An Efficiency Rationale for Bundling of Public Goods*

Hanming Fang[†] Peter Norman[‡]

First Version: October 2002
This Version: April 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. For a binary valuation parametric class of examples, 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. All these result are related to the idea that bundling reduces the variance in the distribution of valuations.

*We thank Ted Bergstrom, Larry Samuelson and Stephen Morris for helpful discussions. We also thank seminar participants at University of Essex and University of Wisconsin 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

Many excludable public goods are provided in bundles. An obvious example is cable TV. Technologically, the local cable company could easily allow customers to choose whatever channels they are willing to pay for without constraints. However, with the exception of some premium channels and some pay-per-view programs, the basic pricing scheme usually consists of a limited number of available packages. 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 often 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, as long as he or she pays the property taxes, 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 the above observations, this paper studies the role of bundling in the efficient provision of (excludable) 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 show that the social surplus maximizing mechanism will always be characterized by some degree of bundling. Thus, while "pure bundling" is not necessarily the optimal solution, there is a straightforward argument in favor of joint provision of multiple public goods.

We consider an environment with two excludable public goods and a numeraire private good.[1] Each consumer in the economy is characterized by a valuation for each of the public goods. The valuation for consuming both public goods is assumed to be the sum of the valuations for the individual goods. This assumption rules out bundling arising from complementarities in the utility function. Similarly, the cost of provision of each good is independent of whether or not the other good is provided. That is, there are no synergies on the cost side.

These separability assumptions on valuations and costs imply that the informationally unconstrained efficient provision rule is as follows: a public good should be provided if and only if its

[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.

sum of valuations exceeds its provision cost. Thus, under perfect information, there is no role for bundling in provision; moreover, no agent will be excluded from using a public good if it is provided.

In this paper we depart from the perfect information assumption and assume that preferences are private information to the individuals. The provision mechanism must therefore be constructed so that truthful revelation of preferences is consistent with equilibrium. Moreover, we allow agents to freely choose whether to participate in the mechanism, and require the provision mechanism to 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 is essentially the only instrument that can be used to make consumers willing to contribute a non-negligible amount to the public goods. In the case of a single public good the constrained optimal mechanism is well approximated by standard third-degree price discrimination: for each agent the designer sets a fixed user fee and the agent is included if and only if she is willing to pay the fee. If all agents are ex ante identical, all agents face the same user fee and the mechanism reduces to average cost pricing.

This characterization of the provision problem for a single good generates a simple intuition for the usefulness of bundling. For simplicity, assume that goods are symmetric and that valuations are drawn from the same distribution for all agents. The best mechanism that does not use the bundling option is approximately a fixed user fee, and the fixed user fee is the same for both goods.

The crucial observation is that the average valuation for the "bundled good" is less dispersed than the distribution of valuations for each good. If the user fee is below the expected valuation, this suggests that fewer consumers are excluded from usage if the goods are sold only as a bundle at a price given by the sum of the user fees. Counter examples to both steps of this reasoning are easy to generate, but for log-concave symmetric distributions with non negative support, we show that bundling leads to fewer exclusions.

However, there is an equally intuitive downside to bundling. While fewer agents are excluded, the reduced dispersion of valuations also implies that the average valuation conditional on being excluded is higher under bundling than under separate provision. Thus whether or not *pure* bundling

[2]All these restrictions are essential to the analysis. Removing either the voluntary participation or the self-financing constraint makes it possible to construct pivot-mechanisms that implement the first-best. If we remove the independence assumption and allow correlation in valuations, an adaption of the analysis in Cremer and McLean [5] can be used to implement the efficient outcome.

dominates separate provision depends on parameters.

It is interesting to note that the welfare comparison between pure bundling and separate provision is non-monotonic in the cost of provision. When the cost is sufficiently low, the welfare gain from fewer exclusions in the bundled mechanism necessarily dominates, so bundling is superior to separate provision mechanisms. On the other hand, when the cost is high enough, it is possible that the public goods can be provided if the goods are bundled, but not under the best mechanism with separate provision, again implying that pure bundling dominates separate provision. For intermediate costs, there may be a range where separate provision is better.

