ..., \theta_i^{j-1}, \theta_i^{j+1}, ..., \theta_i^m\right)$ for a type vector where good j has been removed. Analogously, $\theta^{-j} = \left(\theta_1^{-j}, ..., \theta_n^{-j}\right)$ stands for the type profile with good j coordinate removed for all agents and $\theta^j = \left(\theta_1^j, ..., \theta_n^j\right)$ is the vector collecting the valuations for good j for all agents. Furthermore, $\theta_{-i}^{-j} = \left(\theta_1^{-j}, ..., \theta_{i-1}^{-j}, \theta_{i+1}^{-j}, ..., \theta_n^{-j}\right)$ and $\theta_{-i}^j = \left(\theta_1^j, ..., \theta_{i-1}^j, \theta_{i+1}^j ..., \theta_n^j\right)$ are used for the vectors obtained respectively from θ^{-j} and θ^j by removing agent i. These conventions are used also on the distributions, so, for example, $\mathbf{F}_{-i}^{-j}$ denotes the cumulative distribution of θ_{-i}^{-j}. Conditional distributions are denoted in the natural way: for example $\mathbf{F}_{-i}^{-j}\left(\cdot|\,\theta_i^j\right)$ denotes the joint distribution of θ_{-i}^{-j} conditional on θ_i^j. Since no integrals are taken over subsets of the range of integration, we also conserve space and write $\int_\theta h(\theta)\, d\mathbf{F}(\theta)$ rather than $\int_{\theta\in\Theta^n} h(\theta)\, d\mathbf{F}(\theta)$ when integrating a function h over θ and similarly for integrals over various components of θ.

Proof. Consider a simple anonymous incentive feasible mechanism g. For $k \in \{1, ..., m!\}$ let $P_k : \mathcal{J} \to \mathcal{J}$ be the k-th permutation of $\mathcal{J}$ and $\theta_i^{P_k} = \left(\theta_i^{P_k^{-1}(1)}, ..., \theta_i^{P_k^{-1}(m)}\right) \in \Theta$ be the permutation of θ_i when the goods are permuted according to P_k. Let $\theta^{P_k} = \left(\theta_1^{P_k}, ..., \theta_n^{P_k}\right) \in \Theta^n$ denote the corresponding permutation of θ.[23] For each $k \in \{1, ..., m!\}$ define mechanism $g_k = (\left\{\rho_k^j\right\}_{j\in\mathcal{J}}, \left\{\eta_k^j\right\}_{j\in\mathcal{J}}, t_k)$, where for every $\theta \in \Theta^n$;

1. $\rho_k^j(\theta) = \rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)$ for every $j \in \mathcal{J}$;[24]
2. $\eta_k^j(\theta_i) = \eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)$ for every $j \in \mathcal{J}$;[25]
3. $t_k(\theta_i) = t\left(\theta_i^{P_k}\right)$.

By construction, each g_k is simple. Each g_k is also anonymous by the anonymity of g. Using the definition of g_k and manipulating the result by observing that the labeling of the variables is irrelevant, we get:[26]

$$\begin{aligned}
\mathrm{E}\rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j &= \int_\theta \rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j d\mathbf{F}(\theta) \;/\text{def of } g_k/ = \int_{\theta\in\Theta^n} \rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\theta_i^j d\mathbf{F}(\theta) \\
&= \int_{\theta^j}\left[\int_{\theta^{-j}} \rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\theta_i^j d\mathbf{F}^{-j}\left(\theta^{-j}\middle|\,\theta^j\right)\right] d\mathbf{F}^j\left(\theta^j\right) \qquad \text{(B50)} \\
/\text{relabel}/ &= \int_{\theta^{P_k^{-1}(j)}}\left[\int_{\left(\theta^{-j}\right)^{P_k}} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}^{-j}\left(\left(\theta^{-j}\right)^{P_k}\middle|\, \underbrace{\theta^{P_k^{-1}(j)}}_{j\text{-th argument}}\right)\right] d\mathbf{F}^j\left(\theta^{P_k^{-1}(j)}\right)
\end{aligned}$$

[23] To illustrate, suppose $n = 2, m = 3$, and $\theta = (\theta_1, \theta_2) = ((1,2,0), (3,2,1))$. Consider, for example, purmutation k given by $P_k(1) = 2, P_k(2) = 1, P_k(3) = 3$. Then $P_k^{-1}(1) = 2, P_k^{-1}(2) = 1, P_k^{-1}(3) = 3$ and $\theta_1^{P_k} = \left(\theta_1^{P_k^{-1}(1)}, \theta_1^{P_k^{-1}(2)}, \theta_1^{P_k^{-1}(3)}\right) = (2,1,0), \theta_2^{P_k} = \left(\theta_2^{P_k^{-1}(1)}, \theta_2^{P_k^{-1}(2)}, \theta_2^{P_k^{-1}(3)}\right) = (2,3,1), \theta^{P_k} = \left(\theta_1^{P_k}, \theta_2^{P_k}\right) = ((2,1,0), (2,3,1))$.

