View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

AN EFFICIENCY RATIONALE FOR BUNDLING OF PUBLIC GOODS

By

Hanming Fang and Peter Norman

October 2003

COWLES FOUNDATION DISCUSSION PAPER NO. 1441

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

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

An Efficiency Rationale for Bundling of Public Goods*

Hanming Fang[†] Peter Norman[‡]

First Version: October 2002
This Version: October 1, 2003

Abstract

This paper studies the role of bundling in the efficient provision of excludable public goods. We show that bundling in the provision of unrelated public goods can enhance social welfare. With a large number of goods and agents, first best can be approximated with pure bundling. For a parametric class of problems with binary valuations, we characterize the optimal mechanism, and show that bundling alleviates the free riding problem in large economies and decreases the extent of use exclusions. Both results are related to the idea that bundling makes it possible to reduce the incidence of exclusions because the variance in the relevant valuations decreases.

Keywords: Public Goods Provision; Bundling; Exclusion

JEL Classification Number: H41.

*We thank Mark Armstrong, Ted Bergstrom, Martin Hellwig, Larry Samuelson and Stephen Morris for comments and helpful discussions. We also thank seminar participants at New York University, Ohio State University, University of Essex, University of Illinois, University of Pittsburgh, University of Wisconsin, Yale University, Society for Economic Dynamics 2003 meetings in Paris, and SITE 2003 Summer Workshop in Stanford for useful comments. 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 Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706-1393. Email: pnorman@facstaff.wisc.edu

1 Introduction

Bundling, the practice to package several goods in a bundle rather than providing them separately, is a common phenomenon in many markets. Many of the goods that are provided in bundles are more or less non-rival in consumption. An obvious example is cable TV. Technologically, the local cable company could allow customers to choose whatever channels they are willing to pay for without constraints. In practice, the basic pricing scheme usually consists of a limited number of available packages. While some premium channels and pay-per-view programming are offered for sale separately, the bundled channels are simply not available in any other way than through their respective bundles. Another striking example is access to electronic libraries. Here, the typical contractual arrangement is a site license that allows access to every journal in the electronic library. While it is sometimes possible to download articles on a pay-per-download basis, this is usually very expensive, and contracts that gives access to a subset of journals in the electronic library are rare.

A third example, which was the initial motivation for this paper, is the casual observation that governmental services are provided in bundles. For example, every resident in a municipality is entitled to a bundle of public services provided by the local government including policing, highway maintenance, fire fighting, public schools etc.. Clearly some of the public services in the bundle are of no value at all for many residents. Why, then, cannot residents only subscribe to their desired local public services?

Motivated by these examples, this paper studies the role of bundling in the efficient provision of public goods. We ask a simple question: is there an efficiency rationale to provide unrelated public goods in bundles rather than separately; and if so, why?

We consider an environment with m excludable public goods and a numeraire private good.[1] Each consumer is characterized by a valuation for each of the public goods. The willingness to pay for any subset of the goods is assumed to be the sum of the valuations for the individual goods in the subset. This assumption rules out bundling arising from complementarities in the utility function. Similarly, the cost of provision for each good is independent of which other goods are provided. These separability assumptions on valuations and costs imply that the informationally unconstrained efficient mechanism provides a public good if and only if the sum of all agents' valuations for that particular good exceeds its provision cost, and excludes no consumer from usage. Under perfect information there is thus no role for bundling.

This paper departs from the perfect information assumption, and assumes that preferences are

[1]The term "excludable public goods" refers to a good which is fully non-rival, but where it is possible to costlessly exclude any consumer from usage.

private information to the individuals. The provision mechanism must therefore be constructed so that truthful revelation of preferences is consistent with equilibrium. Agents may also freely choose whether to participate in the mechanism, and the provision mechanism must be self-financing. Finally, we assume that the preference parameters are stochastically independent across individuals. Under these restrictions, the (non-bundling) perfect information social optimum can no longer be implemented.[2]

Ruling out trivial cases, use exclusions are always active in the constrained efficient mechanism. Indeed, if the economy is large, use exclusions are essentially the only instruments available to induce consumers to contribute a non-negligible amount to the public goods.

To gain intuition, we first consider the case where the number of public goods is large, and where the valuation for all goods are stochastically independent. In this case, a pure bundling mechanism can approximate the first best if both the number of goods and the number of consumers are large. This is because the valuation for the average good in the bundle converges in probability to its expectation. If necessary, the designer can therefore extract almost the full surplus from each agent. This implies that, while the threat of exclusion is still what supports the incentives, it is possible to exclude an agent with arbitrarily small probability and still raise enough revenue to provide all goods. Finally, a large number of agents are needed for approximate efficiency because it is only in a large economy that ex ante information is sufficient for an efficient provision decision. If the number of agents are small, there will be a significant probability that a particular good should not be provided at all, a consideration that disappears with many agents.

Intuitively, the desirability of bundling in the many good case comes from the fact that bundling reduces the variance in the distribution of valuations. Whether goods are public or private is irrelevant for this. However, unlike the standard setup with private goods, non-rivalness in consumption means that the society can give access to all goods at no additional costs. The desirability of pure bundling thus relies crucially on the public good assumption.[3]

Next, we turn to 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 consumer valuations. While this is obviously a very special case, the results are suggestive, and the methodology

[2] 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 [6] can be used to implement the efficient outcome.

[3] Armstrong [3] considers a similar many good exercise for private goods. Due to similar law of large numbers reasoning, a monopolist can extract almost the full consumer surplus. The mechanism is a two part tariff, where consumers can pay a fixed fee for the right to purchase any good at marginal cost.

may be useful for more general (symmetric) multidimensional screening problems.

There is an element of bundling in the constrained efficient mechanism for almost all parametrizations of the model. This should be expected. We know from McAfee et al [13] that introducing the bundling instrument increases the profits for a monopolist that is restricted to fixed-price mechanisms. By results in Norman [17] we also know that, in the case with a single good, the constrained efficient mechanism is near a fixed price mechanism of the form considered by McAfee et al [13]. Finally, the (single-dimensional) constrained welfare problem has a Lagrangian characterization (see Hellwig [9] and Norman [17]). 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 insight in McAfee et al [13] 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 for. 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.

It is important to note that, while the existing literature on bundling in private goods focuses on how bundling relaxes the informational constraints and improves sellers' *revenue,* we derive a *constrained efficient* mechanism that involve bundling in the public good setting. Under the typical assumption in the private good bundling literature that goods are produced at constant (or increasing) marginal costs, bundling may enhance revenue, but will be dominated by marginal cost pricing in terms of social efficiency.

The remainder of the paper is structured as follows. Section 2 presents the model and some characterization results to be used later. In Section 3 we consider the case with a large number of goods. Section 4 introduces the special case when valuations are binary and demonstrates by example that a (pure) bundling mechanism may improve efficiency. Section 5 characterizes the optimal mechanism for this special case, and compares our characterization with existing results in the literature, and Section 6 concludes. All proofs are collected in the Appendix.

2 The Model

There are m *excludable* public goods, labeled by $j \in \mathcal{J} = \{1, ..., m\}$ and n consumers, indexed by $i \in \mathcal{I} = \{1, ..., n\}$.[4] All public goods are indivisible projects, and the cost of providing good j, denoted by $C^j(n)$, is independent of which of the other goods are provided. Notice here that n is the size of the economy and not the number of users, so all goods are fully non-rival. The rationale for indexing cost by the number of agents is to be able to analyze large economies, which makes it necessary to normalize per capita costs to avoid trivializing the provision problem. 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 particular economic interpretation, it is best viewed as a way to ensure that the provision problem remains "significant" also with many agents.

Consumer i is fully 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 obviously imagine more general utility functions than (1), but the linear formulation (which is also used by Adams and Yellen [1], McAfee et al [13], and Manelli and Vincent [12]) has the advantage that it rules out bundling due to complementarities in preferences.

The preference vector θ_i is private information to the agent, and we assume that preferences are independently and identically distributed across agents. We denote by F the joint cumulative distribution over θ_i. For brevity of notation, we let $\theta \equiv (\theta_1, ..., \theta_n) \in \mathbf{\Theta} \equiv (\Theta)^n$, 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.

2.1 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

[4] There are two reasons for allowing use exclusions. It allows us to consider large economies, which aids tractability. Moreover, it allows for a more intuitive form of bundling, since different consumers can consume different bundles when exclusions are possible.

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 simply a map from $\mathbf{\Theta}$ to A. We represent a randomized mechanism in analogy with the representation of mixed strategies in Aumann [4]. That is, let $\Xi = [0,1]$, 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 $\mathcal{G} : \mathbf{\Theta} \times \Xi \to A$. A conceptual advantage of this formalization of a random mechanism is that it allows for a useful decomposition.[5] That is, we may write $\mathcal{G}$ as a $(2m+1)$-tuple, $\mathcal{G} = (\{\zeta^j\}_{j\in\mathcal{J}}, \{\omega^j\}_{j\in\mathcal{J}}, \tau)$ where,

$$\begin{aligned} \zeta^j &: \mathbf{\Theta} \times \Xi \to \{0,1\} \\ \omega^j &: \mathbf{\Theta} \times \Xi \to \{0,1\}^n \\ \tau &: \mathbf{\Theta} \to \mathrm{R}^n. \end{aligned} \tag{3}$$

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 referred to as 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 referred to as the *cost-sharing rules,* where $\tau_i(\theta)$ is the transfer from agent i to the mechanism designer given announced valuations θ. In principle, transfers could also be randomized, but, agents are risk neutral with respect to transfers, so there are no gains from this. The pure transfer rule in (3) is therefore without loss of generality.

Because of the separability of the provision costs and the linear utility functions in (1), the ex post efficient provision and inclusion rules are simple: provide public good j if and only if $\sum_{i\in\mathcal{I}} \theta_i^j \geq C^j(n)$ and never exclude any consumer from usage when the good is provided. This is exactly as if each good j were the only public good.

2.2 The Design Problem

Utility is transferable, so we can characterize the constrained efficient allocation rules as the solution to a planning problem. A fictitious social planner seeks to maximize total surplus in the

[5] Because A is finite, there are no technical reasons for choosing this representation. It is chosen only because it generates more convenient notation than either the "natural" representation or the "distributional approach" of Milgrom and Weber [14].

economy, subject incentive compatibility, feasibility, and participation constraints. Let E_{-i} denote the expectation operator with respect to (θ_{-i}, ϑ). *Incentive compatibility*, that is, the requirement that truth-telling is a Bayesian Nash equilibrium in the revelation game induced by $\mathcal{G}$, requires that

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)\theta_i^j - \tau_i(\theta)\right] \geq \mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j(\widehat{\theta}_i,\theta_{-i},\vartheta)\omega_i^j(\widehat{\theta}_i,\theta_{-i},\vartheta)\theta_i^j - \tau_i(\widehat{\theta}_i,\theta_{-i})\right]$$
$$\forall i \in \mathcal{I}, \theta \in \mathbf{\Theta}, \widehat{\theta}_i \in \Theta. \tag{4}$$

Next, we require that the project be self-financing. For simplicity, this is imposed as an *ex ante balanced-budget constraint:*[6]

$$\mathrm{E}\left(\sum_{i\in\mathcal{I}} \tau_i(\theta) - \sum_{j\in\mathcal{J}} \zeta^j(\theta,\vartheta)\, C^j(n)\right) \geq 0. \tag{5}$$

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,

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)\theta_i^j - \tau_i(\theta)\right] \geq 0, \qquad \forall i \in \mathcal{I}, \theta_i \in \Theta. \tag{6}$$

A mechanism is *incentive feasible* if it satisfies (4), (5) and (6). A mechanism is *constrained efficient* if it maximizes the expected social surplus,

$$\sum_{j\in\mathcal{J}} \mathrm{E}\zeta^j(\theta,\vartheta)\left[\sum_{i\in\mathcal{I}} \omega_i^j(\theta,\vartheta)\theta_i^j - C^j(n)\right], \tag{7}$$

over all incentive feasible mechanisms.

All these constraints are noncontroversial if the design problem is interpreted as a private bargaining agreement, but, if the goods are government provided, the participation constraints (6) 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 (ignoring that the reservation utility should then be endogenous).[7]

[6] The ex ante constraint (5) is literally relevant only when the designer can access fair insurance market against budget deficits. However, adapting standard arguments (see Mailath and Postlewaite [11] and Cramton et al [5]), one can show that any allocation implementable with transfers satisfying (5) 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 one or more of the agents in the economy.

[7] Another defense of imposing voluntary participation in the context of government provided goods is to view this as a reduced form for inequality aversion of the planner. See Hellwig [8].

If preferences are observable, or if either (5) or (6) are ignored, the (non-bundling) ex post efficient mechanism is implementable. As discussed above, the ex post efficient provision rule for good j ignores the valuations for all other goods and no one is excluded. The ex post efficient rule is therefore implementable if and only if a single non-excludable public good can be efficiently provided under (4), (5) or (6). But, this is the exact setup of Mailath and Postlewaite [11], and from their adaption of Myerson and Satterthwaite's [16] impossibility result we know that the first best efficiency is only possible in trivial cases. Our setup is therefore a second best problem.

2.3 Simple Anonymous Mechanisms

To simplify the problem, we first exploit the fact that all control variables enter linearly in both the constraints and the objective function and that the problem is symmetric. 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* $j \in \mathcal{J}$,

$$\begin{aligned} \rho^j &: \boldsymbol{\Theta} \to [0,1] \\ \eta^j &: \Theta \to [0,1] \\ t &: \Theta \to \mathrm{R}, \end{aligned} \tag{8}$$

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 (8) relative to (3): (1). Both the inclusion and the transfer rules are the same for all agents; (2). Conditional on θ, the provision probability $\rho^j(\theta)$ is stochastically independent from all other provision probabilities, $\left\{\rho^k(\theta)\right\}_{k\in\mathcal{J}\setminus\{i\}}$, and all inclusion probabilities; (3). The inclusion and transfer rules for any agent i are independent of the realization of θ_{-i}; (4). 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 (8) still allows asymmetric treatment of agents in the provision rules. 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* $\rho^j(\theta) = \rho^j\left(\theta'\right)$ *for every* $j \in \mathcal{J}$, *every* $\theta \in \boldsymbol{\Theta}$, *and every* $\theta' \in \boldsymbol{\Theta}$ *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 (8) that generates the same social surplus.*

Consequently, the remainder of this paper only considers simple anonymous mechanisms. The intuition for why this class of mechanisms are sufficient is simple. Because of the risk neutrality, agents only care about the perceived probability of consuming each public good and the expected transfer. Therefore, there is nothing to gain from making transfers and inclusion probabilities functions of θ_{-i} , or by making inclusion and provision rules conditionally dependent. Mechanisms of the form (8) are therefore sufficient. Moreover, from any non-anonymous incentive feasible mechanism, one can always create a new incentive feasible mechanism that generates the same social surplus by permuting the roles of the agents. There are $n!$ permuted mechanisms, and from these we can create an anonymous incentive feasible mechanism that generates the same surplus by averaging over the $n!$ permutations.[8]

2.4 Symmetric Treatment of the Goods

Our next result, on which we rely heavily in Sections 4 and 5, 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 θ_i is an *exchangeable* random variable, that is $F(\theta_i) = F(\theta_i')$ 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.

The notion of symmetric mechanisms is intuitive, but we nevertheless provide a formal definition for clarity. 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. $t\left(\theta_i^P\right) = t(\theta_i)$.

[8] The exact argument is slightly more complex than simply randomizing with equal probabilities over the $n!$ permutations. The reason is that inclusion and provision probabilities are potentially correlated since they both depend on $\theta_i \in \Theta$.

It is worth pointing out that, in defining the symmetric mechanism, we require that the same permutation of goods 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 $\Theta = \{(h,h),(h,l),(l,h),(l,l)\}$. One type profile in Θ is $\theta = ((h,l),(l,h))$. Applying the only non-identity permutation of the goods to all agents generates a type profile $((l,h),(h,l))$. 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.[9]

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 an simple anonymous and symmetric incentive feasible mechanism that generates the same surplus as g.*

The idea of the proof is similar to that of Proposition 1, except that now it is the role of the goods that are permuted. For concreteness, consider the case with two goods. Suppose that there is an original mechanism, which possibly treats good 1 and 2 differently. We can reverse the role of the goods and obtain an alternative mechanism that generates the same surplus. A symmetric mechanism can be obtained by averaging over the original and the reversed mechanism.[10] One can show that it is also incentive feasible and generates the same surplus as the original mechanism. Proposition 2 generalizes this procedure by permuting the roles of the goods ($m!$ possibilities) and creating a symmetric mechanism by averaging over these.

3 The Case with Many Independent Goods

A relatively straightforward case is when both the number of goods and the number of agents are large. For reasons familiar from the multidimensional screening literature, finding an exactly optimal mechanism is an intractable problem. However, using reasoning similar to Armstrong [3], one can construct an approximately optimal mechanism for the case when both n and m are

[9] 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.

[10] 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.

sufficiently large.[11] This approximately optimal mechanism is a pure bundling mechanism.

The ex post efficient rule is to provide good j if and only if $\sum_{i=1}^{n}\theta_i^j \geq C^j(n)$ and exclude nobody from usage. What is the provision probability for good j under the ex post efficient rule? Assuming that there exists finite numbers μ and σ^2 such that $\mathrm{E}\theta_i^j \leq \mu$ and $\mathrm{Var}\theta_i^j \leq \sigma^2$ for each j, we can appeal to law of large numbers reasoning to get a simple answer.[12] Given these regularity conditions, $\sum_{i=1}^{n}\theta_i^j/n$ converges in probability to $\mathrm{E}\theta_i^j$, and the ex post efficient provision probability converges to either zero or one depending on the relation between $\mathrm{E}\theta_i^j$ and $C^j(n)/n$:

Lemma 1 *Suppose there are finite numbers μ and σ^2 such that $\mathrm{E}\theta_i^j \leq \mu$ and $\mathrm{Var}\theta_i^j \leq \sigma^2$ for all j.*

1. *If there exists N and $\delta > 0$ such that $\mathrm{E}\theta_i^j - C^j(n)/n \geq \delta$ when $n \geq N$ for all j, then $\lim_{n\to\infty}\Pr\left[\sum_{i=1}^{n}\theta_i^j \geq C^j(n)\right] = 1$;*

2. *If there exists N and $\delta > 0$ such that $C^j(n)/n - \mathrm{E}\theta_i^j \geq \delta$ when $n \geq N$ for all j, then $\lim_{n\to\infty}\Pr\left[\sum_{i=1}^{n}\theta_i^j \geq C^j(n)\right] = 0$.*

It is not hard to implement "never provide" , so only the first case is interesting. We therefore assume that the first case applies for all j for the remainder of this section.[13] Also note that if $C^j(n)/n$ has a limit, the only case not covered in Lemma 1 is when $\lim_{n\to\infty}C^j(n)/n = \mathrm{E}\theta_i^j$.