Characterizing the best mechanism under either separate provision or pure bundling is relatively straightforward since these problems are both one-dimensional. The full design problem, however, is truly multi-dimensional and this paper does not solve that problem in general. What we do however, is to solve for the optimal provision mechanism in a special case where the valuation for each good is binary. We then find that there is a strong element of bundling in the optimal solution in the sense that types that have a low valuation for both goods always get lower priority than "mixed" types get for their low valuation good.

There is a considerable literature on bundling of private goods. Adams and Yellen [1], McAfee, McMillan and Whinston [7] showed that a multiproduct monopolistic producer may have incentive to bundle commodities to maximize revenue; Nalebuff [10] showed that in oligopolistic setting, commodity bundling may be used as an entry deterrent. While there are many similarities at the technical level, there are significant qualitative differences between the private and public goods cases. In particular, any bundling in the private goods case is necessarily inefficient, albeit revenue enhancing: a profit maximizing seller with monopoly power may want to bundle two commodities, but marginal cost pricing is always more efficient.[3]

The remainder of the paper is structured as follows. Section 2 presents the general model; Section 3 compares at a general level the optimal separate provision mechanisms and the pure bundle mechanism; Section 4 presents a special binary valuation example of our general model; Section 5 characterizes the optimal mechanism for our binary valuations example and compares our characterization with existing results in the literature; finally, Section 6 concludes. Most technical proofs are collected in the appendix.

[3]Unless there are significant fixed costs, but these fixed costs can then be viewed as a public good.

2 The Model

The objective of the paper is to understand whether there is an efficiency rationale for bundling in the provision of collective goods. To do this, we imagine a fictitious social planner who seeks maximize social surplus in an environment with multiple excludable public goods. There are two important reasons for allowing use exclusions. Firstly, it allows use to consider large economies without making the provision problem "impossible". Secondly, it allows for a more intuitive form of bundling since different consumers can consume different bundles when exclusions are possible.

The environment is set up so ex post efficiency requires that a good is provided if and only if the sum of valuations for that good exceeds the cost of provision for that good. Any role for bundling therefore comes from the constraints facing the planner. The constraints we impose are all standard. Preferences are assumed to be private information, so any implementable mechanism must be incentive compatible. Moreover, the mechanism must be self-financing, and agents are free to opt out of the mechanism.

All these constraints are noncontroversial if thinking about the design problem as characterizing the efficiency frontier of what may be achieved in a private market bargaining agreement, but in particular the voluntary participation constraints may seem questionable if interpreting the goods as government provided. 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 then should be endogenous).

There are two excludable public goods, labeled by $j = 1, 2$; and n agents, indexed by $i \in \mathcal{I} = \{1, ..., n\}$. In order to obtain a tractable characterization we will focus on asymptotic results as the number of agents n goes to infinity. Provision of either public good is a binary decision and the cost of providing good j is denoted $C^j(n)$. Note that the assumption that the cost of provision is independent of the number of *users* implies that the goods are fully non-rival. The rationale for indexing costs by the size of the economy is to avoid to make the problem trivial in a large economy. In other words, we will assume that there exists $c^j > 0$ such that $\lim_{n\to\infty} C^j(n)/n = c^j > 0, j = 1, 2$. The most useful way to think about this assumption is as a normalization of the per capita costs of provision.

Agent i's valuation from good j is denoted by θ_i^j. Valuations are independently and identically distributed across agents.[4] Agent i's valuations of public goods (θ_i^1, θ_i^2) are drawn from a joint

[4]More specific distributional assumptions will be made later when solving the model. We will momentarily work with general distributions because the setup is easier to understand and the exposition of some intermediate results

distribution F. The marginal distribution of good j's valuation is denoted by F^j and the conditionals will be denoted by $F^1(\cdot|\theta_i^2)$ and $F^2(\cdot|\theta_i^1)$ respectively. The set of possible realizations of (θ_i^1, θ_i^2) is denoted $\Theta_i = \Theta^1 \times \Theta^2$, where Θ^j is the set of possible values for θ_i^j. For brevity, we write $\theta_i = (\theta_i^1, \theta_i^2) \in \Theta_i$, $\theta = (\theta_1, ..., \theta_n) \in \Pi_{i\in\mathcal{I}}\Theta_i \equiv \Theta$, $\theta_{-i} = (\theta_1, .., \theta_{i-1}, \theta_{i+1}, ..., \theta_n) \in \Pi_{k\in\mathcal{I}\setminus\{i\}}\Theta_k \equiv \Theta_{-i}$, and $\theta^j = \left(\theta_1^j, ..., \theta_n^j\right) \in \left[\Theta^j\right]^n, j = 1, 2$. By independence, F is the prior distribution over agent i's valuations θ_i perceived by the mechanism designer as well as other agents when the revelation game is played. 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.