[24] This implies that $\rho_k^{P_k^{-1}(j)}\left(\theta^{P_k}\right) = \rho^j(\theta)$ for every $j \in \mathcal{J}$.

[25] This implies that $\eta_k^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right) = \eta^j(\theta_i)$ for every $j \in \mathcal{J}$.

[26] It is important to point out that, in reaching the fourth equality in (B50), we can relabel the integrating varibles (since they are dummies) but not the integrating functions.

where we recall,

$$\left(\theta^{-j}\right)^{P_k} \equiv \left(\theta^{P_k^{-1}(1)}, ..., \theta^{P_k^{-1}(j-1)}, \theta^{P_k^{-1}(j+1)}, ..., \theta^{P_k^{-1}(n)}\right). \tag{B51}$$

By exchangeability, we have

$$\begin{aligned}
& d\mathbf{F}^{-j}\left(\left(\theta^{-j}\right)^{P_k} \Big| \underbrace{\theta^{P_k^{-1}(j)}}_{j\text{-th (vector) argument}}\right) \qquad \text{(B52)}\\
= & \; d\mathbf{F}^{-j}\left(\theta^{P_k^{-1}(1)}, ..., \theta^{P_k^{-1}(j-1)}, \theta^{P_k^{-1}(j+1)}, ..., \theta^{P_k^{-1}(n)} | j\text{-th (vector) argument} = \theta^{P_k^{-1}(j)}\right)\\
= & \; d\mathbf{F}^{-j}\left(\theta^{-j} | j\text{-th (vector) argument} = \theta^{P_k^{-1}(j)}\right)\\
= & \; d\mathbf{F}^{-P_k^{-1}(j)}\left(\theta^{-P_k^{-1}(j)} | P_k^{-1}(j)\text{-th (vector) argument } = \theta^{P_k^{-1}(j)}\right);
\end{aligned}$$

and

$$d\mathbf{F}^j\left(\theta^{P_k^{-1}(j)}\right) = d\mathbf{F}^{\theta^{P_k^{-1}(j)}}\left(\theta^{P_k^{-1}(j)}\right). \tag{B53}$$

Using (B50), (B52) and (B53), we have that

$$\begin{aligned}
& \mathrm{E}\rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j \qquad \text{(B54)}\\
= & \int_{\theta^{P_k^{-1}(j)}}\left[\int_{\left(\theta^{-j}\right)^{P_k}} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}^{-j}\left(\left(\theta^{-j}\right)^{P_k} \Big| \underbrace{\theta^{P_k^{-1}(j)}}_{j\text{-th argument}}\right)\right] d\mathbf{F}^j\left(\theta^{P_k^{-1}(j)}\right)\\
= & \int_{\theta^{P_k^{-1}(j)}}\left[\int_{\left(\theta^{-j}\right)^{P_k}} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}^{-P_k^{-1}(j)}\left(\theta^{-P_k^{-1}(j)} | \underbrace{\theta^{P_k^{-1}(j)}}_{P_k^{-1}(j)\text{-th argument}}\right)\right] d\mathbf{F}^{P_k^{-1}(j)}\left(\theta^{P_k^{-1}(j)}\right)\\
= & \int_{\theta} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}(\theta) = \mathrm{E}\rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\theta_i^{P_k^{-1}(j)}.
\end{aligned}$$

Moreover, exchangeability implies that $\mathrm{E}t_k(\theta_i) = \mathrm{E}t\left(\theta_i^{P_k}\right) = \mathrm{E}t(\theta_i)$. The ex ante utility,

$$\begin{aligned}
\mathrm{E}\left[\sum_{j=1}^m \rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j - t_k(\theta_i)\right] &= \left[\sum_{j=1}^m \mathrm{E}\rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\theta_i^{P_k^{-1}(j)}\right] - \mathrm{E}t(\theta_i)\\
\Big/\ \begin{matrix}\text{Same elements in } \mathcal{J} \text{ and}\\ \left\{P_k^{-1}(1), ..., P_k^{-1}(m)\right\}\end{matrix}\ \Big/ &= \left[\sum_{j=1}^m \mathrm{E}\rho^j(\theta)\,\eta^j(\theta_i)\theta_i^j\right] - \mathrm{E}t(\theta_i), \qquad \text{(B55)}
\end{aligned}$$