Let $\varepsilon > 0$ and consider the simple anonymous mechanism $\widehat{g} = \left(\{\widehat{\rho}^j\}_{j=1}^m, \{\widehat{\eta}^j\}_{j=1}^m, \widehat{t}\right)$, where

$$\begin{aligned}
\widehat{\rho}^j(\theta) &= 1 \text{ for all } j \text{ and all } \theta \in \mathbf{\Theta} \\
\widehat{\eta}^j(\theta_i) &= \begin{cases} 1 & \text{if } \sum_j \theta_i^j \geq \sum_j \frac{C_j(n)}{n} + \varepsilon m \\ 0 & \text{if } \sum_j \theta_i^j < \sum_j \frac{C_j(n)}{n} + \varepsilon m \end{cases} \qquad j \in \{1, ..., m\} \\
\widehat{t}(\theta_i) &= \begin{cases} \sum_j \frac{C_j(n)}{n} + \varepsilon m & \text{if } \sum_j \theta_i^j \geq \sum_j \frac{C_j(n)}{n} + \varepsilon m \\ 0 & \text{if } \sum_j \theta_i^j < \sum_j \frac{C_j(n)}{n} + \varepsilon m. \end{cases}
\end{aligned} \tag{9}$$

The mechanism in (9) is a *pure bundling mechanism.* While expressed as a direct revelation mechanism, we can interpret it as a fixed mechanism where the full bundle is offered to anyone willing to pay price $\sum_j \frac{C_j(n)}{n} + \varepsilon m$. Clearly, truth-telling is a dominant strategy and the participation

[11] While containing no formal large numbers analysis, similar reasoning can also be found in Dana [7]. Jackson and Sonnenschein [10] also show in a related setting that the welfare costs of incentive constraints completely disappear when a large number of decisions are linked.

[12] A sufficient condition for the existence of the bounds μ and σ^2 is that there exists an interval $[a, b]$ such that $\theta_i^j \in [a, b]$ for all j.

[13] If the first best probability of provision converges to one for some goods and zero for others, the analysis still applies as long as there are sufficiently many goods that should be provided in a large economy according to the ex post efficient rule. Goods for which the first best probability of provision converges to zero may simply be dropped from the bundle and the rest of the analysis carries over.

constraints are satisfied given such a mechanism. The only questions are: (1). when does (9) satisfy the feasibility constraint (5)? (2). what are the optimality properties of (9)?

Proposition 3 *Suppose uniform bounds* $\mu < \infty, \sigma^2 < \infty, \delta > 0$ *exist such that* $\mathrm{E}\theta_i^j \leq \mu, \mathrm{VAR}\theta_i^j \leq \sigma^2$ *for every* j*, and* $\mathrm{E}\theta_i^j - C^j(n)/n \geq \delta$ *for every* j *and* $n \geq N$*. Then:*
(1) for any $\varepsilon \in (0, \delta)$ *there exists* $M < \infty$ *such that* $\widehat{g}$ *is incentive feasible for any* n *and any* m *such that* $n \geq N$ *and* $m \geq M$*,*
(2) for every $\varepsilon \in (0, \delta)$ *there exists* N_ε *and* M *(independent of* ε*) such that the difference in per capita surplus between the ex post optimal mechanism and mechanism* $\widehat{g}$ *is less than* ε *for any economy with* $n \geq N_\varepsilon$ *and* $m \geq M$*.*

In words, the simple pure bundling mechanism in (9) can approximate the outcome of the informationally unconstrained efficient mechanism arbitrarily well, provided that the number of goods and the number of consumers are both sufficiently large. In contrast, if the bundling instrument is not available, the probability of exclusion is always bounded away from zero for all agents (see Norman [17]). The result thus illustrates that (pure) commodity bundling improves economic efficiency in large economies with many public goods.

The intuition for the above "double infinity" (n and m both go to infinity) asymptotic results is as follows. By selling usage of the goods only as a bundle, a consumer will buy the good if and only if the average valuation exceeds the ratio of the price over the number of goods. The average valuation converges almost surely to the expectation as the number of goods approach infinity, implying that the probability of excluding an agent can be made negligible even if the "per good" bundle price is near the expected average valuation. Hence, what is crucial for the implementability of $\widehat{g}$ is the number of goods. Indeed, if $C^j(n) = c^j n$ for each j and n, the number of agents is completely irrelevant for part 1 of the result. However, the number of agents play a crucial role for the *desirability* of the pure bundling scheme. With a small number of agents, there is a significant probability that a particular good should not be provided. If n is large, this probability is negligible, implying that the pure bundling provision rule is near the efficient provision rule.

Independence among the elements in θ_i is of course a strong assumption. In many situations it seems reasonable that there are correlations, for example due to all elements in θ_i being correlated with wealth, age, or other "background variables." But, if the variables that induce the correlation are observable, this can easily been taken care of. That is, if (θ_i, z) follows some joint distribution F, what is needed is that the elements in θ_i are conditionally independent given any realized z.[14]

[14] If for each i, F_i is the cdf over θ_i and $\{F_i\}_{i=1}^n$ are independent, mechanism (9) still leads to approximate efficiency.

3.1 What is Special About Public Goods?

It is instructive to compare our large numbers analysis with Armstrong [3], who studies of a multi-product monopolist selling private goods. While Armstrong is concerned with profit maximization, he obtains an approximate full surplus extraction result when there is a large number of stochastically independent goods, so the allocation is almost first best. The mechanism that achieves this is a two part tariff. In essence, the monopolist sells the right to purchase goods at marginal cost. With a large number of goods, the consumer surplus for the average good is near the expected consumer surplus, so the monopolist can extract almost the full surplus.

A pure bundling scheme would not do particularly well when selling private goods. Unless the marginal cost is zero, pure bundling would lead to a large numbers of goods being produced for users who value the goods below cost. For the public goods case, however, the non-rivalness in consumption means that this concern vanishes, and bundling can almost implement the first best.

4 The Model with Binary Valuations

We now turn to a case for which we can characterize the constrained efficient mechanism exactly. Assume that there are two public goods, and that the valuation for good j can either be "high" ($\theta_i^j = h$) or "low" ($\theta_i^j = l$). The individual type space is thus $\Theta = \{(h,h), (h,l), (l,h), (l,l)\}$. For notational brevity we henceforth write $\theta_i = hh$ instead of (h,h), $\theta_i = hl$ instead of (h,l), and so on. To facilitate comparisons with the non-bundling benchmark, we also assume that θ_i^1 and θ_i^2 are independent with $\alpha = \Pr[\theta_i^1 = h] = \Pr[\theta_i^2 = h] \in (0,1)$, implying that the probability distribution F over Θ is:[15]

$$\left\{\alpha^2, \alpha(1-\alpha), \alpha(1-\alpha), (1-\alpha)^2\right\}.$$

Finally, we assume that costs are given by $C^1(n) = C^2(n) = cn$. The most important simplification here is that costs are the same for both goods. Together with the symmetric type space, this implies that we can appeal to Proposition 2 and restrict attention to symmetric mechanisms. Keeping the per capita costs constant simplifies notation, but is not necessary.

If $h \leq c$, "never provide any good" is ex post optimal, which can be trivially implemented. Symmetrically, if $l \geq c$, "always provide both goods" is ex post optimal and can be implemented by charging a constant tax equal to $2c$. We therefore maintain the assumption that $l < c < h$ in order to keep the problem interesting.

[15] Independence across agents is a crucial assumption, but independence across goods is only for ease of comparison with the no-bundling case. The analysis extends with minor modifications with a probability distribution of the form $\left\{\sigma_{hh}, \frac{\sigma_m}{2}, \frac{\sigma_m}{2}, \sigma_{ll}\right\}$, where σ_m is the probability of a "mixed type".

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

$$\mathcal{X}_n = \left\{ x \in \{0, ..., n\}^4 : x_{hh} + x_{hl} + x_{lh} + x_{ll} = n \right\}. \tag{10}$$

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 on form

$$\mathcal{M} = \left(\left\{ \rho^j, \eta^j \right\}_{j=1,2}, t \right), j \in \{1, 2\}, \tag{11}$$

where $\rho^j : \mathcal{X}_n \to [0,1], \eta^j \equiv (\eta^j_{hh}, \eta^j_{hl}, \eta^j_{lh}, \eta^j_{ll}) \in [0,1]^4$ and $t \equiv (t_{hh}, t_{hl}, t_{lh}, t_{ll}) \in \mathrm{R}^4$ satisfy

$$\rho^1 (x_{hh}, x_{hl}, x_{lh}, x_{ll}) = \rho^2 (x_{hh}, x_{lh}, x_{hl}, x_{ll}), \tag{12a}$$

$$\eta^1_{hh} = \eta^2_{hh}, \eta^1_{hl} = \eta^2_{lh}, \eta^1_{lh} = \eta^2_{hl}, \eta^1_{ll} = \eta^2_{ll}, t_{hl} = t_{lh}. \tag{12b}$$

4.1 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 index the mechanism by n. With some abuse of notation, we write a separate provision mechanism for good j in an economy of size n as a triple $(\rho^j_n, \eta^j_n, t^j_n)$, where $\rho^j_n : \{1, ..., n\} \to [0,1]$ and $\rho^j_n(\kappa)$ denotes the probability of provision if κ agents announce a high valuation for good j; $\eta^j_n \in [0,1]$ is the inclusion probability for type l and $t^j_n = (t^j_n(h), t^j_n(l))$ are the transfers.[16]

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 (4), (5) and (6) in Section 2.2 one obtains the following characterization of the constrained optimal separate provision mechanism:

Proposition 4 *Consider a sequence of economies of size* $\{n\}_{n=1}^{\infty}$. *Then,*
(1) if $\alpha h < c, \lim_{n\to\infty} \mathrm{E}\rho^j_n(\kappa) = 0$ *for any sequence of feasible separate provision mechanisms* $\left\{ \rho^j_n, \eta^j_n, t^j_n \right\}$;

[16] In principle, use exclusions of type h agents is of course also feasible. However, such exclusions never occur in an optimal mechanism, since excluding type h tightens the downwards incentive constraint for h.

(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 [17] and the ideas are very similar.[17] Instead of a formal proof, we only provide a heuristic explanation of the result.[18] The key idea is that the incentive constraint

$$\mathrm{E}\left[\rho_n^{*j}(\kappa) | \theta_i^j = h\right] h - t_n^{*j}(h) \geq \mathrm{E}\left[\rho_n^{*j}(\kappa) | \theta_i^j = l\right] \eta_n^{*j} h - t_n^{*j}(l), \tag{13}$$

may be replaced by

$$\mathrm{E}\rho_n^{*j}(\kappa) h - t_n^{*j}(h) \geq \mathrm{E}\rho_n^{*j}(\kappa) \eta_n^{*j} h - t_n^{*j}(l), \tag{14}$$

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, and (again ignoring the effects of being pivotal) this implies that $t_n^{*j}(l) \approx \mathrm{E}\rho_n^{*j}(\kappa) \eta_n^{*j} l$. Because (13) binds in the optimal mechanism, budget balance requires that, approximately,

$$\begin{aligned} \mathrm{E}\rho_n^{*j}(\kappa) c &= \alpha t_n^{*j}(h) + (1-\alpha) t_n^{*j}(l) \approx \alpha \left[t_n^{*j}(l) + \mathrm{E}\rho_n^{*j}(\kappa) h \left(1 - \eta_n^{*j}\right)\right] + (1-\alpha) t_n^{*j}(l) \\ &= t_n^{*j}(l) + \alpha \mathrm{E}\rho_n^{*j}(\kappa) h \left(1 - \eta_n^{*j}\right) = \mathrm{E}\rho_n^{*j}(\kappa) \widehat{\eta}_n^j l + \mathrm{E}\rho_n^{*j}(\kappa) \alpha h \left(1 - \eta_n^{*j}\right). \end{aligned} \tag{15}$$

Hence, $\eta_n^{*j} \approx (\alpha h - c) / (\alpha h - l)$ follows from (15). Inspecting (15), it follows that $\lim_{n\to\infty} \mathrm{E}\rho_n^{*j}(\kappa) = 0$ if $\alpha h < c$ (since $l < c$ 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 4, 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

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

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

[17] Strictly speaking, Proposition 4 is not a special case of the results in Norman [17], which deals with continuous distributions satisfying the "increasing virtual valuation condition" familiar from Myerson [15] and others. Since continuous approximations of discrete distributions violate this regularity condition, there are some qualitative differences between the binary case and the "regular" continuously distributed case. In particular, the solution to the binary case will generically involve randomizations.

[18] Details available on request from the authors.

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 4 is bounded away from first best efficiency. First of all, the asymptotic provision probability is zero when $\alpha h < c$ while efficiency requires that the public good be provided 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, even though the public good is provided asymptotically with probability one.

4.2 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

$$\begin{aligned} t_{hh} &= t_{hl} = t_{lh} = h + l, \text{ and } t_{ll} = 0 \\ \eta_{hh} &= \eta_{hl} = \eta_{lh} = 1, \text{ and } \eta_{ll} = 0 \\ \rho^1(x) &= \rho^2(x) = 1 \text{ for all } x \in \mathcal{X}_n. \end{aligned} \tag{16}$$

That is, type-hh and type-hl agents are taxed the willingness to pay of the mixed type and consume both goods for sure. Type-ll pays nothing and is excluded from usage from both goods.

All incentive and participation constraints are trivially satisfied by mechanism (16). The only question is thus whether the feasibility constraint (5) is satisfied, that is, if

$$\Pr[\{hh, hl, lh\}](h+l) = \alpha(2-\alpha)(h+l) \geq 2c, \tag{17}$$

holds. It is easy to show that:

Claim 1 *Given any $c > 0$ and $\alpha \in (0,1)$ there exists pairs (h,l) with $h > c > l$ such that (17) is satisfied, where at the same time $ah < c$.*

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 (16). The proposed bundling mechanism therefore improves efficiency. The construction of the values of h and l for any $c > 0$ for which (16) outperforms the best separate provision mechanism is depicted in Figure 1. The intuition for the improvement of bundling mechanism is as follows. The revenue maximizing separate provision mechanism is to include only high valuation types. Hence, a fraction

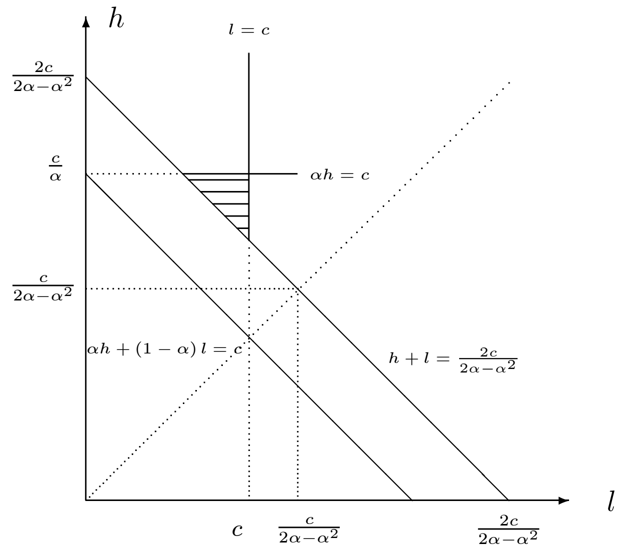

Figure 1: The Bundling Mechanism Outperforms Optimal Non-bundling Mechanism in the Shaded Region.

α of the agents contribute towards each good. In the bundling mechanism (16), only a fraction $(1-\alpha)^2$ are excluded. While the contribution per agent decreases, the total revenue increases if l is sufficiently close to c.

5 The Constrained Optimal Mechanism

5.1 The Mechanism Design Problem

For the binary model described in the previous section, we now set up the design problem to maximize social surplus (7) subject to the incentive compatibility constraints in (4), the feasibility constraint (5) and the participation constraints (6) in a tractable form.

The most involved part of the optimization problem is the provision rule. This is difficult to deal with because $\rho^j(x)$ is weighted by the ex ante probability that x occurs in the objective function to the problem, while the relevant probabilities in the constraints are conditional probabilities. To deal with this, we need to be explicit about the (multinomial) probability distribution of x, in order to eventually be able to link the unconditional and conditional probabilities. 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 $\left(n, \alpha^2, \alpha(1-\alpha), \alpha(1-\alpha), (1-\alpha)^2\right)$.

There are 12 incentive constraints to be satisfied. However, 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". We conjecture that only downwards incentive constraints are relevant and will therefore ignore all upwards constraints as well as the constraints between type hl and lh. Once the solution to the relaxed problem is fully characterized, we will verify that the other omitted constraints are satisfied. Finally, it is easy to check that if type-hh is better off announcing her true type than announcing hl, and type-hl is better off announcing her true type than announcing ll, then there are no incentives for type-hh to announce ll. Together with the symmetry of the mechanism (12), we are thus left with two distinct incentive constraints:

$$\begin{aligned} 2\eta^1_{hh} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll})\,h - t_{hh} \;\geq \\ \eta^1_{hl} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}+1, x_{lh}, x_{ll})h \\ +\eta^1_{lh} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}+1, x_{ll})h - t_{hl}, \end{aligned} \tag{18a}$$

$$\begin{aligned} \eta^1_{hl} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}+1, x_{lh}, x_{ll})h \\ +\eta^1_{lh} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}+1, x_{ll})l - t_{hl} \;\geq \\ \eta^1_{ll} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\,(h+l) - t_{ll}, \end{aligned} \tag{18b}$$

where (18a) states that type-hh agents do not have incentives to mis-report as type hl; and (18b) states that type-hl agents do not have incentives to mis-report as type ll.

Given that all downward incentive constraints and the participation constraint for type ll are fulfilled, it follows by a standard argument that the participation constraints for types hh, hl and lh are also fulfilled. The only relevant participation constraint is thus

$$2\eta^1_{ll} \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)l - t_{ll} \;\geq\; 0. \tag{19}$$

Finally, the budget balance constraint can be simplified considerably due to the simple transfer schemes and the constant per capita costs. That is, using the symmetry of (12) and breaking out n from (5), we can express the feasibility constraint in per capita form as

$$\alpha^2 t_{hh} + \alpha(1-\alpha)\,2t_{hl} + (1-\alpha)^2\, t_{ll} - 2c \sum_{x\in\mathcal{X}_n} \mathbf{a}_n(x)\,\rho^1(x) \geq 0. \tag{20}$$

Again using the symmetry of (12), we can drop one of the goods, and express the relaxed programming problem as:[19]

$$\max_{\{\rho^1,\eta^1,t\}} 2\sum_{x\in\mathcal{X}_n} \mathbf{a}_n(x)\,\rho^1(x)\left[\frac{\left(\eta^1_{hh}x_{hh} + \eta^1_{hl}x_{hl}\right)h + \left(\eta^1_{lh}x_{lh} + \eta^1_{ll}x_{ll}\right)l}{n} - c\right] \tag{21}$$

s.t. (18a)-(18b), (19) and (20),

$$\eta^1_{\theta_i} \geq 0, 1-\eta^1_{\theta_i} \geq 0 \text{ for each } \theta_i \in \Theta, \tag{22}$$

$$\rho^1(x) \geq 0, 1-\rho^1(x) \geq 0 \text{ for each } x \in \mathcal{X}_n, \tag{23}$$

[19]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.

where the social planner's objective function is written in per capita form.

Lemma 2 *There exists at least one optimal solution to (21).*

The proof is standard by first compactifying the constraint set and then applying Weierstrass Theorem. It can be shown that Slater's condition for constraint qualification holds, so the Kuhn-Tucker conditions are necessary for an optimum. Since a solution to (21) exists, these first order conditions therefore provide a characterization of the optimal mechanism, provided that the constraints that we ignored when formulating (21) are satisfied at the candidate solution.