Write $\mathtt{I}_i^j$ as the dummy variable taking value 1 when agent i consumes good j and 0 otherwise, and t_i as the quantity of private goods (or the transfer of "money") that i sacrifices to consume the goods (if at all). Then the utility for agent i of type (θ_i^1, θ_i^2) is

$$\mathtt{I}_i^1\theta_i^1 + \mathtt{I}_i^2\theta_i^2 - t_i. \tag{1}$$

A number of restrictions are imposed with this formulation. Besides the additive separability between "money" and the public good, it also rules out complementarities between the two public goods, and assumes that agents are risk neutral.

2.1 General Mechanisms

In general, the outcome of any mechanism must specify the following: (1). whether or not public good j should be provided, for $j = 1, 2$; (2). if public good j is provided, which agents should be allowed access to it, for $j = 1, 2$; (3). how the costs of public good provision should be shared among the agents. The set of feasible *pure* outcomes is thus

$$A = \underbrace{\{0,1\} \times \{0,1\}}_{\substack{\text{provision/no provision}\\ \text{for goods 1 and 2}}} \times \underbrace{\{0,1\}^n \times \{0,1\}^n}_{\substack{\text{inclusion/no inclusion}\\ \text{for good 1 and 2 for all agents}}} \times \underbrace{\mathrm{R}^n}_{\text{"taxes"}}. \tag{2}$$

By the revelation principle, we will, without loss of generality, restrict our attention to direct mechanisms for which truth-telling is a Bayesian Nash equilibrium. A pure direct mechanism is simply a map from Θ to A. Following Aumann [2], we represent a randomized mechanism as a measurable mapping $g : \Theta \times X \to A$, where $X = [0, 1]$; and $x \in X$ is the outcome of a fictitious lottery. Without loss of generality, x is assumed uniformly distributed and independent of θ.[5]

is clearer.

[5]Because A is finite, there is no technical reason to deviate from the "natural" representation of a randomized mechanism as a map from Θ to the set of probability distributions over A. The representation is chosen only because

A conceptual advantage of representing random mechanisms in analogy with Aumanns' way of representing mixed strategies is that we can decompose any mechanism in a useful way. We denote a mechanism by $\mathcal{M}$ and observe that we may think of a mechanism as a 5−tuple $\mathcal{M} = (\zeta^1, \zeta^2, \omega^1, \omega^2, \tau)$ where,

$$\begin{aligned} \zeta^j &: \Theta \times X \to \{0,1\} \text{ for } j = 1,2 \\ \omega^j &: \Theta \times X \to \{0,1\}^n \text{ for } j = 1,2 \\ \tau &: \Theta \to \mathrm{R}^n. \end{aligned} \tag{3}$$

We refer to ζ^j as the *provision rule* for good j, and interpret $\mathrm{E}_X \zeta^j(\theta, x)$ 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 rules* for good j, and $\mathrm{E}_X \omega_i^j(\theta, x)$ 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 $\tau_i(\theta)$, if positive, is the transfer from agent i to the social planner when announcements are θ. In general, transfers could of course also be randomized, but since agents are risk neutral with respect to transfers, there are no gains from randomizing over transfers. The pure transfer rule in (3) is therefore without loss of generality.