is thus unchanged when changing from g to g_k. The same steps as in (B50) through (B54) (only somewhat simpler) establishes that $\mathrm{E}\rho_k^j(\theta) = \mathrm{E}\rho^{P_k^{-1}(j)}$ for every j, implying that

$$\begin{aligned}
\mathrm{E}\left[\sum_{j=1}^m \rho_k^j(\theta)\,C^j(n) - \sum_i t_k(\theta_i)\right] &= \left[C(n)\,\mathrm{E}\sum_{j=1}^m \rho_k^j(\theta) - \sum_i \mathrm{E}t_k(\theta_i)\right] \qquad \text{(B56)}\\
= \left[C(n)\,\mathrm{E}\sum_{j=1}^m \rho^j(\theta) - \sum_i \mathrm{E}t(\theta_i)\right] &= \mathrm{E}\left[\sum_{j=1}^m \rho^j(\theta)\,C(n) - \sum_i t(\theta_i)\right],
\end{aligned}$$

so the feasibility constraint is unaffected when changing from g to g_k. Next, write Write $U(\theta_i, \theta_i'; g)$ and $U(\theta_i, \theta_i'; g_k)$ for the expected utility from announcing θ_i' when the true type is θ_i in mechanisms g and g_k respectively. Next, by a calculation in the same spirit as (B50) through (B54):

$$\begin{aligned}
\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'\right) &= \int_{\theta_{-i}} \rho_k^j\left(\theta_{-i},\theta_i'\right) d\mathbf{F}_{-i}\left(\theta_{-i}\right) /\text{def of } g_k/ = \int_{\theta_{-i}} \rho^{P_k^{-1}(j)}\left(\left(\theta_{-i},\theta_i'\right)^{P_k}\right) d\mathbf{F}_{-i}\left(\theta_{-i}\right) \\
&= \int_{\theta_{-i}^j}\left[\int_{\theta_{-i}^{-j}} \rho^{P_k^{-1}(j)}\left(\left(\theta_{-i},\theta_i'\right)^{P_k}\right) d\mathbf{F}_{-i}^{-j}\left(\theta_{-i}^{-j}\middle|\theta_{-i}^j\right)\right] d\mathbf{F}_{-i}^j\left(\theta_{-i}^j\right) \qquad \text{(B57)} \\
/\text{relabel}/ &= \int_{\theta_{-i}^{P_k^{-1}(j)}}\left[\int_{\theta_{-i}^{-P_k^{-1}(j)}} \rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) d\mathbf{F}_{-i}^{-j}\left(\left(\theta_{-i}^{-j}\right)^{P_k}\middle|\theta_{-i}^{P_k^{-1}(j)}\right)\right] d\mathbf{F}_{-i}^j\left(\theta_{-i}^{P_k^{-1}(j)}\right) \\
/\text{exchangeability}/ &= \int_{\theta_{-i}^{P_k^{-1}(j)}}\left[\int_{\theta_{-i}^{-P_k^{-1}(j)}} \rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) d\mathbf{F}_{-i}^{-P_k^{-1}(j)}\left(\theta_{-i}^{-P_k^{-1}(j)}\middle|\theta_{-i}^{P_k^{-1}(j)}\right)\right] d\mathbf{F}_{-i}^j\left(\theta_{-i}^{P_k^{-1}(j)}\right) \\
&= \int_{\theta_{-i}} \rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) d\mathbf{F}_{-i}\left(\theta_{-i}\right) = \mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right)
\end{aligned}$$

That is, the perceived probability of getting j when announcing θ_i' in mechanism g_k is the same as the perceived probability of getting good $P_k^{-1}(j)$ when announcing $\left(\theta_i'\right)^{P_k}$, so that

$$\begin{aligned}
U(\theta_i,\theta_i';g_k) &= \mathrm{E}_{-i}\sum_{j=1}^m \rho_k^j\left(\theta_{-i},\theta_i'\right)\eta_k^j(\theta_i')\theta_i^j - t_k\left(\theta_i'\right) \qquad \text{(B58)} \\
&= \sum_{j=1}^m \eta_k^{P_k^{-1}(j)}((\theta_i')^{P_k})\theta_i^j\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) - t\left(\left(\theta_i'\right)^{P_k}\right),
\end{aligned}$$