5.2 Relationship Between Multipliers

Taxes enter linearly into all constraints and are not constrained by boundaries. It is therefore convenient to begin the analysis by taking first order conditions with respect to t_θ. This allows us to express the multiplier of any other constraint as a linear scaling of the multiplier of the feasibility constraint.

The first order conditions with respect to $t = (t_{hh}, t_{hl}, t_{ll})$ are,

$$\begin{aligned} &(\text{w.r.t. } t_{hh}) && -\lambda_{hh} + \Lambda\alpha^2 = 0, \\ &(\text{w.r.t. } t_{hl}) && \lambda_{hh} + \lambda_{hl} + \Lambda 2\alpha(1-\alpha) = 0, \ , \\ &(\text{w.r.t. } t_{ll}) && \lambda_{hl} - \lambda_{ll} + \Lambda(1-\alpha)^2 = 0. \end{aligned} \tag{24}$$

where λ_{hh} and λ_{hl} are the multipliers associated with (incentive compatibility) constraints (18a) and (18b), and λ_{ll} is the multiplier associated with the (participation) constraint (19). Hence:

Lemma 3 *In any solution to (21), the multipliers* $(\lambda_{hh}, \lambda_{hl}, \lambda_{ll}, \Lambda)$ *satisfy:* $\lambda_{hh} = \alpha^2\Lambda, \lambda_{hl} = (2\alpha - \alpha^2)\Lambda,$ *and* $\lambda_{ll} = \Lambda$.

In all its simplicity, Lemma 3 is actually a key step in the solution of (21). Its role is similar to the characterization of incentive compatibility and individual rationality in terms of a single integral constraint in single-dimensional mechanism design problem (i.e., the approach in Myerson [15] and others). In multidimensional problems, it is impossible to collapse all constraints into a single constraint. Instead, Lemma 3 allows us to *indirectly* relate all optimality conditions to a single constraint. The analysis is thus very much as if an objective function is maximized subject to a single constraint.

5.3 Optimal Inclusion Rules

We now characterize the optimal inclusion rules η^1. 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

$$\begin{aligned} G(\Phi) &\equiv (1-\Phi)\,2l + \Phi\left[\frac{2\alpha-\alpha^2}{\alpha(1-\alpha)}l - \frac{\alpha^2}{\alpha(1-\alpha)}h\right], \\ H(\Phi) &\equiv (1-\Phi)\,2l + \Phi\left[\frac{2}{(1-\alpha)^2}l - \frac{2\alpha-\alpha^2}{(1-\alpha)^2}(h+l)\right]. \end{aligned} \tag{25}$$

The result is:

Lemma 4 *Let $\mathcal{M} = (\rho^1, \rho^2, \eta^1, \eta^2, t)$ be a symmetric solution to (21) and let $\Phi = \Lambda/(1+\Lambda)$, where Λ is the associated multiplier on the resource constraint. Also, suppose that $\mathrm{E}\left[\rho^j(x)\,|\theta_i\right] > 0$ for all $\theta_i \in \Theta$ and $j = 1, 2$. Then,*

1. $\eta^1_{hh} = \eta^2_{hh} = \eta^1_{hl} = \eta^2_{lh} = 1$;
2. $\eta^1_{lh} = \eta^2_{hl} = \begin{cases} 1 & \text{if } G(\Phi) > 0 \\ y \in [0,1] & \text{if } G(\Phi) = 0 \\ 0 & \text{if } G(\Phi) < 0; \end{cases}$
3. $\eta^1_{ll} = \eta^2_{ll} = \begin{cases} 1 & \text{if } H(\Phi) > 0 \\ y \in [0,1] & \text{if } H(\Phi) = 0 \\ 0 & \text{if } H(\Phi) < 0. \end{cases}$

To interpret the result, note that $\Phi = \Lambda/(1+\Lambda) \in [0,1]$, and $G(\Phi) \geq 0$ if and only if

$$\Phi\overbrace{\left[\alpha(2-\alpha)\,l - \alpha^2 h\right]}^{\text{Term 1}} + (1-\Phi)\overbrace{2\alpha(1-\alpha)\,l}^{\text{Term 2}} \geq 0. \tag{26}$$

To understand Term 1 in expression (26), consider two candidate inclusion rules. The first candidate is $\eta^1_{lh} = \eta^2_{hl} = \eta^1_{ll} = \eta^2_{ll} = 0$, and $\eta^1_{hh} = \eta^2_{hh} = \eta^1_{hl} = \eta^2_{lh} = 1$. 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 $2h \times \alpha^2 + h \times 2\alpha(1-\alpha) = 2\alpha h$ from each agent. The second candidate inclusion rule is $\eta^1_{lh} = \eta^2_{hl} = \eta^1_{hh} = \eta^2_{hh} = \eta^1_{hl} = \eta^2_{lh} = 1$ and $\eta^1_{ll} = \eta^2_{ll} = 0$. 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 at least $\left[\alpha^2 + 2\alpha(1-\alpha)\right](h+l) = \alpha(2-\alpha)(h+l)$. The change in revenue if increasing η^1_{lh} and η^2_{hl} from 0 to 1 is thus

$$\alpha(2-\alpha)(h+l) - \alpha 2h = \alpha(2-\alpha)\,l - \alpha^2 h, \tag{27}$$

which is Term 1 in expression (26). Term 2 in expression (26), $2\alpha(1-\alpha)l$, on the other hand, captures the marginal increase in per capita surplus from increasing η^1_{lh} and η^2_{hl} 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.

Clearly, if both Term 1 and Term 2 are positive, then both the social planner and monopolistic provider prefers setting $\eta^1_{lh} = \eta^2_{hl} = 1$. On the other hand, if Term 1 is negative, i.e. if $\alpha(2-\alpha)l < \alpha^2 h$, then some algebra on expression (26) shows that $G(\Phi) \geq 0$ if

$$\Phi \leq \Phi^*_{lh} = \frac{(1-\alpha)\,2l}{\alpha(h-l)}. \tag{28}$$

Clearly, $\Phi^*_{lh} > 0$ but $\Phi^*_{lh} < 1$ only when $\alpha(2-\alpha)l < \alpha^2 h$. That is, when there is a conflict of interest between the surplus maximizing social planner and a revenue maximizing monopolistic provider, $\eta^1_{lh} = \eta^2_{hl} = 1$ will be optimal only when Λ, or the shadow price of resources, is sufficiently low. To summarize, item 2 of Lemma 4 could be restated as: there exists a critical value $\Phi^*_{lh} \in (0,1)$ such that $G(\Phi) \geq 0$ if and only if $\Phi \leq \Phi^*_{lh}$.

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

$$(1-\Phi)(1-\alpha)^2\,2l + \Phi\left[2l - \alpha(2-\alpha)(h+l)\right] \geq 0. \tag{29}$$

The term $(1-\alpha)^2\,2l$ is the gain in social surplus when η^1_{ll} and η^2_{ll} are increased from 0 to 1; and the term $2l - \alpha(2-\alpha)(h+l)$ is the revenue effect of 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 $2l - \alpha(2-\alpha)(h+l) > 0$, the $H(\Phi) > 0$ for sure and $\eta^1_{ll} = \eta^2_{ll} = 1$ is optimal. Otherwise, $H(\Phi) \geq 0$ if and only if

$$\Phi \leq \Phi^*_{ll} = \frac{(1-\alpha)^2\,2l}{\alpha(2-\alpha)(h-l)}. \tag{30}$$

Note

$$\frac{\Phi^*_{lh}}{\Phi^*_{ll}} = \frac{2-\alpha}{1-\alpha} > 1. \tag{31}$$

This implies that type-hl or type-lh agents are always "first in line" to get access to the good for which they have a low valuation in the following sense: if $\eta^1_{ll} = \eta^2_{ll} > 0$, then we know that $\Phi \leq \Phi^*_{ll} < \Phi^*_{lh}$, thus $\eta^1_{lh} = \eta^2_{hl} = 1$; symmetrically, if $\eta^1_{lh} = \eta^2_{hl} < 1$, then we know $\Phi \geq \Phi^*_{lh} > \Phi^*_{ll}$, then $\eta^1_{ll} = \eta^2_{ll} = 0$. We summarize the above discussion as:

Lemma 5 *Suppose that* $E\left[\rho^j(x)\,|\theta_i\right] > 0$ *for all* $\theta_i \in \Theta$ *and* $j = 1, 2$. *Let* $\Phi = \Lambda/(1+\Lambda)$. *There exists* $\Phi^*_{ll} < \Phi^*_{lh}$ *such that the optimal inclusion rule satisfies:*

1. *All agents with a high valuation for good j is included with probability one for using good j if it is provided;*

2. *If $\Phi < \Phi^*_{ll} < \Phi^*_{lh}$, then all agents get access to both public goods.*

3. *If $\Phi = \Phi^*_{ll} < \Phi^*_{lh}$, then $\eta^1_{ll} = \eta^2_{ll} \in [0,1]$ and $\eta^1_{lh} = \eta^2_{hl} = 1$.*

4. *If $\Phi^*_{ll} < \Phi < \Phi^*_{lh}$, then $\eta^1_{ll} = \eta^2_{ll} = 0$ and $\eta^1_{lh} = \eta^2_{hl} = 1$.*

5. *If $\Phi^*_{ll} < \Phi = \Phi^*_{lh}$, then $\eta^1_{ll} = \eta^2_{ll} = 0$ and $\eta^1_{lh} = \eta^2_{hl} \in [0,1]$*

6. *If $\Phi^*_{ll} < \Phi^*_{lh} < \Phi$, then $\eta^1_{ll} = \eta^2_{ll} = \eta^1_{lh} = \eta^2_{hl} = 0$*

While Λ is still unknown, we now possess a simple characterization of the optimal inclusions as a function of the still unknown multiplier on the resource constraint.

5.4 Optimal Provision Rules

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

$$\begin{aligned} Q^1\left(\frac{x}{n},\Phi\right) &\equiv \frac{x_{hh}}{n}h + \frac{x_{hl}}{n}h + \frac{x_{lh}}{n}\frac{\max\{0,G(\Phi)\}}{2} + \frac{x_{ll}}{n}\frac{\max\{0,H(\Phi)\}}{2} - c. \\ Q^2\left(\frac{x}{n},\Phi\right) &\equiv \frac{x_{hh}}{n}h + \frac{x_{lh}}{n}h + \frac{x_{hl}}{n}\frac{\max\{0,G(\Phi)\}}{2} + \frac{x_{ll}}{n}\max\frac{\{0,H(\Phi)\}}{2} - c. \end{aligned} \tag{32}$$

These functions have a natural interpretation. To see this, first consider the case where $\Phi = 0$, in which case [see definitions in (25)] $G(0) = H(0) = 2l$. 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.[20] We can thus think of $Q^j(x/n,\Phi)$ as a weighted average of social surplus and net revenue if the good is provided when the state is x.

The constrained optimal provision rule can be fully described in terms of these two functions:

Lemma 6 *Let $\mathcal{M}$ be an optimal solution to (21) and $\Phi = \Lambda/(1+\Lambda)$ where Λ is the multiplier associated with the constraint (20) 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

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

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.[21] In contrast, in the model without bundling, $\alpha h > c$ is the necessary and sufficient for asymptotic probability one provision. To see this, write $\mu = \left(\alpha^2, \alpha(1-\alpha), \alpha(1-\alpha), (1-\alpha)^2\right)$ as the asymptotic proportions of agents with different valuation combinations hh, hl, lh, and ll; and write $\Phi_n = \Lambda_n / (1+\Lambda_n)$ where Λ_n is the associated multiplier on the resource constraint in the optimal solution when the number of agents in the economy is n. By continuity of Q^j,

$$\begin{aligned}\lim_{n\to\infty} Q^j\left(\frac{x}{n}, \Phi_n\right) &= Q^j(\mu, \Phi) \\ &= \alpha h + \alpha(1-\alpha)\max\{0, G(\Phi)\} + (1-\alpha)^2 \max\{0, H(\Phi)\} - c \qquad (33)\end{aligned}$$

where $\Phi = \lim_{n\to\infty} \Phi_n$. Thus, $\alpha h > c$ is a sufficient condition for $Q^j(\mu, \Phi) > 0$ (and hence for asymptotic provision with probability 1).

5.5 The Main Result

In this section, we provide a full characterization of the asymptotic properties of a sequence of optimal mechanisms. We index the mechanisms by the size of the economy and write $\left\{\rho_n^j, \eta_n^j, t_n\right\}_{j=1}^2$, where $\rho_n^j : \mathcal{X}_n \to [0,1]$ is the provision rule for good j, and $\eta_n^1 = \left(\eta_n^1(lh), \eta_n^1(ll)\right)$ are the probabilities that type-lh and ll agents are allowed access to good 1 conditional on provision, and $\eta_n^2 = \left(\eta_n^2(hl), \eta_n^2(ll)\right)$ are the probabilities that type-hl and ll agents are allowed access to good 2 conditional on provision; and t_n is the transfer rule. Note that, by Lemma 4, the other types are included with probability 1 in any optimal mechanism. Our main result is:

Proposition 5 *Let $\left\{\rho_n^j, \eta_n^j, t_n\right\}_{j=1}^2$ be a sequence of optimal mechanism. Then, the following holds:*

1. *if* $\max\{2\alpha h, \alpha(2-\alpha)(h+l)\} > 2c$, *then* $\lim_{n\to\infty} \mathrm{E}\rho_n^j(x) \to 1$ *for* $j = 1, 2$;
2. *if* $\max\{2\alpha h, \alpha(2-\alpha)(h+l)\} < 2c$, *then* $\lim_{n\to\infty} \mathrm{E}\rho_n^j(x) \to 0$ *for* $j = 1, 2$;
3. *if* $\alpha(2-\alpha)(h+l) > 2c$, *then there exists* $N < \infty$ *such that* $\eta_n^1(lh) = \eta_n^2(hl) = 1$ *for every* $n \geq N$, $\eta_n^1(ll) = \eta_n^2(ll)$ *for every* n *and*

$$\lim_{n\to\infty} \eta_n^1(ll) = \lim_{n\to\infty} \eta_n^2(ll) = \eta_{ll}^*,$$

[21] Recall that in the example in Section 4.2, the proposed bundling mechanism achieves provision with probability one for cases when $\alpha h < c$.

Bundling╲Exclusion	No Exclusion	Exclusion
No Bundling	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite [11])	$\mathrm{E}\rho_n^{j*} \to 0$, if $\alpha h < c$ $\mathrm{E}\rho_n^{j*} \to 1$, if $\alpha h > c$ (Norman [17])
Bundling Allowed	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite [11])	$\mathrm{E}\rho_n^{j*} \to 0$, if $\max\{2\alpha h, \alpha(2-\alpha)(h+l)\} < 2c$; $\mathrm{E}\rho_n^{j*} \to 1$, if $\max\{2\alpha h, \alpha(2-\alpha)(h+l)\} > 2c$ (This Paper)

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

where

$$\eta_{ll}^* = \frac{\alpha(2-\alpha)(h+l) - 2c}{\alpha(2-\alpha)(h+l) - 2l} \in (0,1);$$

4. *If* $2\alpha h > 2c > \alpha(2-\alpha)(h+l)$*, then there exists* $N < \infty$ *such that* $\eta_n^1(ll) = \eta_n^2(ll) = 0$ *for all* $n \geq N$ *and* $\eta_n^1(lh) = \eta_n^2(hl)$ *for every* n *and*

$$\lim_{n\to\infty} \eta_n^1(lh) = \lim_{n\to\infty} \eta_n^2(hl) = \eta_{lh}^*$$

where

$$\eta_{lh}^* = \frac{2\alpha h - 2c}{2\alpha h - \alpha(2-\alpha)(h+l)} \in (0,1).$$

Allowing for bundling improves efficiency in two dimensions. First, in some cases, public goods that are not feasible to provide separately can be provided with probability one when bundling is allowed. This is shown in Table 1, which also illustrates the need for studying excludable public goods when considering large economies.[22]

Secondly, even for public goods that can be provided without bundling, the optimal bundling mechanism still improves efficiency by increasing the probability of inclusion for low-valuation agents. Specifically, suppose that $\alpha h > c$ so that both public goods will be asymptotically provided with probability one with or without bundling. From Proposition 4, we know that under the best separate provision mechanism, the ex ante probability for access is

$$\alpha + (1-\alpha)\frac{\alpha h - c}{\alpha h - l}, \tag{34}$$

[22]Mailath and Postlewaite [11] considers a single-dimensional problem. 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.

where α is the probability of a high valuation, in which case the consumer gets access for sure, $1-\alpha$ is the probability of a low valuation, in which case the inclusion probability is $(\alpha h - c)/(\alpha h - l)$). In contrast, Proposition 5 implies that the ex ante probability for access in the bundling regime in the case where in the case where $2c > \alpha(2-\alpha)(h+l)$ is

$$\alpha + (1-\alpha)\,\alpha \frac{2\alpha h - 2c}{2\alpha h - \alpha(2-\alpha)(h+l)}. \tag{35}$$

Again, consumers with high valuations get access for sure, while low valuation consumers get access with probability $(2\alpha h - 2c)/[2\alpha h - \alpha(2-\alpha)(h+l)]$. Simple algebra shows that (35) is larger than (34).[23] Fewer consumers are thus excluded in the optimal bundling mechanism. A similar calculation applies to the case where $2c < \alpha(2-\alpha)(h+l)$.

6 Conclusion

This paper studies the role of bundling in the optimal provision of multiple excludable public goods in large economies. We show that bundling in the provision of unrelated public goods can enhance social welfare. For a parametric class of examples with binary valuations, we characterize the optimal mechanism and show that allowing for bundling alleviates the well-known free riding problem in large economies and increases the probability of public good provision. The basic intuition, which we formalized in the case with many goods, is that bundling reduces the variance of valuations, and that, due to the non-rivalness, there are no direct efficiency losses from providing goods to agents with a low willingness to pay. Bundling therefore leads to fewer exclusions, which in turn implies an increase in revenue. In our simplistic model this will in some cases change the probability of provision from zero to one when allowing the designer to bundle. In a richer environment with a quantity dimension, this should translate to increasing the provision levels.

We believe that there are two interesting directions in which the model of this paper could be extended. First, can we characterize the optimal mechanism for the provision of multiple public goods when the valuation distributions are more general? We believe that this is possible. In particular, as long as all goods are binary and the problem is symmetric, a natural ordering of the types exist no matter how many goods there are. While still very simplistic, this extension would allow an analysis of how many different bundles would be offered, and one could also address to what extent the mixing in the current paper is an artefact of the minimal type space. Secondly, it

[23] The difference between the two access probabilities is $(1-\alpha)$ multiplied by

$$\frac{2\alpha h - 2c}{2h - (2-\alpha)(h+l)} - \frac{\alpha h - c}{\alpha h - l} = \frac{\alpha(\alpha h - c)(h-l)}{[2h - (2-\alpha)(h+l)][\alpha h - l]} > 0.$$

does not seem crucial to have non-rivalness in all goods. One could therefore use our setup to ask to what extent public provision of a private good could be rationalized as a way 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] Aumann, Robert. J. "Mixed and Behavioral Strategies in Infinite Extensive Games," *Annals of Mathematical Studies,* 52, 1964, 627-650.