Denote by E_{-i} the expectation operator with respect to (θ_{-i}, x) conditional on i's valuation θ_i. Agent i's conditional expected payoff when the announcements are $\widehat{\theta}$ given her true valuations $\theta_i = (\theta_i^1, \theta_i^2)$ is

$$\mathrm{E}_{-i}\left[\sum_{j=1,2} \zeta^j(\widehat{\theta}, x)\omega_i^j(\widehat{\theta}, x)\theta_i^j - \tau_i(\widehat{\theta})\right] \qquad \forall i \in \mathcal{I}, \theta_i \in \Theta_i. \tag{4}$$

Incentive compatibility, that is, the requirement that truth-telling is a Bayesian Nash equilibrium in the revelation game induced by $\mathcal{M}$, requires

$$\begin{aligned} \mathrm{E}_{-i}\left[\sum_{j=1,2} \zeta^j(\theta, x)\omega_i^j(\theta, x)\theta_i^j - \tau_i(\theta)\right] &\geq \mathrm{E}_{-i}\left[\sum_{j=1,2} \zeta^j(\widehat{\theta}_i, \theta_{-i}, x)\omega_i^j(\widehat{\theta}_i, \theta_{-i}, x)\theta_i^j - \tau_i(\widehat{\theta}_i, \theta_{-i})\right] \\ \forall i &\in \mathcal{I}, \theta \in \Theta, \widehat{\theta}_i \in \Theta_i. \end{aligned} \tag{5}$$

Balanced-budget constraint requires that the taxes collected should be sufficient to finance the

it generates more convenient notation than either the "natural" representation or a representation following the "distributional approach" of Milgrom and Weber [8]).

provision cost. We impose the *ex ante balanced-budget constraint:*[6]

$$\mathrm{E}\left(\sum_{i}\tau_{i}\left(\theta\right)-\sum_{j=1,2}\zeta^{j}\left(\theta,x\right)C^{j}\left(n\right)\right)\geq 0. \tag{6}$$

Finally, we impose *voluntary participation* or *individual rationality* constraint. When deciding whether to participate in the mechanism, agents know their own valuations, but do not know the realizations of other agents' valuations and the outcome of the fictitious lottery x. That is, individual rationality is imposed at the interim stage as

$$\mathrm{E}_{-i}\left[\sum_{j=1,2}\zeta^{j}(\theta,x)\omega_{i}^{j}(\theta,x)\theta_{i}^{j}-\tau_{i}(\theta)\right]\geq 0,\qquad\forall i\in I,\theta_{i}\in\Theta_{i}. \tag{7}$$

A mechanism is *incentive feasible* if it satisfies (5), (6) and (7). A mechanism is *constrained efficient* if it is incentive feasible and maximizes the expected social surplus

$$\sum_{j=1,2}\mathrm{E}\zeta^{j}(\theta,x)\left[\sum_{i\in\mathcal{I}}\omega_{i}^{j}(\theta,x)\theta_{i}^{j}-C^{j}\left(n\right)\right]. \tag{8}$$

As a comparison benchmark, the perfect information ex post efficient rule is as follows: public good j is provided if and only if $\sum_{i\in\mathcal{I}}\theta_{i}^{j}\geq C^{j}\left(n\right)$, and if it is provided, no agents should be excluded from its usage. It is well understood that, with more than two agents, such rule is implementable only in trivial cases: either no public good should be provided in all situations, or the public good can be financed by charging the lowest possible valuation. In any other cases we can apply Mailath and Postlewaite's [6] adaptation of the fundamental bargaining inefficiency result from Myerson and Satterthwaite [9] to conclude that ex post efficiency is impossible to achieve.

2.2 Simple Anonymous Mechanisms

To characterize the constrained efficient mechanism, we need to maximize (8) subject to constraints (5), (6) and (7). We first exploit the facts that all control variables enter linearly in both the constraints and the objective function and that the problem is symmetric to reduce the dimensionality of the problem.

[6] The ex ante balanced-budget constraint is literally relevant only when the designer can access fair insurance market against budget deficits. However, standard arguments (see Mailath and Postlewaite [6] and Cramton et al [4]) show that, for any allocation implementable with transfers satisfying the ex ante balanced-budget constraint, there exists a transfer rule that satisfies the ex post balanced-budget constraint (i.e. resource feasibility is guaranteed for every realization of θ) that implements the same allocation.