whereas

$$U(\theta_i,\theta_i';g) = \sum_{j=1}^m \eta_k^j(\theta_i')\theta_i^j\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'\right) - t\left(\theta_i'\right) \Rightarrow \qquad \text{(B59)}$$

$$U(\theta_i,\theta_i';g)\Big|_{\substack{\theta_i=\theta_i^{P_k}\\ \theta_i'=\theta_i'^{P_k}}} = \sum_{j=1}^m \eta_k^j((\theta_i')^{P_k})\theta_i^{P_k^{-1}(j)}\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'^{P_k}\right) - t\left(\left(\theta_i'\right)^{P_k}\right)$$

$$= \sum_{j=1}^m \eta_k^{P_k^{-1}(j)}((\theta_i')^{P_k})\theta_i^j\mathrm{E}_{-i}\rho_k^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) - t\left(\left(\theta_i'\right)^{P_k}\right) = U(\theta_i,\theta_i';g_k),$$

which establishes that type θ_i who announces θ_i' in mechanism g_k gets the same utility as type $\theta_i^{P_k}$ who announces $\left(\theta_i'\right)^{P_k}$ in mechanism g. Hence incentive compatibility and individual rationality of g_k follows from incentive compatibility and individual rationality of g. Now, construct a new mechanism $\widetilde{g} = (\{\widetilde{\rho}^j\}_{j\in\mathcal{J}}, \{\widetilde{\eta}^j\}_{j\in\mathcal{J}}, \widetilde{t})$ by letting

$$\begin{aligned}
\widetilde{\rho}^j\left(\theta\right) &= \tfrac{1}{m!}\sum_{k=1}^{m!}\rho_k^j\left(\theta\right) = \tfrac{1}{m!}\sum_{k=1}^{m!}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right) \qquad \text{(B60)} \\
\widetilde{\eta}^j\left(\theta_i\right) &= \frac{\sum_{k=1}^{m!}\eta_k^j\left(\theta_i\right)\mathrm{E}_{-i}\rho_k^j\left(\theta\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j\left(\theta\right)} = \frac{\sum_{k=1}^{m!}\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)} \\
\widetilde{t}\left(\theta_i\right) &= \tfrac{1}{m!}t_k\left(\theta_i\right) = \tfrac{1}{m!}t\left(\theta_i^{P_k}\right)
\end{aligned}$$

let $P:\mathcal{J}\to\mathcal{J}$ be an arbitrary perturbation of the set of goods. Then,

$$\widetilde{\rho}^{P^{-1}(j)}\left(\theta^P\right) = \tfrac{1}{m!}\sum_{k=1}^{m!}\rho^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta^P\right)^{P_k}\right) = \tfrac{1}{m!}\sum_{k=1}^{m!}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right) = \widetilde{\rho}^j\left(\theta\right), \qquad \text{(B61)}$$

since the sets $\left\{\rho^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta^P\right)^{P_k}\right)\right\}_{k=1}^{m!}$ and $\left\{\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)\right\}_{k=1}^{m!}$ are identical. Furthermore

$$\begin{aligned}\widetilde{\eta}^{P^{-1}(j)}\left(\theta_i^P\right) &= \frac{\sum_{k=1}^{m!}\eta_k^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta_i^P\right)^{P_k}\right)\mathrm{E}_{-i}\rho_k^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta^P\right)^{P_k}\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta^P\right)^{P_k}\right)} \\ &= \frac{\sum_{k=1}^{m!}\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)} = \widetilde{\eta}^j\left(\theta_i\right)\end{aligned} \tag{B62}$$

for the same reason. It is obvious that $\widetilde{t}\left(\theta_i^P\right) = \widetilde{t}\left(\theta_i\right)$, which together with (B61) and (B62) establishes that $\widetilde{g}$ is symmetric. To complete the proof we need to show that $\widetilde{g}$ is incentive feasible and generates the same surplus as g. We note that