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

[6] Cremer, Jacques and Richard P. McLean. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions", *Econometrica,* 56(6), November 1988, pages 1247-57.

[7] Dana, James D. "The Organization and Scope of Agents: Regulating Multiproduct Industries," *Journal of Economic Theory,* **59**(2)**,** 1993, 288-309.

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

[9] Hellwig, Martin, F. "Public Good Provision with Many Participants," *Review of Economic Studies,* **70**(3)**,** 2003, 589-614.

[10] Jackson, Matthew O. and Hugo F. Sonnenschein. "The Linking of Collective Decisions and Efficiency," *manuscript*, California Institute of Technology, March 2003.

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

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

[13] 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.

[14] Milgrom, Paul and Robert Weber. "Distributional Strategies fro Games with Incomplete Information," *Mathematics of Operations Research,* 10, 1985,

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

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

[17] Norman, Peter. "Efficient Mechanisms for Public Goods with Use Exclusion." Forthcoming, *Review of Economic Studies*, 2003.

A Appendix: Proofs

Proof of Proposition 1.

Claim A1 *For any incentive feasible mechanism $\mathcal{G}$ of the form (3), there exist an incentive feasible mechanism*

$$G = \left(\left(\rho^j, \eta_1^j, ..., \eta_n^j \right)_{j \in \mathcal{J}}, (t_i)_{i \in \mathcal{I}} \right), \tag{A1}$$

that generates the same social surplus, where $\rho^j : \mathbf{\Theta} \to [0,1]$ 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,

$$\rho^j(\theta) = \mathrm{E}_{\Xi}\zeta^j(\theta,\vartheta) = \int_0^1 \zeta^j(\theta,\vartheta)\, d\vartheta \tag{A2}$$

$$\eta_i^j(\theta_i) = \begin{cases} \frac{\mathrm{E}_{-i}\zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)}{\mathrm{E}_{-i}\zeta^j(\theta,\vartheta)} = \frac{\int_{\mathbf{\Theta}_{-i}}\int_0^1 \zeta^j(\theta,\vartheta)\omega_i^j(\theta,\vartheta)d\vartheta d\mathbf{F}(\theta_{-i})}{\int_{\mathbf{\Theta}_{-i}}\int_0^1 \zeta^j(\theta,\vartheta)d\vartheta d\mathbf{F}(\theta_{-i})} & \text{if } \int_{\mathbf{\Theta}_{-i}}\int_0^1 \zeta^j(\theta,\vartheta)\, d\vartheta d\mathbf{F}(\theta_{-i}) > 0 \\ k & \text{if } \int_{\mathbf{\Theta}_{-i}}\int_0^1 \zeta^j(\theta,\vartheta)\, d\vartheta d\mathbf{F}(\theta_{-i}) = 0 \end{cases}$$

$$t_i(\theta_i) = \mathrm{E}_{-i}\tau(\theta) = \int_{\mathbf{\Theta}_{-i}} \tau(\theta)\, d\mathbf{F}(\theta_{-i}),$$

for each $\theta \in \mathbf{\Theta}, j \in \mathcal{J}$ and $i \in \mathcal{I}$. This is a mechanism of the form in (A1), and we will call it G. Use of the law of iterated expectations on $\rho^j(\theta)$ and $t_i(\theta_i)$ shows that the feasibility constraint (5) is unaffected when switching from $\mathcal{G}$ to G. It remains to show that the surplus is unchanged, and that (4) and (6) continue to hold under G. The utility of agent i of type $\theta_i \in \Theta$ who announces $\hat{\theta}_i \in \Theta$ is

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\theta_i - \tau\left(\hat{\theta}_i, \theta_{-i}\right)\right] \text{ in mechanism } \mathcal{G} \tag{A3}$$

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \rho^j\left(\hat{\theta}_i, \theta_{-i}\right)\eta_i^j\left(\hat{\theta}_i\right)\theta_i - t_i\left(\hat{\theta}_i\right)\right] \text{ in mechanism } G. \tag{A4}$$

If $\int_{\mathbf{\Theta}_{-i}}\int_0^1 \zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right) d\vartheta d\mathbf{F}(\theta_{-i}) = 0$, we trivially have that

$$\mathrm{E}_{-i}\rho^j\left(\hat{\theta}_i, \theta_{-i}\right)\eta_i^j\left(\hat{\theta}_i\right)\theta_i = 0 = \mathrm{E}_{-i}\left[\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\theta_i\right]; \tag{A5}$$

whereas if $\int_{\mathbf{\Theta}_{-i}}\int_0^1 \zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right) d\vartheta d\mathbf{F}(\theta_{-i}) > 0$, we have that

$$\begin{aligned} \mathrm{E}_{-i}\rho^j\left(\hat{\theta}_i, \theta_{-i}\right)\eta_i^j\left(\hat{\theta}_i\right)\theta_i &= \mathrm{E}_{-i}\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\frac{\mathrm{E}_{-i}\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)}{\mathrm{E}_{-i}\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)} \\ &= \mathrm{E}_{-i}\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\zeta^j\left(\hat{\theta}_i, \theta_{-i}, \vartheta\right)\theta_i. \end{aligned} \tag{A6}$$

Trivially, $\mathrm{E}_{-i} t_i(\theta_i) = t_i(\theta_i) = \mathrm{E}_{-i}\tau(\theta)$, which combined with (A5) and (A6) implies that the payoffs in (A3) and (A4) are identical. Since the equality between (A3) and (A4) were established for any i, θ_i and $\hat{\theta}_i$, it follows that all incentive and participation constraints (4) and (6) hold for mechanism G given that they are satisfied in mechanism $\mathcal{G}$. Moreover, [again by (A5) and (A6)]

$$\mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i\right] = \mathrm{E}_{-i}\left[\sum_{j\in\mathcal{J}} \omega_i^j(\theta,\vartheta)\,\zeta^j(\theta,\vartheta)\,\theta_i\right], \tag{A7}$$

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

$$\mathrm{E}\left[\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i\right] = \mathrm{E}\left[\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \zeta^j(\theta,\vartheta)\,\omega_i^j(\theta,\vartheta)\,\theta_i\right], \tag{A8}$$

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

$$\sum_{j\in\mathcal{J}} \mathrm{E}\rho^j(\theta)\left[\sum_{i\in\mathcal{I}} \eta_i^j(\theta_i)\,\theta_i - C^j(n)\right] = \sum_{j\in\mathcal{J}} \mathrm{E}\zeta^j(\theta,\vartheta)\left[\sum_{i\in\mathcal{I}} \omega_i^j(\theta,\vartheta)\,\theta_i - C^j(n)\right]. \tag{A9}$$

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

Claim A2 *For every incentive feasible mechanism of the form (A1), there exists an anonymous simple incentive feasible mechanism g of the form (8) that generates the same surplus.*

Proof. Consider an incentive feasible simple mechanism G on form (A1). 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 $\theta \in \mathbf{\Theta}$, let $\theta^{P_K} = \left(\theta_{P_k^{-1}(1)}, ..., \theta_{P_k^{-1}(n)}\right) \in \mathbf{\Theta}$ denote the corresponding k-th permutation of θ.[24] 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

$$\begin{aligned} \rho_k^j(\theta) &= \rho^j\left(\theta^{P_k}\right) \quad \forall\, \theta \in \mathbf{\Theta},\ j \in \mathcal{J}, \\ \eta_{ki}^j(\theta_i) &= \eta_{P_k^{-1}(i)}^j(\theta_i) \quad \forall\, \theta_i \in \Theta,\ j \in \mathcal{J}, i \in \mathcal{I}, \\ t_{ki}(\theta_i) &= t_{P_k^{-1}(i)}(\theta_i) \quad \forall\, \theta_i \in \Theta, i \in \mathcal{I}, \end{aligned} \tag{A10}$$

[24] To illustrate, suppose $n = 3, m = 2, \theta = (\theta_1, \theta_2, \theta_3) = ((1,2), (3,2), (2,1))$. Consider, for example, permutation 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 $\theta^{P_k} = \left(\theta_{P_k^{-1}(1)}, \theta_{P_k^{-1}(2)}, \theta_{P_k^{-1}(3)}\right) = (\theta_2, \theta_1, \theta_3) = ((3,2), (1,2), (2,1))$.

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

$$\begin{aligned}
\widetilde{\rho}^j(\theta) &= \frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta) \quad \forall\, \theta \in \mathbf{\Theta},\ j \in \mathcal{J} \qquad \text{(A11)}\\
\widetilde{\eta}_i^j(\theta_i) &= \frac{\sum_{k=1}^{n!} \mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]} \quad \forall\, \theta_i \in \Theta, i \in \mathcal{I},\ j \in \mathcal{J}\\
\widetilde{t}_i(\theta_i) &= \frac{1}{n!}\sum_{k=1}^{n!} t_{ki}(\theta_i) \quad \forall\, \theta_i \in \Theta, i \in \mathcal{I}.
\end{aligned}$$

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 $\mathrm{E}_{-i}\left[\widetilde{\rho}^j(\theta)\right] = \mathrm{E}_{-i'}\left[\widetilde{\rho}^j(\theta)\right]$; 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 $\{t_{ki}(\theta_i')\}_{k=1}^{n!}$ 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 $k = 1, ..., n!$,

$$\sum_{j\in\mathcal{J}} \mathrm{E}\left\{\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_{ki}^j(\theta_i)\theta_i^j - C^j(n)\right]\right\} = \sum_{j\in\mathcal{J}}\mathrm{E}\left\{\rho^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_i^j(\theta_i)\theta_i^j - C^j(n)\right]\right\}. \qquad \text{(A12)}$$

Hence,

$$\begin{aligned}
\sum_{j\in\mathcal{J}}\mathrm{E}\left\{\widetilde{\rho}^j(\theta)\left[\sum_{i\in\mathcal{I}}\widetilde{\eta}^j(\theta_i)\theta_i^j - C^j(n)\right]\right\} &= \sum_{j\in\mathcal{J}}\mathrm{E}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}}\frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\rho_k^j(\theta)}\theta_i^j - C^j(n)\right]\right\}\\
= \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}}\mathrm{E}_{\theta_i}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\mathrm{E}_{-i}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\theta_i^j\right\} &- \mathrm{E}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]C^j(n)\\
= \frac{1}{n!}\sum_{k=1}^{n!}\sum_{j\in\mathcal{J}}\mathrm{E}\left\{\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_{ki}^j(\theta_i)\theta_i^j - C^j(n)\right]\right\} &= \sum_{j\in\mathcal{J}}\mathrm{E}\left\{\rho^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_i^j(\theta_i)\theta_i^j - C^j(n)\right]\right\}, \qquad \text{(A13)}
\end{aligned}$$

where the last equality follows from (A12). 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

$$\begin{aligned} \mathrm{E}\sum_{i\in\mathcal{I}}\widetilde{t}_i(\theta_i)-\sum_{j\in\mathcal{J}}\mathrm{E}\widetilde{\rho}^j(\theta)C^j(n) &= \mathrm{E}\sum_{i\in\mathcal{I}}\frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\theta_i)-\sum_{j\in\mathcal{J}}\mathrm{E}\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)C^j(n) \\ &= \mathrm{E}\sum_{i\in\mathcal{I}}t_i(\theta_i)-\sum_{j\in\mathcal{J}}\mathrm{E}\rho^j(\theta)C^j(n), \end{aligned} \tag{A14}$$

so g is feasible if the original mechanism G is. Second, we note that incentive compatibility holds for any permuted mechanism, i.e.,

$$\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\theta_i^j-t_{ki}(\theta_i)\geq\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j-t_{ki}(\widehat{\theta}_i,\theta_{-i}) \tag{A15}$$

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

$$\begin{aligned} &\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\widetilde{\rho}^j(\theta)\widetilde{\eta}^j(\theta_i)\theta_i^j-\widetilde{t}(\theta_i)=\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]\frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]}\theta_i^j-\frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\theta_i) \\ =\;& \frac{1}{n!}\sum_{k=1}^{n!}\left[\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\theta_i^j-t_{ki}(\theta_i)\right]\geq\frac{1}{n!}\sum_{k=1}^{n!}\left[\mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j-t_{ki}(\widehat{\theta}_i,\theta_{-i})\right] \\ =\;& \mathrm{E}_{-i}\sum_{j\in\mathcal{J}}\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j-\frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\widehat{\theta}_i,\theta_{-i})=\sum_{j\in\mathcal{J}}\mathrm{E}_{-i}\widetilde{\rho}^j(\widehat{\theta}_i,\theta_{-i})\widetilde{\eta}_i^j\left(\widehat{\theta}_i\right)\theta_i^j-\widetilde{t}(\widehat{\theta}_i), \end{aligned} \tag{A16}$$

where the inequality follows from (A15). Hence g is incentive compatible. Finally, g also satisfies the participation constraints because (see the second line in (A16)) all the permuted mechanisms satisfy participation constraints. Proposition 1 follows by combining Claims A1 and A2. ■

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 $\int_{\theta\in\mathbf{\Theta}}h(\theta)\,d\mathbf{F}(\theta)$ 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 $k\in\{1,...,m!\}$, write with some abuse of notation $P_k:\mathcal{J}\to\mathcal{J}$ as the k-th permutation of $\mathcal{J}$, and let $\theta_i^{P_k}=\left(\theta_i^{P_k^{-1}(1)},...,\theta_i^{P_k^{-1}(m)}\right)\in$

Θ denote the permutation of θ_i when the goods are permuted according to P_k. Write $\theta^{P_k} = \left(\theta_1^{P_k}, ..., \theta_n^{P_k}\right) \in \mathbf{\Theta}$ denote the corresponding permutation of θ.[25] 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 $\theta \in \mathbf{\Theta}$;

1. $\rho_k^j(\theta) = \rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)$ for every $j \in \mathcal{J}$;[26]
2. $\eta_k^j(\theta_i) = \eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)$ for every $j \in \mathcal{J}$;[27]
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:[28]

$$\begin{aligned}
\mathrm{E}\rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j &= \int_\theta \rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j d\mathbf{F}(\theta)\ /\text{def of } g_k/ = \int_{\theta\in\mathbf{\Theta}} \rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\theta_i^j d\mathbf{F}(\theta) \\
&= \int_{\theta^j}\left[\int_{\theta^{-j}} \rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\theta_i^j d\mathbf{F}^{-j}\left(\theta^{-j}\middle|\,\theta^j\right)\right] d\mathbf{F}^j\left(\theta^j\right) \qquad \text{(A17)} \\
/\text{relabel}/ &= \int_{\theta^{P_k^{-1}(j)}}\left[\int_{(\theta^{-j})^{P_k}} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}^{-j}\left(\left(\theta^{-j}\right)^{P_k}\middle|\,\underbrace{\theta^{P_k^{-1}(j)}}_{j\text{-th argument}}\right)\right] d\mathbf{F}^j\left(\theta^{P_k^{-1}(j)}\right)
\end{aligned}$$

where we recall,

$$\left(\theta^{-j}\right)^{P_k} \equiv \left(\theta^{P_k^{-1}(1)}, ..., \theta^{P_k^{-1}(j-1)}, \theta^{P_k^{-1}(j+1)}, ..., \theta^{P_k^{-1}(n)}\right). \qquad \text{(A18)}$$

By exchangeability, we have

$$\begin{aligned}
& d\mathbf{F}^{-j}\left(\left(\theta^{-j}\right)^{P_k}\middle|\,\underbrace{\theta^{P_k^{-1}(j)}}_{j\text{-th (vector) argument}}\right) \qquad \text{(A19)} \\
= \;& d\mathbf{F}^{-j}\left(\theta^{P_k^{-1}(1)}, ..., \theta^{P_k^{-1}(j-1)}, \theta^{P_k^{-1}(j+1)}, ..., \theta^{P_k^{-1}(n)}|j\text{-th (vector) argument} = \theta^{P_k^{-1}(j)}\right) \\
= \;& d\mathbf{F}^{-j}\left(\theta^{-j}|j\text{-th (vector) argument} = \theta^{P_k^{-1}(j)}\right) \\
= \;& d\mathbf{F}^{-P_k^{-1}(j)}\left(\theta^{-P_k^{-1}(j)}|P_k^{-1}(j)\text{-th (vector) argument } = \theta^{P_k^{-1}(j)}\right);
\end{aligned}$$

and

$$d\mathbf{F}^j\left(\theta^{P_k^{-1}(j)}\right) = d\mathbf{F}^{\theta^{P_k^{-1}(j)}}\left(\theta^{P_k^{-1}(j)}\right). \qquad \text{(A20)}$$

[25]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 $\theta_1^{P_k} = \left(\theta_1^{P_k^{-1}(1)}, \theta_1^{P_k^{-1}(2)}, \theta_1^{P_k^{-1}(3)}\right) = (2,1,0), \theta_2^{P_k} = \left(\theta_2^{P_k^{-1}(1)}, \theta_2^{P_k^{-1}(2)}, \theta_2^{P_k^{-1}(3)}\right) = (2,3,1), \theta^{P_k} = \left(\theta_1^{P_k}, \theta_2^{P_k}\right) = ((2,1,0), (2,3,1))$.

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

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

[28]It is important to point out that, in reaching the fourth equality in (A17), we can relabel the integrating varibles (since they are dummies) but not the integrating functions.