Consider a simplified class of mechanisms with a generic element $M = (\rho^1, \rho^2, \eta^1, \eta^2, t)$ where

$$\begin{aligned} \rho^j &: \Theta \to [0,1] \text{ for } j = 1,2, \\ \eta^j &: \Theta_i \to [0,1] \text{ for } j = 1,2, \\ t &: \Theta_i \to \mathrm{R}. \end{aligned} \tag{9}$$

Here ρ^j is the provision rule, η^j is the inclusion rule and t is the transfer rule. Relative to (3), a number of restrictions are built into the specification in (9): first, the inclusion and transfer rules are the same for all agents; second, conditional on the realization of θ, the provision probabilities $\rho^1(\theta), \rho^2(\theta)$ are stochastically independent of the inclusion probabilities $\{\eta^1(\theta_i), \eta^2(\theta_i)\}_{i=1,\dots n}$; and third, the inclusion and transfer rules for any agent i are independent of the realization of θ_{-i}. A mechanism as specified in (9) is called a *simple* mechanism. Furthermore, a simple mechanism is *anonymous* if $\rho^j(\theta) = \rho^j(\theta')$ for $j = 1,2$ and every $\theta, \theta' \in \Theta$ where θ' is a permutation of θ. The index of an agent is completely irrelevant in anonymous simple mechanisms.

We now use the symmetry of the problem and the linearity in the payoffs to show that it suffices for the social planner to focus on simple anonymous mechanisms.

Proposition 1 *For any incentive feasible mechanism $\mathcal{M}$ of the form (3), there exists an anonymous simple incentive feasible mechanism M of the form (9) that generates the same value of the planner's objective function.*

Consequently, we consider in the remainder of this paper only simple anonymous mechanisms of the form (9). The proof of Proposition 1 uses tedious but straightforward arguments and it is relegated to the appendix. The intuition for why we can, with no loss of generality, only consider the simple anonymous mechanisms is simple. First, because of the risk neutrality, all agents care only about their expected probability of consuming each public good and the expected transfer; thus nothing is gained by making transfers and inclusion probabilities functions of θ_{-i} or by making inclusion and provision rules conditionally dependent. Second, all agents may be treated symmetrically (i.e., η^j and t is the same for all agents) since from any asymmetric mechanism, we can create new asymmetric mechanisms that generates the same surplus by permuting the roles of the agents. But then, if we randomize over all $n!$ permuted asymmetric mechanisms, we obtain an incentive feasible symmetric mechanism that generates the same surplus.[7]

[7] The exact argument has to be slightly modified because the inclusion probabilities and provision probabilities are potentially correlated since they both depend on $\theta_i \in \Theta_i$.

It can also be shown that, if the provision costs $C^j(n)$ and conditional valuation distributions F^j of the two public goods are identical, then the two goods may also be treated symmetrically:

Proposition 2 *Suppose that* $\Theta^1 = \Theta^2, F^1(\cdot|\theta^2) = F^2(\cdot|\theta^1)$ *for any* $\theta^1, \theta^2 \in \Theta^1 = \Theta^2$ *and* $C^1(n) = C^2(n)$. *Then, given any simple anonymous incentive feasible mechanism* $\widetilde{M}$, *there exists an incentive feasible mechanism* M *that generates the same social surplus with the property that* $\rho^1(\theta^1, \theta^2) = \rho^2(\theta^2, \theta^1)$ *for all* $\theta \in \Theta$ *and* $\eta^1(\theta_i^1, \theta_i^2) = \eta^2(\theta_i^2, \theta_i^1)$ *for all* $\theta_i \in \Theta_i$.

The idea of the proof is similar to that of Proposition 1. From any mechanism, one can reverse the roles of the goods and construct another mechanism that generates the same surplus. A symmetric mechanism can then be constructed by averaging the initial and the reversed mechanism.

3 Pure Bundle Versus Separate Provision Mechanisms

In this section, we compare two extreme mechanisms: in the pure bundle mechanism, agents can access either both or none of the public goods; and in the separate provision mechanism, the provision of the two goods are separately considered. This comparison provides useful intuition for the optimal mechanism in the binary valuation example we will study in Section 4. Note that, under both separate provision and pure bundle mechanisms, the designer faces a single-dimensional problem; this allows us to compare these two extremes by appealing to known results about provision of a single excludable public good.

3.1 Constrained Optimal Separate Provision Mechanism

In a separate provision mechanism, the mechanism designer deals with the mechanism design problem for one public good as if the other good did not exist. It is best to imagine two separate mechanism designers respectively in charge of the provision of the two goods, with the common knowledge that the two designers do not share information obtained from the agents. Specifically, a separate provision mechanism for good j is a mechanism where: (1). the provision and inclusion probability for j depends only on valuation announcements for j; (2). the financing of j is separate from that of the other good.