$$\begin{aligned}&\mathrm{E}\widetilde{\rho}^j\left(\theta\right)\widetilde{\eta}^j\left(\theta_i\right)\theta_i^j = \tfrac{1}{m!}\textstyle\sum_{k=1}^{m!}\mathrm{E}\rho_k^j\left(\theta\right)\frac{\sum_{k=1}^{m!}\eta_k^j(\theta_i)\mathrm{E}_{-i}\rho_k^j(\theta)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j(\theta)}\theta_i^j \\ &= \tfrac{1}{m!}\mathrm{E}_{\theta_i}\textstyle\sum_{k=1}^{m!}\left[\mathrm{E}_{-i}\rho_k^j\left(\theta\right)\frac{\sum_{k=1}^{m!}\eta_k^j(\theta_i)\mathrm{E}_{-i}\rho_k^j(\theta)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j(\theta)}\theta_i^j\right] = \tfrac{1}{m!}\mathrm{E}\left[\sum_{k=1}^{m!}\eta_k^j\left(\theta_i\right)\rho_k^j\left(\theta\right)\theta_i^j\right] \\ &\Rightarrow \mathrm{E}\textstyle\sum_{j=1}^{n}\left[\widetilde{\rho}^j\left(\theta\right)\widetilde{\eta}^j\left(\theta_i\right)\theta_i^j - \widetilde{t}\left(\theta_i\right)\right] = \tfrac{1}{m!}\sum_{k=1}^{m!}\mathrm{E}\left[\sum_{j=1}^{m}\eta_k^j\left(\theta_i\right)\rho_k^j\left(\theta\right)\theta_i^j - t_k\left(\theta_i\right)\right] \\ &/\text{(B54) \& (B55)}/ = \mathrm{E}\left[\textstyle\sum_{j=1}^{m}\eta^j\left(\theta_i\right)\rho^j\left(\theta\right)\theta_i^j - t\left(\theta_i\right)\right],\end{aligned} \tag{B63}$$

which establishes that the ex ante utility from $\widetilde{g}$ and g are the same for all agents. Moreover,

$$\begin{aligned}&\mathrm{E}\left[\textstyle\sum_{j=1}^{m}\widetilde{\rho}^j\left(\theta\right)C^j\left(n\right) - \sum_{i=1}^{n}\widetilde{t}\left(\theta_i\right)\right] = \mathrm{E}\left[C\left(n\right)\sum_{j=1}^{m}\tfrac{1}{m!}\sum_{k=1}^{m!}\rho_k^j\left(\theta\right) - \sum_{i=1}^{n}\sum_{k=1}^{m!}\tfrac{1}{m!}t_k\left(\theta_i\right)\right] \\ &= \textstyle\sum_{k=1}^{m!}\mathrm{E}\left[C\left(n\right)\sum_{j=1}^{m}\rho_k^j\left(\theta\right) - \sum_{i=1}^{n}t_k\left(\theta_i\right)\right] \; / \text{ (B56)}/ \\ &= \tfrac{1}{m!}\textstyle\sum_{k=1}^{m!}\mathrm{E}\left[\sum_{j=1}^{m}\rho^j\left(\theta\right)C\left(n\right) - \sum_{i=1}^{n}t\left(\theta_i\right)\right] = \mathrm{E}\left[\sum_{j=1}^{m}\rho^j\left(\theta\right)C^j\left(n\right) - \sum_{i=1}^{n}t\left(\theta_i\right)\right],\end{aligned} \tag{B64}$$

so the budget balance constraint is unaffected. All incentive compatibility constraints hold since,

$$U(\theta_i,\theta_i';\widetilde{g}) = \sum_{j=1}^{m}\widetilde{\eta}^j(\theta_i')\theta_i^j\mathrm{E}_{-i}\widetilde{\rho}^j\left(\theta_{-i},\theta_i'\right) - \widetilde{t}\left(\theta_i'\right) \tag{B65}$$

$$\begin{aligned}&= \frac{\sum_{k=1}^{m!}\eta_k^j(\theta_i')\mathrm{E}_{-i}\rho_k^j(\theta_{-i},\theta_i')}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j(\theta_{-i},\theta_i')}\mathrm{E}_{-i}\left[\tfrac{1}{m!}\textstyle\sum_{k=1}^{m!}\rho_k^j\left(\theta_{-i},\theta_i'\right)\right] - \tfrac{1}{m!}\sum_{k=1}^{m!}t_k\left(\theta_i'\right) \\ &= \tfrac{1}{m!}\textstyle\sum_{k=1}^{m!}\left[\eta_k^j\left(\theta_i'\right)\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'\right) - t_k\left(\theta_i'\right)\right] \\ /\text{ (B58)}/ &= \tfrac{1}{m!}\textstyle\sum_{k=1}^{m!}U(\theta_i,\theta_i';g_k) \;\le /\text{ IC for each } k/\; \tfrac{1}{m!}\sum_{k=1}^{m!}U(\theta;g_k) = U(\theta;\widetilde{g}).\end{aligned}$$

By the same calculation, $U(\theta;\widetilde{g}) = \frac{1}{m!}\sum_{k=1}^{m!}U(\theta;g_k) \ge 0$, since all participation constraints hold for each k. This completes the proof. ∎