Using (A17), (A19) and (A20), we have that

$$\begin{aligned}
& \mathrm{E}\rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j \qquad\qquad (A21)\\
= & \int_{\theta^{P_k^{-1}(j)}} \left[\int_{(\theta^{-j})^{P_k}} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}^{-j}\left(\left(\theta^{-j}\right)^{P_k} \Big| \underbrace{\theta^{P_k^{-1}(j)}}_{j\text{-th argument}} \right) \right] d\mathbf{F}^j\left(\theta^{P_k^{-1}(j)}\right) \\
= & \int_{\theta^{P_k^{-1}(j)}} \left[\int_{(\theta^{-j})^{P_k}} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}^{-P_k^{-1}(j)}\left(\theta^{-P_k^{-1}(j)} | \underbrace{\theta^{P_k^{-1}(j)}}_{P_k^{-1}(j)\text{-th argument}} \right) \right] d\mathbf{F}^{P_k^{-1}(j)}\left(\theta^{P_k^{-1}(j)}\right) \\
= & \int_{\theta} \rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\,\theta_i^{P_k^{-1}(j)} d\mathbf{F}(\theta) = \mathrm{E}\rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\theta_i^{P_k^{-1}(j)}.
\end{aligned}$$

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,

$$\begin{aligned}
\mathrm{E}\left[\sum_{j=1}^m \rho_k^j(\theta)\,\eta_k^j(\theta_i)\theta_i^j - t_k(\theta_i)\right] &= \left[\sum_{j=1}^m \mathrm{E}\rho^{P_k^{-1}(j)}(\theta)\,\eta^{P_k^{-1}(j)}(\theta_i)\theta_i^{P_k^{-1}(j)}\right] - \mathrm{E}t(\theta_i) \qquad (A22)\\
\Big/ \begin{array}{c}\text{Same elements in } \mathcal{J} \text{ and} \\ \left\{P_k^{-1}(1),...,P_k^{-1}(m)\right\}\end{array} \Big/ &= \left[\sum_{j=1}^m \mathrm{E}\rho^j(\theta)\,\eta^j(\theta_i)\theta_i^j\right] - \mathrm{E}t(\theta_i),
\end{aligned}$$

is thus unchanged when changing from g to g_k. The same steps as in (A17) through (A21) (only somewhat simpler) establishes that $\mathrm{E}\rho_k^j(\theta) = \mathrm{E}\rho^{P_k^{-1}(j)}$ for every j, implying that

$$\begin{aligned}
\mathrm{E}\left[\sum_{j=1}^m \rho_k^j(\theta)\,C^j(n) - \sum_i t_k(\theta_i)\right] &= \left[C(n)\,\mathrm{E}\sum_{j=1}^m \rho_k^j(\theta) - \sum_i \mathrm{E}t_k(\theta_i)\right] \qquad (A23)\\
= \left[C(n)\,\mathrm{E}\sum_{j=1}^m \rho^j(\theta) - \sum_i \mathrm{E}t(\theta_i)\right] &= \mathrm{E}\left[\sum_{j=1}^m \rho^j(\theta)\,C(n) - \sum_i t(\theta_i)\right],
\end{aligned}$$

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 (A17) through (A21):

$$\begin{aligned}
\mathrm{E}_{-i}\rho_k^j(\theta_{-i},\theta_i') &= \int_{\theta_{-i}} \rho_k^j(\theta_{-i},\theta_i')\,d\mathbf{F}_{-i}(\theta_{-i}) \,/\text{def of } g_k/ = \int_{\theta_{-i}} \rho^{P_k^{-1}(j)}\left((\theta_{-i},\theta_i')^{P_k}\right) d\mathbf{F}_{-i}(\theta_{-i}) \\
&= \int_{\theta_{-i}^j} \left[\int_{\theta_{-i}^{-j}} \rho^{P_k^{-1}(j)}\left((\theta_{-i},\theta_i')^{P_k}\right) d\mathbf{F}_{-i}^{-j}\left(\theta_{-i}^{-j}\Big|\,\theta_{-i}^j\right)\right] d\mathbf{F}_{-i}^j\left(\theta_{-i}^j\right) \qquad (A24)\\
/\text{relabel}/ \quad &= \int_{\theta_{-i}^{P_k^{-1}(j)}} \left[\int_{\theta_{-i}^{-P_k^{-1}(j)}} \rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) d\mathbf{F}_{-i}^{-j}\left(\left(\theta_{-i}^{-j}\right)^{P_k}\Big|\,\theta_{-i}^{P_k^{-1}(j)}\right)\right] d\mathbf{F}_{-i}^j\left(\theta_{-i}^{P_k^{-1}(j)}\right) \\
/\text{exchangeability}/ \quad &= \int_{\theta_{-i}^{P_k^{-1}(j)}} \left[\int_{\theta_{-i}^{-P_k^{-1}(j)}} \rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) d\mathbf{F}_{-i}^{-P_k^{-1}(j)}\left(\theta_{-i}^{-P_k^{-1}(j)}\Big|\,\theta_{-i}^{P_k^{-1}(j)}\right)\right] d\mathbf{F}_{-i}^j\left(\theta_{-i}^{P_k^{-1}(j)}\right) \\
&= \int_{\theta_{-i}} \rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) d\mathbf{F}_{-i}(\theta_{-i}) = \mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right)
\end{aligned}$$

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 $(\theta_i')^{P_k}$, so that

$$U(\theta_i,\theta_i';g_k) = \mathrm{E}_{-i}\sum_{j=1}^{m}\rho_k^j\left(\theta_{-i},\theta_i'\right)\eta_k^j(\theta_i')\theta_i^j - t_k\left(\theta_i'\right) \tag{A25}$$
$$= \sum_{j=1}^{m}\eta_k^{P_k^{-1}(j)}((\theta_i')^{P_k})\theta_i^j\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) - t\left((\theta_i')^{P_k}\right),$$

whereas

$$U(\theta_i,\theta_i';g) = \sum_{j=1}^{m}\eta_k^j(\theta_i')\theta_i^j\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'\right) - t\left(\theta_i'\right) \Rightarrow \tag{A26}$$
$$U(\theta_i,\theta_i';g)\big|_{\substack{\theta_i=\theta_i^{P_k}\\ \theta_i'=\theta_i'^{P_k}}} = \sum_{j=1}^{m}\eta_k^j((\theta_i')^{P_k})\theta_i^{P_k^{-1}(j)}\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'^{P_k}\right) - t\left((\theta_i')^{P_k}\right)$$
$$= \sum_{j=1}^{m}\eta_k^{P_k^{-1}(j)}((\theta_i')^{P_k})\theta_i^j\mathrm{E}_{-i}\rho_k^{P_k^{-1}(j)}\left(\theta_{-i},\theta_i'^{P_k}\right) - t\left((\theta_i')^{P_k}\right) = U(\theta_i,\theta_i';g_k),$$

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 $(\theta_i')^{P_k}$ 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

$$\widetilde{\rho}^j(\theta) = \frac{1}{m!}\sum_{k=1}^{m!}\rho_k^j(\theta) = \frac{1}{m!}\sum_{k=1}^{m!}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right) \tag{A27}$$
$$\widetilde{\eta}^j(\theta_i) = \frac{\sum_{k=1}^{m!}\eta_k^j(\theta_i)\,\mathrm{E}_{-i}\rho_k^j(\theta)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j(\theta)} = \frac{\sum_{k=1}^{m!}\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)}$$
$$\widetilde{t}(\theta_i) = \frac{1}{m!}t_k(\theta_i) = \frac{1}{m!}t\left(\theta_i^{P_k}\right)$$

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

$$\widetilde{\rho}^{P^{-1}(j)}\left(\theta^P\right) = \frac{1}{m!}\sum_{k=1}^{m!}\rho^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta^P\right)^{P_k}\right) = \frac{1}{m!}\sum_{k=1}^{m!}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right) = \widetilde{\rho}^j(\theta), \tag{A28}$$

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

$$\widetilde{\eta}^{P^{-1}(j)}\left(\theta_i^P\right) = \frac{\sum_{k=1}^{m!}\eta_k^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta_i^P\right)^{P_k}\right)\mathrm{E}_{-i}\rho_k^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta^P\right)^{P_k}\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^{P_k^{-1}\left(P^{-1}(j)\right)}\left(\left(\theta^P\right)^{P_k}\right)} \tag{A29}$$
$$= \frac{\sum_{k=1}^{m!}\eta^{P_k^{-1}(j)}\left(\theta_i^{P_k}\right)\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho^{P_k^{-1}(j)}\left(\theta^{P_k}\right)} = \widetilde{\eta}^j(\theta_i)$$

for the same reason. It is obvious that $\widetilde{t}\left(\theta_i^P\right) = \widetilde{t}(\theta_i)$, which together with (A28) and (A29) 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

$$\mathrm{E}\widetilde{\rho}^j(\theta)\widetilde{\eta}^j(\theta_i)\theta_i^j = \frac{1}{m!}\sum_{k=1}^{m!}\mathrm{E}\rho_k^j(\theta)\frac{\sum_{k=1}^{m!}\eta_k^j(\theta_i)\mathrm{E}_{-i}\rho_k^j(\theta)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j(\theta)}\theta_i^j \tag{A30}$$

$$= \frac{1}{m!}\mathrm{E}_{\theta_i}\sum_{k=1}^{m!}\left[\mathrm{E}_{-i}\rho_k^j(\theta)\frac{\sum_{k=1}^{m!}\eta_k^j(\theta_i)\mathrm{E}_{-i}\rho_k^j(\theta)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j(\theta)}\theta_i^j\right] = \frac{1}{m!}\mathrm{E}\left[\sum_{k=1}^{m!}\eta_k^j(\theta_i)\rho_k^j(\theta)\theta_i^j\right]$$

$$\Rightarrow \mathrm{E}\sum_{j=1}^{n}\left[\widetilde{\rho}^j(\theta)\widetilde{\eta}^j(\theta_i)\theta_i^j - \widetilde{t}(\theta_i)\right] = \frac{1}{m!}\sum_{k=1}^{m!}\mathrm{E}\left[\sum_{j=1}^{m}\eta_k^j(\theta_i)\rho_k^j(\theta)\theta_i^j - t_k(\theta_i)\right]$$

$$/(\text{A21})\ \&\ (\text{A22})/ \quad = \mathrm{E}\left[\sum_{j=1}^{m}\eta^j(\theta_i)\rho^j(\theta)\theta_i^j - t(\theta_i)\right],$$

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

$$\mathrm{E}\left[\sum_{j=1}^{m}\widetilde{\rho}^j(\theta)C^j(n) - \sum_{i=1}^{n}\widetilde{t}(\theta_i)\right] = \mathrm{E}\left[C(n)\sum_{j=1}^{m}\frac{1}{m!}\sum_{k=1}^{m!}\rho_k^j(\theta) - \sum_{i=1}^{n}\sum_{k=1}^{m!}\frac{1}{m!}t_k(\theta_i)\right] \tag{A31}$$

$$= \frac{1}{m!}\sum_{k=1}^{m!}\mathrm{E}\left[C(n)\sum_{j=1}^{m}\rho_k^j(\theta) - \sum_{i=1}^{n}t_k(\theta_i)\right] /\ (\text{A23})/ = \frac{1}{m!}\sum_{k=1}^{m!}\mathrm{E}\left[\sum_{j=1}^{m}\rho^j(\theta)C(n) - \sum_{i=1}^{n}t(\theta_i)\right]$$

$$= \mathrm{E}\left[\sum_{j=1}^{m}\rho^j(\theta)C^j(n) - \sum_{i=1}^{n}t(\theta_i)\right],$$

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

$$U(\theta_i,\theta_i';\widetilde{g}) = \sum_{j=1}^{m}\widetilde{\eta}^j(\theta_i')\theta_i^j\mathrm{E}_{-i}\widetilde{\rho}^j\left(\theta_{-i},\theta_i'\right) - \widetilde{t}\left(\theta_i'\right) \tag{A32}$$

$$= \frac{\sum_{k=1}^{m!}\eta_k^j\left(\theta_i'\right)\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'\right)}{\sum_{k=1}^{m!}\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'\right)}\mathrm{E}_{-i}\left[\frac{1}{m!}\sum_{k=1}^{m!}\rho_k^j\left(\theta_{-i},\theta_i'\right)\right] - \frac{1}{m!}\sum_{k=1}^{m!}t_k\left(\theta_i'\right)$$

$$= \frac{1}{m!}\sum_{k=1}^{m!}\left[\eta_k^j\left(\theta_i'\right)\mathrm{E}_{-i}\rho_k^j\left(\theta_{-i},\theta_i'\right) - t_k\left(\theta_i'\right)\right]$$

$$/\ (\text{A25})/ \quad = \frac{1}{m!}\sum_{k=1}^{m!}U(\theta_i,\theta_i';g_k) \leq /\ \text{IC for each } k/\ \frac{1}{m!}\sum_{k=1}^{m!}U(\theta;g_k) = U(\theta;\widetilde{g}).$$

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

Proof of Lemma 1.

Proof. **[Part 1]** Pick an arbitrary $\varepsilon > 0$ and assume that there exists $\delta > 0$ and $N < \infty$ such that $\mathrm{E}\theta_i^j - C^j(n)/n \geq \delta$ for every $n \geq N$. Applying Chebyshev's inequality, we have

$$\begin{aligned}\Pr\left[\sum_{i=1}^{n}\theta_i^j \leq C^j(n)\right] &\leq \Pr\left[\sum_{i=1}^{n}\theta_i^j \leq n(\mathrm{E}\theta_i^j - \delta)\right] = \Pr\left[\sum_{i=1}^{n}\theta_i^j - n\mathrm{E}\theta_i^j \leq -n\delta\right] \\ &\leq \Pr\left[\left|\sum_{i=1}^{n}\theta_i^j - n\mathrm{E}\theta_i^j\right| \geq n\delta\right] = \frac{\mathrm{Var}\left(\sum_i \theta_i^j\right)}{n^2\delta^2} \leq \frac{\sigma^2}{n\delta^2}.\end{aligned} \tag{A33}$$

Hence, for every $\varepsilon > 0$ we can find some N' such that the probability that the ex post efficient rule provides good j is at least $1-\varepsilon$, which establishes part 1 of the claim. **[Part 2]** The argument is symmetric and omitted. ■

<u>*Proof of Proposition 3.*</u>

Proof. **[Part 1]** Let $\varepsilon \in (0,\delta)$. Since we assume that there exists N such that $\mathrm{E}\theta_i^j - C^j(n)/n \geq \delta$ for each j and $n \geq N$ we have that

$$\begin{aligned}1 - \mathrm{E}\widehat{\eta}_i^j(\theta_i) &= \Pr\left[\sum_j \theta_i^j < \sum_{j=1}^{m}\frac{C^j(n)}{n} + \varepsilon m\right] \leq \Pr\left[\sum_j \theta_i^j < \sum_{j=1}^{m}\mathrm{E}\theta_i^j - \delta m + \varepsilon m\right] \\ &\leq \Pr\left[\left|\sum_j \theta_i^j - \sum_{j=1}^{m}\mathrm{E}\theta_i^j\right| < m(\delta-\varepsilon)\right],\end{aligned} \tag{A34}$$

for every $n \geq N$. By the assumption that $\mathrm{Var}\theta_i^j \leq \sigma^2$ for every j we can apply Chebyshev's inequality to conclude that

$$1 - \mathrm{E}\widehat{\eta}_i^j(\theta_i) \leq \frac{\mathrm{Var}\left(\sum_j \theta_i^j\right)}{(\delta-\varepsilon)^2 m^2} \leq \frac{\sigma^2}{(\delta-\varepsilon)^2 m} \to 0 \text{ as } m \to \infty \tag{A35}$$

Hence, there exists M such that $\mathrm{E}\widehat{\eta}_i^j(\theta_i) \geq 1 - \frac{\varepsilon}{2(\mu+\varepsilon)}$ for every $m \geq M$. It follows that

$$\begin{aligned}\mathrm{E}\sum_{i=1}^{n}\widehat{t}(\theta_i) - \sum_{j=1}^{m}C^j(n) &\geq n\left(1 - \frac{\varepsilon}{2(\mu+\varepsilon)}\right)\left(\sum_{j=1}^{m}\frac{C^j(n)}{n} + \varepsilon m\right) - \sum_{j=1}^{m}C^j(n) \\ &= -\frac{n\varepsilon}{2(\mu+\varepsilon)}\sum_{j=1}^{m}\frac{C^j(n)}{n} + n\left(1 - \frac{\varepsilon}{2(\mu+\varepsilon)}\right)\varepsilon m \\ &\geq -\frac{n\varepsilon}{2(\mu+\varepsilon)}m\mu + n\left(\frac{2\mu+\varepsilon}{2(\mu+\varepsilon)}\right)\varepsilon m = \frac{nm\varepsilon}{2(\mu+\varepsilon)}[\mu+\varepsilon] > 0,\end{aligned} \tag{A36}$$

for $m \leq M$, so (5) is satisfied. The remaining constraints hold trivially, so $\widehat{g}$ is incentive feasible.

[Part 2] Let $\rho^* = \{\rho^{*j}\}_{j=1}^{m}$ denote the ex post efficient provision rules (transfers are irrelevant for efficiency and we no consumer is excluded from usage). Let $A^j = \left\{\theta \in \mathbf{\Theta} \middle| \sum_{i\in\mathcal{I}}\theta_i^j - C^j(n)\right\}$

denote the set of type profiles for which the ex post efficient rule provides good j. The per capita surplus generated by ρ^* may this be written as

$$s^* \equiv \sum_{j\in\mathcal{J}} \int_{\theta\in\mathbf{\Theta}} \rho^{*j}(\theta) \left[\sum_{i\in\mathcal{I}} \frac{\theta_i^j}{n} - \frac{C^j(n)}{n}\right] d\mathbf{F}(\theta) = \sum_{j\in\mathcal{J}} \int_{\theta\in A^j} \left[\sum_{i\in\mathcal{I}} \frac{\theta_i^j}{n} - \frac{C^j(n)}{n}\right] d\mathbf{F}(\theta). \tag{A37}$$

Define $B_i \equiv \left\{\theta_i \in \Theta | \sum_j \theta_i^j \geq \sum_j \frac{C_j(n)}{n} + \varepsilon m\right\}$.We note that

$$\int_{\Theta} \sum_{j\in\mathcal{J}} \widehat{\eta}^j(\theta_i)\theta_i^j dF(\theta_i) = \int_{B^i} \sum_{j\in\mathcal{J}} \theta_i^j dF(\theta_i) \geq \Pr(B_i) \int_{\theta_i\in\Theta} \sum_{j\in\mathcal{J}} \theta_i^j dF(\theta_i), \tag{A38}$$

Since $\sum_{j\in\mathcal{J}} \theta_i^j > \sum_{j\in\mathcal{J}} \theta_i'^j$ when $\theta_i \in B_i$ and $\theta_i' \notin B_i$. The per capita surplus generated by mechanism $\widehat{g}$ can then be decomposed as

$$\begin{aligned}
\widehat{s} &= \sum_{j\in\mathcal{J}} \int_{\theta\in\mathbf{\Theta}} \left[\sum_{i\in\mathcal{I}} \frac{\widehat{\eta}^j(\theta_i)\theta_i^j}{n} - \frac{C^j(n)}{n}\right] d\mathbf{F}(\theta) \qquad \text{(A39)}\\
&= \sum_{i\in\mathcal{I}} \int_{\theta_{-i}\in\mathbf{\Theta}_{-i}} \int_{\theta_i\in B_i} \sum_{j\in\mathcal{J}} \frac{\theta_i^j}{n} dF(\theta_i)\, d\mathbf{F}_{-i}(\theta_{-i}) - \sum_{j\in\mathcal{J}} \frac{C^j(n)}{n}\\
&\geq \Pr[B_i] \left[\sum_{i\in\mathcal{I}} \int_{\theta\in\mathbf{\Theta}} \sum_{j\in\mathcal{J}} \frac{\theta_i^j}{n} d\mathbf{F}(\theta) - \sum_{j\in\mathcal{J}} \frac{C^j(n)}{n}\right] - (1-\Pr[B_i]) \sum_{j\in\mathcal{J}} \frac{C^j(n)}{n}\\
&= \Pr[B_i] \sum_{j\in\mathcal{J}} \int_{\theta\in A^j} \left[\sum_{i\in\mathcal{I}} \frac{\theta_i^j}{n} - \frac{C^j(n)}{n}\right] d\mathbf{F}(\theta) + \Pr[B_i] \sum_{j\in\mathcal{J}} \int_{\theta\in\Theta\setminus A^j} \left[\sum_{i\in\mathcal{I}} \frac{\theta_i^j}{n} - \frac{C^j(n)}{n}\right] d\mathbf{F}(\theta)\\
&\quad - (1-\Pr[B_i]) \sum_{j\in\mathcal{J}} \frac{C^j(n)}{n} \geq \Pr[B_i]\, s^* - \left[\Pr[B_i]\left[1-\Pr\left(A^j\right)\right] + (1-\Pr[B_i])\right] \sum_{j\in\mathcal{J}} \frac{C^j(n)}{n}.
\end{aligned}$$

By applications of Chebyshev's inequality $\Pr[B_i] \to 1$ and $\Pr\left(A^j\right) \to 1$ as $n \to \infty$, which implies that $\widehat{s} \to s^*$ as $n \to \infty$. ■.