Since Proposition 1 holds also for a single public good, we only need to focus on simple anonymous mechanisms. The optimal separate provision mechanism is a special case of the general characterization in Norman [11] (see his Proposition 2 and 3). The asymptotic result as $n \to \infty$ is as follows:

Let F^j denote the marginal distribution over θ_i^j and assume that the virtual valuation $\theta_i^j - \left[1 - F^j(\theta_i^j)\right] / f^j(\theta_i^j)$ is strictly increasing. Let

$$p_m^j = \arg\max_{p \in \Theta^j} p\left[1 - F^j(p)\right]. \tag{10}$$

Then: (1). if $p_m^j \left[1 - F^j(p_m^j)\right] > \lim_{n\to\infty} C^j(n)/n$, then the ex ante probability of provision under the optimal mechanism converges to 1 as $n \to \infty$; (2). if $p_m^j \left[1 - F^j(p_m^j)\right] < \lim_{n\to\infty} C^j(n)/n$, then the ex ante probability of provision in any feasible mechanism converges to zero as $n \to \infty$.[8] The term p_m^j has a natural interpretation as the price that a profit maximizing monopolistic provider would charge for access if the good is provided for sure. The basic intuition for the result is as follows: as $n \to \infty$, the average probability of being pivotal goes to zero ; thus it becomes impossible to price discriminate between agents with the same inclusion probability. Under the assumption that virtual valuations are increasing, the best inclusion rule is a threshold rule where agents are included if an only if their valuations are above the threshold. Hence, all agents above the threshold must be charged with the same price. Norman [11] also shows (in his Proposition 4 and Lemma 3) that average cost pricing is asymptotically optimal whenever the provision probability under the optimal mechanism converges to unity, where the notion of "asymptotic optimality" is that the difference in per capita surplus between the best average cost pricing mechanism and the optimal mechanism can be made arbitrarily small as n is sufficiently large. Average cost pricing in this context is simply to find the smallest p^{*j} such that $C^j(n)/n = p^{*j}\left[1 - F^j(p^{*j})\right]$.

3.2 Example 1: Pure Bundle versus Separate Provision Mechanisms with Independent Uniform Valuation Distributions

We now provide an example to illustrate the basic ideas. Assume that θ_i^1 and θ_i^2 are both uniformly distributed on $[0, 1]$ and stochastically independent; and $C^j(n) = cn$ for $j = 1, 2$ where $c < 1/4$.

Under these parametric assumptions, the "monopoly price" defined in (10) is $p_m^j = 1/2$; and $p_m^j \left[1 - F^j(p_m^j)\right] = 1/4 > c$. Thus, Norman [11]'s results discussed in Section 3.1 imply that, under the separate provision mechanisms, it is asymptotically optimal to provide both goods with

[8] More generally, when the valuation of different agents are drawn from different distributions (as in Norman [11]), the monopoly prices will be individual-specific.

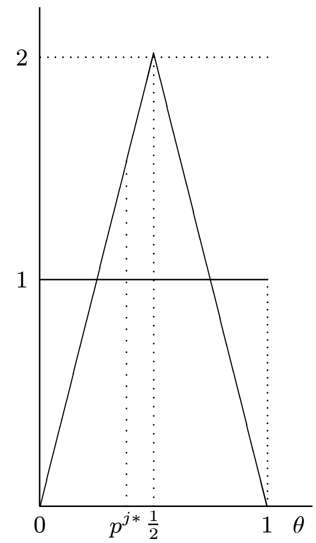

Figure 1: Bundling vs. Non-Bundling with Uniform Valuations.

probability 1 and charge a user fee equal to the average cost:[9]

$$p^{j*} = \frac{1}{2} - \sqrt{\frac{1}{4} - c}. \tag{11}$$

Note that $p^{j*} < 1/2 = p_m^j$, which is simply that a social planner sets a lower price than a profit maximizing monopolist.