<u>*Proof of Lemma 2.*</u>

Proof. For each $x \in \mathcal{X}_n, j = 1, 2, \theta_i \in \Theta$ we have that $\rho^j(x) \in [0,1], \eta_{\theta_i}^j \in [0,1]$. Next, we note that if $t_{ll} < 0$ and all constraints are satisfied, then a deviation where taxes are changed from t to $t' = (t_{hh}, t_{hl}, t_{lh}, 0)$ and where inclusion and provision rules are unchanged will satisfy all constraints in the relaxed program (21). Similarly, if all constraints hold and $t_{lh} < -l - h$ the deviation

$$t' = (t_{hh}, t_{hl}, -l-h, \max(0, t_{ll})) \tag{A40}$$

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

$$t' = (-3h - l, \max(t_{hl}, -l-h), \max(t_{lh}, -l-h), \max(0, t_{ll})) \tag{A41}$$

Constraint	Multiplier
Type hh IC (18a)	λ_{hh}
Type hl (lh) IC (18b)	λ_{hl}
Type ll IR (19)	λ_{ll}
Feasibility (20)	Λ
$\eta^1_{\theta_i} \geq 0$	γ_{θ_i}
$1-\eta^1_{\theta_i} \geq 0$	ϕ_{θ_i}
$\rho^1(x) \geq 0$	$\gamma(x)$
$1-\rho^1(x) \geq 0$	$\phi(x)$

Table 2: Notation of multipliers.

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 (21) 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. ■

Notation for optimality conditions to program (21).

The proofs that follow make direct use of the Kuhn-Tucker conditions to the optimization problem (21). For easy reference, Table 2 summarizes our notation for the multipliers associated with each constraint.

Proof of Lemma 4.

Proof. **[Step 1]**Consider first the Kuhn-Tucker optimality conditions with respect to η^1_{hh}. They are given by

$$2\sum_{x\in\mathcal{X}_n} \mathbf{a}_n(x)\rho^1(x)\frac{x_{hh}h}{n} + 2\lambda_{hh}\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll})h + \gamma_{hh} - \phi_{hh} = 0$$
$$\gamma_{hh}\eta^1_{hh} = 0, \phi_{hh}(1-\eta^1_{hh}) = 0, \gamma_{hh} \geq 0, \phi_{hh} \geq 0. \tag{A42}$$

All terms except $\gamma_{hh} - \phi_{hh}$ in the first order condition are strictly positive, so $\gamma_{hh} - \phi_{hh} < 0$. The only possibility for this is that $\phi_{hh} > 0$, which requires that $\eta^1_{hh} = 1$ for the complementary slackness constraint to be fulfilled. $\eta^2_{hh} = 1$ follows from proposition 2.

[Step 2] The first order condition with respect to η_{hl}^1 reads

$$2\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,h\frac{x_{hl}}{n}-\lambda_{hh}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,h \tag{A43}$$
$$+\quad\lambda_{hl}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,h+\gamma_{hl}-\phi_{hl}=0.$$

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

$$\begin{aligned}\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,h\frac{x_{hl}}{n} &= \sum_{x\in\mathcal{X}_n:x_{hl}\geq 1}\frac{n}{x_{hl}}\alpha(1-\alpha)\,\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,\rho^1(x)\,h\frac{x_{hl}}{n}\\ &= \alpha(1-\alpha)\,h\sum_{x\in\mathcal{X}_n:x_{hl}\geq 1}\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,\rho^1(x)\\ &= \alpha(1-\alpha)\,h\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,. \end{aligned}\tag{A44}$$

By assumption, $\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})>0$, so

$$\begin{aligned}\widehat{\gamma}_{hl}^1 &= \frac{\gamma_{hl}}{\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})}>0\\ \widehat{\phi}_{hl}^1 &= \frac{\phi_{hl}}{\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})}>0.\end{aligned}\tag{A45}$$

Substituting (A44) into (A43) and using Lemma 3, we obtain the condition

$$\begin{aligned}2\alpha(1-\alpha)\,h-\lambda_{hh}h+\lambda_{hl}h+\widehat{\gamma}_{hl}^1-\widehat{\phi}_{hl}^1 &= 2\alpha(1-\alpha)\,h-\alpha^2\Lambda h+\alpha(2-\alpha)\Lambda h+\widehat{\gamma}_{hl}^1-\widehat{\phi}_{hl}^1\\ &= 2\alpha(1-\alpha)\,h+2\alpha\Lambda h+\widehat{\gamma}_{hl}^1-\widehat{\phi}_{hl}^1=0.\end{aligned}\tag{A46}$$

By (A45), the "rescaled multipliers" are well-defined, weakly positive, and equal to zero if and only if the "original multiplier" is equal to zero. Since $2\alpha(1-\alpha)\,h+2\alpha\Lambda h>0$, we conclude that $\widehat{\phi}_{hl}^1>0$. Hence $\eta_{hl}^1=1$ for all x by the complementarity slackness condition. By Proposition 2, $\eta_{lh}^2=1$ follows. Steps 1 and 2 thus proves part (1) of the lemma.

[Step 3] To economize on derivations, we immediately observe that

$$\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl},x_{lh}+1,x_{ll})=\sum_{x\in\mathcal{X}_n:x_{lh}\geq 1}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})\rho^1(x), \tag{A47}$$

and write the optimality condition for η_{lh}^1 as

$$\begin{aligned}&2\textstyle\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{lh}}{n}l-\lambda_{hh}\sum_{x\in\mathcal{X}_n:x_{lh}\geq 1}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})\rho^1(x)h\\ &+\lambda_{hl}\textstyle\sum_{x\in\mathcal{X}_n:x_{lh}\geq 1}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})\rho^1(x)l+\gamma_{lh}-\phi_{lh}\;=\;0.\end{aligned}\tag{A48}$$

Since $\mathbf{a}_n(x) = \frac{n}{x_{lh}}\alpha(1-\alpha)\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})$ holds for $x_{lh} \geq 1$ we have

$$\begin{aligned}\sum_{x\in\mathcal{X}_n} \mathbf{a}_n(x)\rho^1(x)\frac{x_{lh}}{n}l &= \sum_{x\in\mathcal{X}_n:x_{lh}\geq 1} \mathbf{a}_n(x)\rho^1(x)\frac{x_{lh}}{n}l \\ &= \sum_{x\in\mathcal{X}_n:x_{lh}\geq 1} \frac{n}{x_{lh}}\alpha(1-\alpha)\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\rho^1(x)\frac{x_{lh}}{n}l \\ &= \alpha(1-\alpha)l\sum_{x\in\mathcal{X}_n:x_{lh}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\rho^1(x)\end{aligned} \tag{A49}$$

Substituting into (A48) and simplifying, one obtains

$$\begin{aligned}0 &= 2\alpha(1-\alpha)l - \lambda_{hh}h + \lambda_{hl}l + \widehat{\gamma}_{lh} - \widehat{\phi}_{lh} = 2\alpha(1-\alpha)l - \alpha^2 h\Lambda + \left(2\alpha-\alpha^2\right)\Lambda l + \widehat{\gamma}_{lh} - \widehat{\phi}_{lh} \\ &= \alpha(1-\alpha)(1+\Lambda)\left\{(1-\Phi)2l + \Phi\left[\frac{\left(2\alpha-\alpha^2\right)}{\alpha(1-\alpha)}l - \frac{\alpha^2}{\alpha(1-\alpha)}h\right] + \frac{\widehat{\gamma}_{lh} - \widehat{\phi}_{lh}}{(1+\Lambda)\alpha(1-\alpha)}\right\} \\ &= \alpha(1-\alpha)(1+\Lambda)\left[G(\Phi) + \frac{\widehat{\gamma}_{lh} - \widehat{\phi}_{lh}}{(1+\Lambda)\alpha(1-\alpha)}\right]\end{aligned} \tag{A50}$$

where $\widehat{\gamma}_{lh}(x)$ and $\widehat{\phi}_{lh}(x)$ are respectively $\gamma_{lh}(x)$ and $\phi_{lh}(x)$ multiplied by $1/\mathrm{E}\left[\rho^1(x)|\theta_i = lh\right]$. We thus conclude that $G(\Phi) > 0$ must imply that $\widehat{\phi}_{lh} > 0$, hence by complementary slackness, $\eta^1_{lh} = 1$. Symmetrically, $G(\Phi) < 0$ must imply that $\widehat{\gamma}_{lh} > 0$, hence $\eta^1_{lh} = 0$. If $G(\Phi) = 0$, then the value of both multipliers must be zero, which imposes no restrictions on η^1_{lh}. Proposition 2 implies that $\eta^2_{hl} = \eta^1_{lh}$, which completes the proof of part (2) of the lemma.

[Step 4] Finally, we consider the optimality condition for η^1_{ll}. Using an identity similar to (A47), we can write the first order condition for η^1_{ll} as

$$\begin{aligned}&2\sum_{x\in\mathcal{X}_n:x_{ll}\geq 1} \mathbf{a}_n(x)\rho^1(x)\frac{x_{ll}}{n} - \lambda_{hl}\sum_{x\in\mathcal{X}_n:x_{ll}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)\rho^1(x)(h+l) \\ &+\lambda_{ll}\sum_{x\in\mathcal{X}_n:x_{ll}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)\rho^1(x)2l + \gamma_{ll} - \phi_{ll} = 0.\end{aligned} \tag{A51}$$

Using the multinomial identity $\mathbf{a}_n(x) = \frac{n}{x_{ll}}(1-\alpha)^2\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1)$ we can rewrite the first order condition as

$$0 = (1-\alpha)^2 2l + \Lambda\left[2l - \left(2\alpha-\alpha^2\right)(h+l)\right] + \widehat{\gamma}_{ll} - \widehat{\phi}_{ll} \tag{A52}$$

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

where $\widehat{\gamma}_{ll}$ and $\widehat{\phi}_{ll}$ are respectively γ_{ll} and ϕ_{ll} multiplied by $1/\mathrm{E}\left[\rho^1(x)|\theta_i = ll\right]$. Arguing as in the previous case completes the proof. ■

Proof of Lemma 6.

Proof. Without loss, we only consider good 1. The first order condition with respect to $\rho^1(x)$ is

$$\begin{aligned}
&2\mathbf{a}_n(x)\left[\tfrac{(\eta^1_{hh}x_{hh}+\eta^1_{hl}x_{hl})h+(\eta^1_{lh}x_{lh}+\eta^1_{ll}x_{ll})l}{n}-c\right]+\lambda_{hh}\left[2\eta^1_{hh}\mathbf{a}_{n-1}(x_{hh}-1,x_{hl},x_{lh},x_{ll})h\right]\\
&-\lambda_{hh}\left[\eta^1_{hl}\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})h-\eta^1_{hl}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})h\right]\\
&+\lambda_{hl}\left[\eta^1_{hl}\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})h+\eta^1_{lh}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})l\right]\\
&-\lambda_{hl}\left[\eta^1_{ll}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh},x_{ll}-1)(h+l)\right]+\lambda_{ll}2\eta^1_{ll}\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh},x_{ll}-1)l\\
&-\Lambda\mathbf{a}_n(x)2c+\gamma(x)-\phi(x)=0,
\end{aligned}\tag{A53}$$

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,

$$\begin{aligned}
\mathbf{a}_{n-1}(x_{hh}-1,x_{hl},x_{lh},x_{ll})&=\tfrac{\mathbf{a}_n(x)}{\alpha^2}\tfrac{x_{hh}}{n}, &\quad \mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})&=\tfrac{\mathbf{a}_n(x)}{\alpha(1-\alpha)}\tfrac{x_{hl}}{n}\\
\mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh}-1,x_{ll})&=\tfrac{a(x)}{\alpha(1-\alpha)}\tfrac{x_{lh}}{n}, &\quad \mathbf{a}_{n-1}(x_{hh},x_{hl},x_{lh},x_{ll}-1)&=\tfrac{1}{(1-\alpha)^2}\tfrac{x_{ll}}{n},
\end{aligned}\tag{A54}$$

exploiting the relationships between multipliers in Lemma 3, and substituting $\eta^1_{hh}=\eta^1_{hl}=1$ due to Lemma 4, we can simplify (A53) to

$$\begin{aligned}
&2\left[\tfrac{(x_{hh}+x_{hl})h+(\eta^1_{lh}x_{lh}+\eta^1_{ll}x_{ll})l}{n}-c\right](1-\Phi)\\
&+\alpha^2\Phi\left[2\tfrac{1}{\alpha^2}\tfrac{x_{hh}}{n}h-\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{hl}}{n}h-\eta^1_{lh}\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{lh}}{n}h\right]\\
&+\alpha(2-\alpha)\Phi\left[\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{hl}}{n}h+\eta^1_{lh}\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{lh}}{n}l-\eta^1_{ll}\tfrac{1}{(1-\alpha)^2}\tfrac{x_{ll}}{n}(h+l)\right]\\
&+\Phi 2\eta^1_{ll}\tfrac{1}{(1-\alpha)^2}\tfrac{x_{ll}}{n}l-\Phi 2c+\tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)}=0,
\end{aligned}\tag{A55}$$

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 (A55) and simplifying we get

$$\begin{aligned}
&2\tfrac{x_{hh}}{n}h+2\tfrac{x_{hl}}{n}h-2c+\tfrac{x_{lh}}{n}\eta^1_{lh}\overbrace{\left\{(1-\Phi)2l+\Phi\left[\frac{\alpha(2-\alpha)}{\alpha(1-\alpha)}l-\frac{\alpha^2}{\alpha(1-\alpha)}h\right]\right\}}^{G(\Phi)}\\
&+\tfrac{x_{lh}}{n}\eta^1_{ll}\overbrace{\left\{(1-\Phi)2l+\Phi\left[\frac{2}{(1-\alpha)^2}l-\frac{\alpha(2-\alpha)}{(1-\alpha)^2}\frac{x_{ll}}{n}(h+l)\right]\right\}}^{H(\Phi)}+\tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)}\\
&/(32)/=2Q^1\left(\tfrac{x}{n},\Phi\right)+\tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)}=0,
\end{aligned}\tag{A56}$$

where the equality uses (from Lemma 4) that $\eta^1_{lh}=0$ if $G(\Phi)<0$ and $\eta^1_{ll}=0$ if $H(\Phi)<0$. The result follows. ■

Proof of Proposition 5.

Lemma A1 *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. Fix an arbitrary $\epsilon > 0$. Let $Y_i(\theta_i; \Phi_n)$ be a transformation of the random variable θ_i given by

$$Y_i(\theta_i; \Phi_n) = \begin{cases} h - c & \text{if } \theta_i \in \{hh, hl\} \\ \max\{0, G(\Phi_n)\} - c & \text{if } \theta_i = lh \\ \max\{0, H(\Phi_n)\} - c & \text{if } \theta_i = ll \end{cases}. \tag{A57}$$

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

$$\mathrm{E}_{\theta_i} Y_i(\theta_i; \Phi_n) = \alpha h + \alpha(1-\alpha)\max\{0, G(\Phi_n)\} + (1-\alpha)^2 \max\{0, H(\Phi_n)\} - c = Q^1(\mu, \Phi_n). \tag{A58}$$

Since for any sequence of realizations $\{y_i(\theta_i; \Phi_n)\}_{i=1}^n$

$$\sum_{i=1}^n \frac{y_i(\theta_i; \Phi_n)}{n} = \frac{x_{hh}}{n}h + \frac{x_{hl}}{n}h + \frac{x_{lh}}{n}\max\{0, G(\Phi_n)\} + \frac{x_{ll}}{n}\max\{0, H(\Phi_n)\} - c = Q^1\left(\frac{x}{n}, \Phi_n\right), \tag{A59}$$

we can apply Chebyshev's inequality to obtain

$$\begin{aligned} \Pr\left(\left|Q^1\left(\frac{x}{n}, \Phi_n\right) - Q^1(\mu, \Phi_n)\right| \geq \epsilon\right) &= \Pr\left(\left|\sum_{i=1}^n \frac{y_i(\theta_i; \Phi_n)}{n} - \mathrm{E}_{\theta_i} Y_i(\theta_i; \Phi_n)\right| \geq \epsilon\right) \\ &\leq \frac{\mathrm{Var}[Y_i(\theta_i; \Phi_n)]}{n\epsilon^2} \leq \frac{\sigma^2}{n\epsilon^2}. \end{aligned} \tag{A60}$$

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 $n \geq N = \sigma^2/\epsilon^3 < \infty$. ■

Lemma A2 *Let Y be a random variable with Binomial (n, p) distribution. For any $\epsilon > 0$ and $p \in (0, 1)$ there exists $N < \infty$ such that the binomial distribution with parameters p, n satisfies*

$$\Pr(Y = y) = \frac{n!}{y!(n - y!)} p^y (1-p)^{n-y} \leq \epsilon$$

for every $n \geq N$ and $y \in \{0, ..., n\}$.

Proof. Omitted. ■

Lemma A3 *For every $\epsilon > 0$ there exists N such that $\left|\rho_i^1(\theta_i) - \rho_i^1(\theta_i')\right| \leq \epsilon$ for every $\theta_i, \theta_i' \in \Theta$ in any truth-telling mechanism for any economy where $n \geq N$.*

Proof. Omitted. ■

The implication of Lemma A3 is as follows: as $n \to \infty$, the perceived provision probability of public goods are little affected by agent i's own announcement; thus such perceived provision probability must be near the ex ante probability of providing the good.

Lemma A4 *For every $\epsilon > 0$, there exists N such that, for all $n \geq N$, $\left|\mathrm{E}\rho^1(x) - \rho_i^1(\theta_i)\right| \leq \epsilon$ for all $\theta_i \in \Theta$ in any truth-telling mechanism.*

Proof. Fix $\epsilon > 0$ arbitrarily. Let N be such that $\left|\rho_i^1(\theta_i) - \rho_i^1(\theta_i')\right| \leq \epsilon$ for every $n \geq N, \theta_i, \theta_i' \in \Theta$. Then

$$\begin{aligned} & \left|\mathrm{E}\rho^1(x) - \rho_i^1(\theta_i)\right| \\ = & \left|\alpha^2\rho_i^1(hh) + \alpha(1-\alpha)\rho_i^1(hl) + \alpha(1-\alpha)\rho_i^1(lh) + \left(1-\alpha^2\right)\rho_i^1(ll) - \rho_i^1(\theta_i)\right| \\ \leq & \alpha^2\left|\rho_i^1(hh) - \rho_i^1(\theta_i')\right| + \alpha(1-\alpha)\left|\rho_i^1(hl) - \rho_i^1(\theta_i)\right| \\ & +\alpha(1-\alpha)\left|\rho_i^1(lh) - \rho_i^1(\theta_i)\right| + (1-\alpha)^2\left|\rho_i^1(ll) - \rho_i^1(\theta_i)\right| \\ \leq & \alpha^2\epsilon + \alpha(1-\alpha)\epsilon + \alpha(1-\alpha)\epsilon + (1-\alpha)^2\epsilon = \epsilon. \end{aligned} \tag{A61}$$

■

(Proof of Proposition 5, continued). Now we use the above lemmas to prove Proposition *5*. We prove the four parts of the proposition in order.

(Part 1) We first prove part 1. Note from (33), we know that $Q^1(\mu, \Phi_n) \geq \alpha h - c$ for any $\Phi_n \in [0,1]$, hence $\lim_{n\to\infty} Q^1(\mu, \Phi_n) \geq \alpha h - c$. Thus if $\alpha h > c$, part 1 of the proposition immediately follows from Lemmas 6 and A1. Suppose instead that $\alpha(2-\alpha)(h+l) > 2c \geq 2\alpha h$. Then,

$$\begin{aligned} Q^1(\mu, \Phi_n) &= \alpha h + \alpha(1-\alpha)\max\{0, G(\Phi_n)\} + (1-\alpha)^2\max\{0, H(\Phi_n)\} - c \\ &\geq \alpha h - c + \alpha(1-\alpha)G(\Phi_n) \\ &= \alpha h - c + \alpha(1-\alpha)\left\{l(1-\Phi_n) + \Phi_n\left[\frac{2\alpha-\alpha^2}{2\alpha(1-\alpha)}l - \frac{\alpha^2}{2\alpha(1-\alpha)}h\right]\right\} \\ &= (1-\Phi_n)\left[\alpha h + \alpha(1-\alpha)l - c\right] + \Phi_n\left[\frac{\alpha(2-\alpha)(l+h)}{2} - c\right]. \end{aligned} \tag{A62}$$

Observe that

$$\alpha h + \alpha(1-\alpha)l = \frac{\alpha(2-\alpha)(l+h)}{2} + \frac{\alpha^2}{2}(h-l) > \frac{\alpha(2-\alpha)(l+h)}{2}. \tag{A63}$$

Hence, $Q^1(\mu, \Phi_n) \geq \frac{\alpha(2-\alpha)(l+h)}{2} - c > 0$ if $\alpha(2-\alpha)(h+l) > 2c$, then for all $\Phi_n \in [0,1]$, implying that $\lim_{n\to\infty} Q^1(\mu, \Phi_n) > 0$. Thus by Lemmas 6 and A1, $\lim_{n\to\infty} \mathrm{E}\rho_n^j(x) \to 1$ for $j = 1, 2$. This proves Part 1.

(Part 2) We now prove part 2. Suppose to the contrary that there exists a (sub) sequence of optimal incentive compatible, balanced-budget voluntary mechanism with provision rules for public good 1, $\rho_n^1(x)$, such that $\lim_{n\to\infty} \mathrm{E}\rho_n^1(x) = \rho > 0$. We will now derive a contradiction that the mechanism can not have a balanced budget.

Now we can use the definition of $\rho_i^j(\theta_i)$ in (B95) to re-write the incentive compatibility constraint (18b), after using the characterization of inclusion rule in Lemma 5, as

$$\rho_i^1(hl)\,h + \rho_i^1(lh)\,\eta_{lh}^1 l - t_{hl} \geq \rho_i^1(ll)\,\eta_{ll}^1(h+l) - t_{ll} \geq \rho_i^1(ll)\,\eta_{ll}^1(h-l)\,, \tag{A64}$$

where the second inequality comes from the participation constraint (19). Pick an arbitrary $\epsilon > 0$. Then, by Lemma A4, there exists finite N such that for every $n \geq N_1$ and each $\theta_i \in \Theta$, for $j = 1, 2$,

$$\left|\rho_i^j(\theta_i) - \mathrm{E}\rho_n^1(x)\right| < \epsilon_1 \equiv \frac{\epsilon}{3h}. \tag{A65}$$

Substituting (A65) into (A64), we obtain that for all $n \geq N_1$,

$$\left[\mathrm{E}\rho_n^1(x) + \epsilon_1\right]\left(h + \eta_{lh}^1 l\right) - t_{hl} \geq \left[\mathrm{E}\rho_n^1(x) - \epsilon_1\right]\eta_{ll}^1(h-l)\,, \tag{A66}$$

which implies that

$$\begin{aligned} t_{hl} &\leq \mathrm{E}\rho_n^1(x)\left[h\left(1-\eta_{ll}^1\right) + \left(\eta_{lh}^1 + \eta_{ll}^1\right)l\right] + \epsilon_1\left[h + \eta_{lh}^1 l + \eta_{ll}^1(h-l)\right] \\ &< \mathrm{E}\rho_n^1(x)\left[h\left(1-\eta_{ll}^1\right) + \left(\eta_{lh}^1 + \eta_{ll}^1\right)l\right] + \underbrace{3h\epsilon_1}_{\epsilon}. \end{aligned} \tag{A67}$$

Similarly, incentive constraints (18a) can be rewritten as:

$$t_{hh} \leq 2\rho_i^1(hh)\,h - \left[\rho_i^1(hl) + \rho_i^1(lh)\,\eta_{lh}^1\right]h + t_{hl}. \tag{A68}$$

Again, by Lemma A4, there exist N_2 such that for all $n > N_2$,

$$\begin{aligned} t_{hh} &< 2\left[\mathrm{E}\rho_n^1(x)\right]h - \mathrm{E}\rho_n^1(x)\left(1+\eta_{lh}^1\right)h + t_{hl} + \epsilon \\ &= \mathrm{E}\rho_n^1(x)\left(1-\eta_{lh}^1\right)h + t_{hl} + \epsilon \\ &< \mathrm{E}\rho_n^1(x)\left(1-\eta_{lh}^1\right)h + \mathrm{E}\rho_n^1(x)\left[h\left(1-\eta_{ll}^1\right) + \left(\eta_{lh}^1+\eta_{ll}^1\right)l\right] + \epsilon \\ &= \mathrm{E}\rho_n^1(x)\left[\left(2-\eta_{ll}^1-\eta_{lh}^1\right)h + \left(\eta_{lh}^1+\eta_{ll}^1\right)l\right] + \epsilon. \end{aligned} \tag{A69}$$

Finally, from the participation constraint (19), there exists N_3 such that for all $n > N_3$,

$$t_{ll} < 2\mathrm{E}\rho_n^1(x)\,\eta_{ll}^1 l + \epsilon. \tag{A70}$$

Now consider two cases:

<u>CASE 1:</u> $\eta_{ll}^1 = \eta_{ll}^2 = 0$ and $\eta_{lh}^1 = \eta_{hl}^2 = \eta_m \in (0,1)$. In this case, we have $t_{ll} = 0$ from type-ll's participation constraint. Using (A67)-(A70), we can bound the total expected tax revenue as follows:

$$\begin{aligned} &\alpha^2 t_{hh} + \alpha(1-\alpha)(t_{hl}+t_{lh}) + (1-\alpha)^2 t_{ll} \qquad\qquad (A71) \\ <\ &\alpha^2\left\{\mathrm{E}\rho_n(x)\left[(2-\eta_m)h + \eta_m l\right] + \epsilon\right\} + 2\alpha(1-\alpha)\left\{\mathrm{E}\rho_n(x)(h+\eta_m l) + \epsilon\right\} \\ =\ &\mathrm{E}\rho_n(x)\left\{\left[\alpha^2(2-\eta_m) + 2\alpha(1-\alpha)\right]h + \left[\alpha^2 + 2\alpha(1-\alpha)\right]\eta_m l\right\} + \epsilon' \\ =\ &\mathrm{E}\rho_n(x)\underbrace{\left\{\left[\alpha^2(2-\eta_m) + 2\alpha(1-\alpha)\right]h + \alpha(2-\alpha)\eta_m l\right\}}_{\equiv Z_1(\eta_m)} + \epsilon' \end{aligned}$$

Note that

$$\frac{\partial Z_1(\eta_m)}{\partial \eta_m} = \alpha(2-\alpha)l - \alpha^2 h = [\alpha(2-\alpha)(h+l)] - 2\alpha h. \tag{A72}$$

Therefore,

$$Z_1(\eta_m) < \begin{cases} Z(1) = \alpha(2-\alpha)(h+l) & \text{if } 2\alpha h \le \alpha(2-\alpha)(h+l) \\ Z(0) = 2\alpha h & \text{if } 2\alpha h > \alpha(2-\alpha)(h+l), \end{cases} \tag{A73}$$

which implies that

$$\alpha^2 t_{hh} + \alpha(1-\alpha)(t_{hl}+t_{lh}) + (1-\alpha)^2 t_{ll} < \mathrm{E}\rho_n(x)\max\{2\alpha h, \alpha(2-\alpha)(h+l)\} + \epsilon'. \tag{A74}$$

Thus if $\max\{2\alpha h, \alpha(2-\alpha)(h+l)\} < 2c$, then the budget balance condition can not be satisfied when n is sufficiently large.

<u>CASE 2:</u> $\eta_{ll}^1 = \eta_{ll}^2 = \eta_l \in (0,1), \eta_{lh}^1 = \eta_{hl}^2 = 1$. We can again use (A67)-(A70) to bound the total expected tax revenue as follows:

$$\begin{aligned}
& \alpha^2 t_{hh} + \alpha(1-\alpha)(t_{hl}+t_{lh}) + (1-\alpha)^2 t_{ll} && \text{(A75)}\\
< \; & \alpha^2\{\mathrm{E}\rho_n(x)[(1-\eta_l)h + (1+\eta_l)l] + \epsilon\} + \alpha(1-\alpha)2\mathrm{E}\rho_n(x)\{[h(1-\eta_l)+(1+\eta_l)l]+\epsilon\} \\
& + (1-\alpha)^2[2\mathrm{E}\rho_n(x)\eta_l l + \epsilon] \\
= \; & \mathrm{E}\rho_n(x)\left\{\left[\alpha^2 + 2\alpha(1-\alpha)\right](1-\eta_l)h + \left[\alpha^2+2\alpha(1-\alpha)\right](1+\eta_l)l + 2(1-\alpha)^2\eta_l l\right\} + \epsilon \\
= \; & \mathrm{E}\rho_n(x)\left[\underbrace{\alpha(2-\alpha)(1-\eta_l)h + \alpha(2-\alpha)(1+\eta_l)l + 2(1-\alpha)^2\eta_l l}_{Z_2(\eta_l)}\right] + \epsilon
\end{aligned}$$

Note that $Z_2(0) = \alpha(2-\alpha)(h+l)$ and $Z_2(1) = 2\alpha(2-\alpha)l + 2(1-\alpha)^2 l = 2l$. Since $Z_2(\eta_l)$ is linear in η_l, we have

$$Z_2(\eta_l) \le \max\{Z_2(0), Z_2(1)\} = \max\{\alpha(2-\alpha)(h+l), 2l\}. \tag{A76}$$

If $\max\{2\alpha h, \alpha(2-\alpha)(h+l)\} < 2c$, then $\max\{\alpha(2-\alpha)(h+l), 2l\} < 2c$ since by assumption $l < c$. Therefore there exists N' such that for all $n > N'$, the budget balance condition will not be satisfied under any incentive compatible voluntary mechanism.

(PART 3) Suppose to the contrary that there does not exist N such that $\eta_n^1(lh) = \eta_n^2(hl) = 1$ for all $n \ge N$. Then, taking a subsequence if necessary, we have that $\eta_n^1(lh) = \eta_n^2(hl) < 1$ for all n, which, by Lemma 5, implies that $\eta_n^j(ll) = 0$ for all n in the sequence. The per capita surplus generated by the optimal mechanism $\mathcal{M}_n$ in the n^{th} economy in the sequence, denoted by $S(\mathcal{M}_n)$, is then

$$\begin{aligned}
\frac{S(\mathcal{M}_n)}{n} &= \frac{2\mathrm{E}\rho_n^1(x)\left[(x_{hh}+x_{hl})h + \left(\eta_n^1(lh)x_{lh} + \eta_n^1(ll)x_{ll}\right)l - cn\right]}{n} && \text{(A77)}\\
&\le \frac{2\mathrm{E}\left[(x_{hh}+x_{hl})h + x_{lh}l - \rho_n^1(x)cn\right]}{n} = 2[\alpha h + \alpha(1-\alpha)l] - 2\mathrm{E}\rho_n^1(x)c
\end{aligned}$$

From Part 2, we know $\mathrm{E}\rho_n^1(x) \to 1$ as $n \to \infty$. Thus each $\varepsilon > 0$ there exists N such that

$$\frac{S(\mathcal{M}_n)}{n} \leq 2[\alpha h + \alpha(1-\alpha)l - c] + \varepsilon. \tag{A78}$$

Now we show that $\mathcal{M}_n$ can be dominated by an alternative mechanism as $n \to \infty$. Consider a sequence of mechanisms $\left\{\widetilde{\mathcal{M}}_n\right\}_{n=1}^{\infty}$, where, for each n,

$$\begin{aligned}
\widetilde{\eta}_n^1(lh) &= \widetilde{\eta}_n^2(hl) = 1 \qquad \text{(A79)}\\
\widetilde{\eta}_n^1(ll) &= \widetilde{\eta}_n^2(ll) = \eta_{ll}^* = \frac{\alpha(2-\alpha)[h+l]-2c}{\alpha(2-\alpha)[h+l]-2l}\\
\widetilde{t}_n(hh) &= \widetilde{t}_n(hl) = \widetilde{t}_n(lh) = (1-\eta_{ll}^*)(h+l) + \eta_{ll}^* 2l\\
\widetilde{t}_n(ll) &= 2\eta_{ll}^* l\\
\widetilde{\rho}_n^j(x) &= 1 \text{ for all } x \in \mathcal{X}_n
\end{aligned}$$

We observe that the participation constraint for type ll holds with equality since

$$\mathrm{E}_{-i}\sum_{j=1,2}\widetilde{\rho}_n^j(x)\widetilde{\eta}_n^j(ll)\,l - \widetilde{t}_n(ll) = 2\eta_{ll}^* l - 2\eta_{ll}^* l = 0. \tag{A80}$$

The downward incentive constraint for type hl also holds with equality since

$$\begin{aligned}
&\mathrm{E}_{-i}\left[\widetilde{\rho}_n^1(x)\widetilde{\eta}_n^1(hl)\,h + \widetilde{\rho}_n^2(x)\widetilde{\eta}_n^2(hl)\,l - \widetilde{t}_n(hl)\right] = h + l - \widetilde{t}_n(hl) \qquad \text{(A81)}\\
= \;& h + l - [(1-\eta_{ll}^*)(h+l) + \eta_{ll}^* 2l] = \eta_{ll}^*(h+l) - \eta_{ll}^* 2l\\
= \;& \mathrm{E}_{-i}\left[\widetilde{\rho}_n^1(x)\widetilde{\eta}_n^1(hl)\,h + \widetilde{\rho}_n^2(x)\widetilde{\eta}_n^2(hl)\,l - \widetilde{t}_n(hl)\middle|\, \theta_i = ll\right].
\end{aligned}$$

Similarly, the downward incentive constraints and participation constraints for all other types of agents also hold. Finally, $\widetilde{\mathcal{M}}_n$ is also budget balanced for all n since, with some algebra, one can show that

$$\mathrm{E}\left(\sum_{i\in\mathcal{I}}\widetilde{t}_n(\theta_i) - \sum_{j=1,2}\widetilde{\rho}_n^j(x)\,cn\right) = 0. \tag{A82}$$

Now, the expected per capita surplus generates by $\widetilde{\mathcal{M}}_n$ is

$$\begin{aligned}
\frac{S\left(\widetilde{\mathcal{M}}_n\right)}{n} &= \alpha^2 2h + 2\alpha(1-\alpha)(h+l) + (1-\alpha)^2\eta_{ll}^* 2l - 2c \qquad \text{(A83)}\\
&= 2[\alpha h + \alpha(1-\alpha)l - c] + (1-\alpha)^2\eta_{ll}^* 2l
\end{aligned}$$

Let $\varepsilon = (1-\alpha)^2\eta_{ll}^* l > 0$, we know from (A78) that there exists $N < \infty$ such that

$$\frac{S(\mathcal{M}_n)}{n} \leq 2[\alpha h + \alpha(1-\alpha)l - c] + \varepsilon = \frac{S\left(\widetilde{\mathcal{M}}_n\right)}{n} - \varepsilon < \frac{S\left(\widetilde{\mathcal{M}}_n\right)}{n}, \tag{A84}$$

which implies that mechanisms $\mathcal{M}_n$ could not be optimal for $n \geq N$, a contradiction.

Now we have concluded that in the sequence $\{\mathcal{M}_n\}$, $\eta_n^1(lh) = \eta_n^2(hl) = 1$ for every $n \geq N$. What is left to show is that $\eta_n^1(ll)$ does converge to η_{ll}^* in the sequence $\{\mathcal{M}_n\}$. Suppose first that there exists a subsequence such that $\eta_n^1(ll) \to \eta' < \eta_{ll}^*$. An argument as the one above shows that, for every $\varepsilon > 0$, there exists $N < \infty$ such that

$$\frac{S(\mathcal{M}_n)}{n} \leq 2\left[\alpha h + \alpha(1-\alpha)l + (1-\alpha)^2 \eta' l - c\right] + \varepsilon. \tag{A85}$$

Again consider the alternative sequence of mechanisms $\left\{\widetilde{\mathcal{M}}_n\right\}$ constructed above. Pick $\varepsilon = (1-\alpha)^2(\eta_{ll}^* - \eta')l$, we find that

$$\frac{S\left(\widetilde{\mathcal{M}}_n\right)}{n} - \frac{S(\mathcal{M}_n)}{n} \geq (1-\alpha)^2\left(\eta_{ll}^* - \eta'\right)2l - \varepsilon = (1-\alpha)^2\left(\eta_{ll}^* - \eta'\right)l > 0. \tag{A86}$$

thus again contradicts the optimality of the mechanism $\mathcal{M}_n$ is better when n is sufficiently large.