Now suppose that the mechanism designer continues to provide both goods with probability one, but *only* offers the *bundle* at a user fee of $p^{1*} + p^{2*}$.[10] That is, agents have to decide whether they are willing to pay $p^{1*} + p^{2*}$ for *both* goods and do not the option to subscribe to one good only. Write $\theta_i^b = \theta_i^1 + \theta_i^2$ as agent i's the valuation of bundle, and f^b and F^b as its density function and cumulative probability function respectively. Note that θ_i^b has a triangular distribution on support $[0, 2]$, thus f^b is

$$f^b(\theta_i^b) = \begin{cases} \theta_i^b & \text{if } \theta_i^b \leq 1 \\ 2 - \theta_i^b & \text{if } \theta_i^b > 1. \end{cases}$$

Figure 1 plots p.d.f of the *average* valuation, $\theta_i^b/2$, and the single good valuation θ_i^j. Under both the separate provision mechanism and the proposed bundle mechanism, the utility loss compared

[9]That is, p^{*j} is the smallest solution to $c = p^j \left[1 - F^j(p^j)\right] = p^j(1 - p^j)$.

[10]This particular bundled mechanism is suboptimal, but allows a more straighforward comparisson with the best mechanism with separate provision than the best pure bundling mechanism.

to the first best is the surplus loss due to exclusions. Under the separate provision mechanism, the measure of agents excluded from good j is represented by the rectangular area between 0 and p^{j*} below the uniform density, namely $F^j\left(p^{*j}\right)$; and under the proposed bundle mechanism, it is given by the triangular area between 0 and p^{j*} below the density of $\theta_i^b/2$, namely $F^b\left(2p^{j*}\right)$.

Since $p^{j*} < 1/2$, we have that $F^b(2p^{j*}) = 2\left[p^{j*}\right]^2 < p^{j*}$. Thus an immediate result is that there are always fewer exclusions for either good under the proposed bundle mechanism. Equally transparent from Figure 1 is that the expected (per-good) valuation *conditional on exclusion* is higher in the bundling mechanism.[11] This results in a trade-off where the net gain or loss from bundling depends on p^{j*}, which in turn is determined by the cost c. The total expected surplus loss due to exclusions may be calculated as

$$2\int_0^{p^{j*}} \theta_i^j d\theta_i^j = \left[p^{j*}\right]^2$$

under the separate provision mechanism and

$$\int_0^{2p^{j*}} (\theta_i^b)^2 d\theta_i^b = \frac{8\left[p^{j*}\right]^3}{3}$$

under the proposed bundle mechanism. Therefore, the total expected surplus loss due to exclusion is smaller under the proposed bundle mechanism if and only if $p^{j*} < 3/8$, which, from Eq. (11), would be the case if $c < 15/64$. Thus, if $c < 15/64$, the proposed bundle mechanism dominates the separate provision mechanisms from the social planner's view point; if $15/64 < c < 1/4$, then the proposed bundle mechanism is dominated.

As already mentioned, the proposed bundle mechanism is not the best mechanism with bundling. The reason is that, since there is fewer exclusions when bundling at price $2p^{j*}$, there is a budget surplus. The mechanism can therefore be improved upon by lowering the access charge. However, as c approaches $1/4$ this positive effect becomes negligible, whereas the loss arising from the relative increase in the expected valuation conditional on being excluded increases as c increases. Hence, there is a range of costs where the optimal separate provision mechanism dominates the optimal pure bundle mechanism.

[11] Clearly, the expected valuation conditional on being excluded is $\frac{p^{j*}}{2}$ in the case with separate provision. In the case with bundling we calculate it as

$$\frac{\int_0^{2p^{j*}} (\theta_i^b)^2 d\theta_i^b}{\int_0^{2p^{j*}} \theta_i^b d\theta_i^b} = \frac{4}{3}p^{j*}.$$

Hence, the expected valuation for *either good* conditional on being excluded is $\frac{2}{3}p^{j*}$.

3.3 Symmetric Log-Concave Densities

While welfare comparisons are ambiguous in general, the example in Section **??** suggests that bundling tends to result in fewer exclusions than separate provision and thus increase mechanism designer's revenue (which in our model has welfare consequences due to the break-even constraint). Not surprisingly, this is not true for arbitrary distributions. However, in this section we will demonstrate that the logic from the uniform example can at least be extended to the class of symmetric log-concave densities with non-negative support.

We first show that a profit maximizing monopolist will charge a price at which more than half of its customers are included in expectation.