Finally, suppose there is a subsequence such that $\eta_n^j(ll) \to \eta' > \eta_{ll}^*$. We now argue that such a mechanism could not be budget balanced. Let

$$\varepsilon = \frac{(\eta' - \eta_{ll}^*)\left[\alpha(2-\alpha)(h+l) - 2l\right]}{4(1+\alpha)} > 0. \tag{A87}$$

Then, since $\eta_n^1(ll) + \eta_n^2(ll) \to 2\eta'$ it follows that to satisfy the participation constraint for type ll for all n there must be some N_1 such that $t_n(ll) \leq 2\eta' l + \varepsilon$ for all $n \geq N_1$. Moreover, there exists N_2 such that $\eta_n^1(lh) = \eta_n^2(hl) = 1$for $n \geq N_2$. Thus the incentive constraint that type hl does not imitate type ll reduces to

$$\begin{aligned}
\rho_{in}^1(hl)h + \rho_{in}^2(hl)l - t_n(hl) &\geq \rho_{in}^1(ll)\eta_n^1(ll)h + \rho_{in}^2(ll)\eta_n^2(ll)l - t_n(ll) \\
\Rightarrow\quad t_n(hl) &\leq t_n(ll) + \left[\rho_{in}^1(hl) - \rho_{in}^1(ll)\eta_n^1(ll)\right]h + \left[\rho_{in}^2(hl) - \rho_{in}^2(ll)\eta_n^2(ll)\right]l
\end{aligned} \tag{A88}$$

By Lemma A3, $\lim_{n\to\infty}\rho_{in}^j(hl) = \lim_{n\to\infty}\rho_{in}^j(hl) = \lim_{n\to\infty}\mathrm{E}\rho_n^j(x) = 1$. This, together with the assumption that $\lim_{n\to\infty}\eta_n^1(ll) = \eta'$, implies that there exists N_3 such that

$$t_n(hl) \leq t_n(ll) + \left(1-\eta'\right)(h+l) + \varepsilon. \tag{A89}$$

Similarly, the incentive constraint that type hh does not announce hl implies that $t_n(hh) \leq t_n(hl) + \varepsilon$. Hence, the expected per capita revenue of the mechanism satisfies

$$\begin{aligned}
& \alpha^2 t_n(hh) + 2\alpha(1-\alpha)t_n(hl) + (1-\alpha)^2 t_n(ll) \\
\leq\ & \left[\alpha^2 + 2\alpha(1-\alpha)\right]t_n(hl) + \alpha^2\varepsilon + (1-\alpha)^2 t_n(ll) \\
\leq\ & \left[\alpha^2 + 2\alpha(1-\alpha)\right]\left[t_n(ll) + \left(1-\eta'\right)(h+l) + \varepsilon\right] + (1-\alpha)^2 t_n(ll) + \alpha^2\varepsilon \\
=\ & t_n(ll) + \left[\alpha^2 + 2\alpha(1-\alpha)\right]\left(1-\eta'\right)(h+l) + 2\alpha\varepsilon \\
\leq\ & 2\eta' l + \varepsilon + \left[\alpha^2 + 2\alpha(1-\alpha)\right]\left(1-\eta'\right)(h+l) + 2\alpha\varepsilon \\
=\ & \eta' 2l + \left(1-\eta'\right)\alpha(2-\alpha)(h+l) + \varepsilon(1+2\alpha).
\end{aligned} \tag{A90}$$

Since there exists $N_4 < \infty$ such that $\mathrm{E}\left[\rho_n^1(x) + \rho_n^2(x)\right] c \geq 2c - \varepsilon$, we have

$$\begin{aligned}
& \alpha^2 t_n(hh) + 2\alpha(1-\alpha) t_n(hl) + (1-\alpha)^2 t_n(ll) - \mathrm{E}\left[\rho_n^1(x) + \rho_n^2(x)\right] c \qquad \text{(A91)} \\
\leq\ & \left(\eta' - \eta_{ll}^*\right)(2l - \alpha(2-\alpha)(h+l)) + 2\varepsilon(1+\alpha) \\
=\ & \left(\eta' - \eta_{ll}^*\right)(2l - \alpha(2-\alpha)(h+l)) + \frac{\left(\eta' - \eta_{ll}^*\right)\left[\alpha(2-\alpha)(h+l) - 2l\right]}{2} \\
=\ & -\frac{\left(\eta' - \eta_{ll}^*\right)\left[\alpha(2-\alpha)(h+l) - 2l\right]}{2} < 0.
\end{aligned}$$

Hence, the mechanism must violate the balanced-budget constraint for $n \geq \max\{N_1, N_2, N_3, N_4\}$. We conclude that there can be no subsequence of optimal mechanisms such that $\eta_n^j(ll) \to \eta' \neq \eta_{ll}^*$, proving the claim.

(Part 4) This part is proved analogous to Part 3. Suppose to the contrary that in the sequence of mechanisms $\{\mathcal{M}_n\}$, there exists no N such that $\eta_n^1(ll) = \eta_n^2(ll) = 0$ for all $n \geq N$. Then there must be a subsequence where $\eta_n^1(ll) = \eta_n^2(ll) > 0$, which from Lemma (5) we know that $\eta_n^1(lh) = \eta_n^2(hl) = 1$ for all n along the subsequence. Hence $\lim_{n\to\infty} \eta_n^1(lh) = \lim_{n\to\infty} \eta_n^2(hl) = 1$ and $\lim_{n\to\infty} \eta_n^1(ll) = \lim_{n\to\infty} \eta_n^2(ll) = \eta' \geq 0$. Let

$$\varepsilon = \frac{2c - \alpha(2-\alpha)(h+l)}{4(1+\alpha)} > 0. \qquad \text{(A92)}$$

We can then use the same calculations as in Part 3 to conclude that there exists $N < \infty$ such that the revenues collected satisfy

$$\begin{aligned}
\alpha^2 t_n(hh) + 2\alpha(1-\alpha) t_n(hl) + (1-\alpha)^2 t_n(ll) &< \eta' 2l + \left(1-\eta'\right)\alpha(2-\alpha)(h+l) + \varepsilon(1+2\alpha) \\
&\leq \alpha(2-\alpha)(h+l) + \varepsilon(1+2\alpha). \qquad \text{(A93)}
\end{aligned}$$

Moreover, there exists N_2 such that $\mathrm{E}\left[\rho_n^1(x) + \rho_n^2(x)\right] c \geq 2c - \varepsilon$, hence

$$\begin{aligned}
& \alpha^2 t_n(hh) + 2\alpha(1-\alpha) t_n(hl) + (1-\alpha)^2 t_n(ll) - \mathrm{E}\left[\rho_n^1(x) + \rho_n^2(x)\right] c \qquad \text{(A94)} \\
\leq\ & \alpha(2-\alpha)(h+l) + \varepsilon(2+2\alpha) - 2c = -\frac{2c - \alpha(2-\alpha)(h+l)}{2} < 0,
\end{aligned}$$

violating the balanced-budget constraint. Establishing that $\lim_{n\to\infty} \eta_n^1(lh) = \lim_{n\to\infty} \eta_n^2(hl) = \eta_{lh}^*$ proceeds along the same lines as those in Part 3. ■

B Omitted Proofs

Lemma A2 *Let Y be a random variable with Binomial (n, p) distribution. For any $\epsilon > 0$ and $p \in (0,1)$ there exists $N < \infty$ such that the binomial distribution with parameters p, n satisfies*

$$\Pr(Y = y) = \frac{n!}{y!\,(n-y!)} p^y (1-p)^{n-y} \leq \epsilon$$

for for every $n \geq N$ and $y \in \{0, ..., n\}$.

Proof. Fix an arbitrary $\epsilon > 0$. The most probable value for y is the unique integer $y^*(n)$ satisfying $np - 1 \leq y^*(n) \leq np + 1$, and the corresponding probability is

$$\Pr(y^*(n)) = \frac{n!}{y^*(n)!\,[n - y^*(n)]!} p^{y^*(n)} (1-p)^{n-y^*(n)}.$$

Let

$$s(r) = \frac{r!}{\sqrt{2\pi} e^{-r} r^{r+1/2}}.$$

By Stirling's Formula, for every $\epsilon > 0$ there exists $R(\epsilon)$ such that $|s(r) - 1| < \epsilon$ for all $r \geq R(\epsilon)$. Observing that

$$\begin{aligned} n! &= s(n)\sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}, \\ y^*(n)! &= s(y^*(n))\sqrt{2\pi} e^{-y^*(n)} y^*(n)^{y^*(n)+\frac{1}{2}}, \\ [n - y^*(n)]! &= s(n - y^*(n))\sqrt{2\pi} e^{-(n-y^*(n))} [n - y^*(n)]^{n-y^*(n)+\frac{1}{2}}, \end{aligned}$$

we obtain

$$\begin{aligned} &\frac{n!}{y^*(n)!\,[n - y^*(n)]!} \\ = \;& \frac{s(n)}{s(y^*(n))\,s(n - y^*(n))} \frac{\sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}}{\sqrt{2\pi} e^{-y^*(n)} y^*(n)^{y^*(n)+\frac{1}{2}} \sqrt{2\pi} e^{-(n-y^*(n))} [n - y^*(n)]^{n-y^*(n)+\frac{1}{2}}} \\ = \;& \frac{s(n)}{s(y^*(n))\,s(n - y^*(n))} \frac{n^{n+\frac{1}{2}}}{\sqrt{2\pi} y^*(n)^{y^*(n)+\frac{1}{2}} [n - y^*(n)]^{n-y^*(n)+\frac{1}{2}}}. \end{aligned}$$

Note that for any $p \in (0,1)$, $\lim_{n\to\infty} y^*(n) = \infty$ and $\lim_{n\to\infty} [n - y^*(n)] = \infty$. Hence, there exists $N < \infty$ such that $y^*(n) \geq R(\epsilon)$ and $n - y^*(n) \geq R(\epsilon)$, implying that $s(n) \leq 1+\epsilon$, $s(y^*(n)) \geq 1-\epsilon$, and $s(n - y^*(n)) \geq 1 - \epsilon$. We can thus bound the probability of $y^*(n)$ by.

$$\begin{aligned} &\Pr(y^*(n)) \\ = \;& \frac{s(n)}{s(y^*(n))\,s(n - y^*(n))} \frac{n^{n+\frac{1}{2}}}{\sqrt{2\pi} y^*(n)^{y^*(n)+\frac{1}{2}} [n - y^*(n)]^{n-y^*(n)+\frac{1}{2}}} p^{y^*(n)} (1-p)^{n-y^*(n)} \\ \leq \;& \frac{(1+\epsilon)}{(1-\epsilon)^2} \frac{n^{n+\frac{1}{2}}}{\sqrt{2\pi} y^*(n)^{y^*(n)+\frac{1}{2}} [n - y^*(n)]^{n-y^*(n)+\frac{1}{2}}} p^{y^*(n)} (1-p)^{n-y^*(n)} \\ = \;& \frac{(1+\epsilon)}{(1-\epsilon)^2} \frac{p^{y^*(n)} (1-p)^{n-y^*(n)}}{\sqrt{2n\pi} \left[\frac{y^*(n)}{n}\right]^{y^*(n)+\frac{1}{2}} \left[\frac{n-y^*(n)}{n}\right]^{n-y^*(n)+\frac{1}{2}}}. \end{aligned}$$

Since $y^*(n)/n = \arg\max_{p\in[0,1]} p^{y^*(n)}(1-p)^{n-y^*(n)}$, we know that

$$\frac{p^{y^*(n)}(1-p)^{n-y^*(n)}}{\left[\frac{y^*(n)}{n}\right]^{y^*(n)}\left[\frac{n-y^*(n)}{n}\right]^{n-y^*(n)}} \leq 1.$$

Therefore,

$$\begin{aligned}\Pr(y^*(n)) &\leq \frac{(1+\epsilon)}{(1-\epsilon)^2}\frac{1}{\sqrt{2n\pi}\left[\frac{y^*(n)}{n}\right]^{\frac{1}{2}}\left[\frac{n-y^*(n)}{n}\right]^{\frac{1}{2}}} \\ &\leq \frac{(1+\epsilon)}{(1-\epsilon)^2}\frac{1}{\sqrt{2n\pi\left(p-\frac{1}{n}\right)\left(1-p-\frac{1}{n}\right)}} \to 0 \text{ as } n\to\infty.\end{aligned}$$

Hence, there exists $N' < \infty$ such that

$$\frac{(1+\epsilon)}{(1-\epsilon)^2}\frac{1}{\sqrt{2n\pi\left(p-\frac{1}{n}\right)\left(1-p-\frac{1}{n}\right)}} \leq \epsilon.$$

Implying that $\Pr(y^*(n)) \leq \epsilon$ for any $n \geq \max\{N, N'\}$. Since ϵ was arbitrary the result follows. ■

Now let

$$\rho_i^j(\theta_i) = \mathrm{E}\left[\rho^j(x)\,|\theta_i\right] \tag{B95}$$

be agent i's perceived probability that public good j will be provided when agent i announces type θ_i. The following lemma shows that as $n\to\infty$, agent i's announcement would not affect the perceived probability of provision, i.e., the probability of any individual agent being pivotal approaches zero as $n\to\infty$:

Lemma A3 *For every $\epsilon > 0$ there exists N such that $\left|\rho_i^1(\theta_i) - \rho_i^1(\theta_i')\right| \leq \epsilon$ for every $\theta_i, \theta_i' \in \Theta$ in any truth-telling mechanism for any economy where $n \geq N$.*

Proof. We only prove the result for $(\theta_i, \theta_i') = (hh, ll)$. The proof for other type combinations proceed step by step in the same way and are left to the reader. Using the now-standard recursive formula for multinomial probability mass function, we have

$$\begin{aligned}\rho_i^1(hh) &= \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\left[\rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll})\right] \\ \rho_i^1(ll) &= \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\left[\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\right].\end{aligned}$$

Let $\overline{\rho}^1$ maximize the difference between $\rho_i^1(hh)$ and $\rho_i^1(ll)$ and let $\overline{\rho}_i^1(hh)$ and $\overline{\rho}_i^1(ll)$ be the perceived provision probabilities when the provision rule is $\overline{\rho}^1$. That is,

$$\overline{\rho}^1 \in \arg\max_{\rho^1:\mathcal{X}_n\to[0,1]} \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\left[\rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll}) - \rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\right], \tag{B96}$$

It is clear that the solution to (B96) is given by

$$\overline{\rho}^1(x) = \begin{cases} 1 & \text{if } \mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll}) \geq \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1) \\ 0 & \text{if } \mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll}) < \mathbf{a}_{-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1). \end{cases} \tag{B97}$$

Using the explicit formula for $\mathbf{a}_{n-1}(x)$, we can express (B97) as

$$\overline{\rho}^1(x) = \begin{cases} 1 & \text{if } \frac{x_{hh}}{\alpha^2} \geq \frac{x_{ll}}{(1-\alpha)^2} \\ 0 & \text{if } \frac{x_{hh}}{\alpha^2} < \frac{x_{ll}}{(1-\alpha)^2}. \end{cases} \tag{B98}$$

Fix an arbitrary $\epsilon > 0$ and let $m = x_{hl} + x_{lh} \leq n-1$. Since m is a binomial random variable with parameters $p = 2\alpha(1-\alpha)$ and $n-1$, we know, by law of large numbers, that there exists $N < \infty$ such that

$$\Pr\left(\frac{m}{n-1} \geq 2\alpha(1-\alpha) + \epsilon\right) \leq \frac{\epsilon}{2} \tag{B99}$$

for every $n \geq N$. Moreover, conditional on m, x_{hh} is binomially distributed with parameters $p' = \alpha^2 / [1 - 2\alpha(1-\alpha)]$ and $n-1-m$. Thus, we know from (B98) that, conditional on m, there exists a single value $\overline{x}_{hh}(m)$ such that $\overline{\rho}^1(\overline{x}_{hh}(m)+1, x_{hl}, x_{lh}, x_{ll}) = 1$ and $\overline{\rho}^1(\overline{x}_{hh}(m), x_{hl}, x_{lh}, x_{ll}+1) = 0$; and for all other realizations the of x_{hh}, the provision probability is unaffected by agent i's announcement. Lemma A2 implies that there exists $N' < \infty$ such that

$$\Pr\left(x_{hh} = \overline{x}_{hh}(m) \middle| m\right) \leq \frac{\epsilon}{2} \tag{B100}$$

for all n such that $n-1-m \geq N'$.

Now let $n^* = \max\left\{N, \frac{N'}{1-2\alpha(1-\alpha)-\epsilon} + 1\right\} < \infty$. Then, $N' \leq (n-1)[1 - 2\alpha(1-\alpha) - \epsilon]$ for all $n \geq n^*$. Hence, for all $n \geq n^*$,

$$\begin{aligned} &\Pr[n-1-m \leq N'] = \Pr[m \geq (n-1) - N'] \\ &\leq \Pr[m \geq (n-1) - (n-1)[1 - 2\alpha(1-\alpha) - \epsilon]] \\ &= \Pr\left[\tfrac{m}{n-1} \geq 2\alpha(1-\alpha) + \epsilon\right] \leq \tfrac{\epsilon}{2} \end{aligned} \tag{B101}$$

where the last equality follows from (B99). Hence, for $n \geq n^*$, $n-1-m \leq N'$ with probability of at least $1 - \epsilon/2$. Thus, for $n \geq n^*$,

$$\begin{aligned} &\overline{\rho}_i^1(hh) - \overline{\rho}_i^1(ll) = \textstyle\sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\left[\overline{\rho}^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll}) - \overline{\rho}^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\right] \\ &= \textstyle\sum_{m=0}^{n-1} \Pr(m) \Pr\left(x_{hh} = \overline{x}_{hh}(m) \middle| m\right) \\ &= \textstyle\sum_{m=0}^{n-1-N'} \Pr(m) \Pr\left(x_{hh} = \overline{x}_{hh}(m) \middle| m\right) + \sum_{m=n-N'}^{n-1} \Pr(m) \Pr\left(x_{hh} = \overline{x}_{hh}(m) \middle| m\right) \\ &\leq \textstyle\sum_{m=0}^{n-1-N'} \Pr(m) \frac{\epsilon}{2} + \sum_{m=n-N'}^{n-1} \Pr(m) \\ &= \tfrac{\epsilon}{2} \Pr[n-1-m \geq N'] + \Pr[n-1-m \leq N'] \leq \epsilon \end{aligned} \tag{B102}$$

where the second equality follows from the definition of $\overline{x}_{hh}(m)$; the first inequality follows from (B100); and the last inequality follows from (B101).

Similarly, let $\underline{\rho}^1$ solve

$$\underline{\rho}^1 \in \arg \min_{\rho:\mathcal{X}_n \to [0,1]} \sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x) \left[\rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll}) - \rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\right], \quad \text{(B103)}$$

and let $\underline{\rho}_i^1(hh)$ and $\underline{\rho}_i^1(ll)$ be the associated perceived provision probabilities when the provision rule $\underline{\rho}^1$. A solution to (B103) is

$$\underline{\rho}^1(x) = \begin{cases} 1 & \text{if } \frac{x_{hh}}{\alpha^2} < \frac{x_{ll}}{(1-\alpha)^2} \\ 0 & \text{if } \frac{x_{hh}}{\alpha^2} \geq \frac{x_{ll}}{(1-\alpha)^2}, \end{cases} \quad \text{(B104)}$$

which is just reversing of provision rule $\overline{\rho}^1$. Hence, conditional on m, $\underline{\rho}^1(\overline{x}_{hh}(m)+1, x_{hl}, x_{lh}, x_{ll}) = 0$ and $\underline{\rho}^1(\overline{x}_{hh}(m), x_{hl}, x_{lh}, x_{ll}+1) = 1$; and for all other values for x_{hh}, agent i's announcement does not affect the provision probability. It thus immediately follows from out previous calculations that

$$\underline{\rho}_i^1(hh) - \underline{\rho}_i^1(ll) = -\sum_{m=0}^{n-1} \Pr(m) \Pr(x_{hh} = \overline{x}_{hh}(m) | m) \geq -\epsilon. \quad \text{(B105)}$$

It follows from (B102) and (B105) that, for any conceivable provision rule,

$$-\epsilon \leq \underline{\rho}_i^1(hh) - \underline{\rho}_i^1(ll) \leq \rho_i^1(hh) - \rho_i^1(ll) \leq \overline{\rho}_i^1(hh) - \overline{\rho}_i^1(ll) \leq \epsilon.$$

■