Lemma 1 *Suppose that f is log-concave and symmetric with support $[\underline{\theta}, \overline{\theta}]$, where $\underline{\theta} \geq 0$. Let $\widetilde{\theta} = \left(\underline{\theta} + \overline{\theta}\right)/2 = \mathrm{E}\theta_i^j$, and $p_m = \arg\max_p p[1 - F(p)]$. Then $p_m \leq \widetilde{\theta}$. Moreover, if either $f(\widetilde{\theta}) > 1/\left(\overline{\theta} - \underline{\theta}\right)$ or $\underline{\theta} > 0$, then $p_m < \widetilde{\theta}$.*

Notice that the only case covered in Lemma 1 where $p_m = \widetilde{\theta}$ is when f is uniform with support $\left[0, \overline{\theta}\right]$.

Next, we state a result about convolutions of log-concave random variables due to Proschan [12]. Roughly speaking, the result is that if a random variable is generated by taking an average of two draws from a symmetric log-concave distribution, then the distribution of the average is more "peaked" than the underlying distribution of the two draws in the sense that the probability of a given size deviation from the mean is smaller for the average than for the underlying distribution.[12]

Lemma 2 ***(Proschan [12])*** *Let f be a symmetric log-concave density and let x_1 and x_2 be independently distributed with density f. Denote the support by $[\underline{x}, \overline{x}]$ and the mode by $\widetilde{x} = \left(\underline{x} + \overline{x}\right)/2$. Furthermore, let F_A denote the cumulative of $\left(x_1 + x_2\right)/2$. Then, for any $\underline{x} < t < \widetilde{x}$ we have that $F_A\left(t\right) < F\left(t\right)$.*

Now suppose that $\theta_i^j, j \in \{1, 2\}$ has a symmetric log-concave density with support $[\underline{\theta}, \overline{\theta}]$. We can use Proschan's theorem to obtain:

Proposition 3 *Suppose f^j is symmetric, log-concave, has a non-negative support. and independent from each other. Then*

1. *Bundling at price $2p_m$ gives at least as high revenues as selling each good at price p_m.*

[12] We are grateful to Rustam Ibragimov for the reference of Proschan's results.

2. *If $p_m < \widetilde{\theta}$, then there is a strict gain in revenue from bundling.*

Proof. The expected revenue from bundling at price $2p_m$ is

$$2p_m \Pr\left[\theta_i^1 + \theta_i^2 \geq 2p_m\right] = 2p_m \Pr\left[\frac{\theta_i^1 + \theta_i^2}{2} \geq p_m\right] = 2p_m \left[1 - F_A(p_m)\right].$$

The result is then immediate from Lemma 2 and Lemma 1. ■

Proposition 3 is on revenue maximization for a monopolistic provider. But, Norman [11] show that whether an excludable public good can be provided asymptotically (as $n \to \infty$) in the optimal mechanism is completely determined by whether or not a profit maximizing monopolist can break even. Thus, if we assume that $C^j(n) = cn$ for $j \in \{1, 2\}$, and let $p_m^A = \arg\max p\left[1 - F_A(p)\right]$, then an immediate corollary of Proposition 3 is as follows:[13]

Corollary 1 *Suppose that $p_m < \widetilde{\theta}$. Then, for any c satisfying $p_m\left[1 - F(p_m)\right] < c < p_m^A\left[1 - F_A(p_m^A)\right]$,*

1. *the ex ante probability of provision converges to zero for both goods in any separate provision mechanism;*

2. *the ex ante probability of provision converges to one in the optimal pure bundle mechanism.*

To give a concrete example, note that for the uniform $[0, 1]$ example in the previous section, if $c > 1/4$, we know from Norman [11] that the probability of provision converges to zero in the optimal separate provision mechanism. However, if the two public goods are provided in a bundle, then $p_m^A = 1/\sqrt{6}$ and $p_m^A\left[1 - F_A(p_m^A)\right] = 2/\left(3\sqrt{6}\right) \approx 0.272$.[14] Thus for any $c \in \left(1/4, 2/\left(3\sqrt{6}\right)\right)$, the public goods can be provided with probability one in the optimal pure bundle mechanism.

To summarize our discussion in this section, pure bundle mechanism may generate higher expected social surplus than the optimal separate provision mechanisms. First of all, bundle mechanisms exclude fewer consumers under some technical conditions (symmetric log-concavity of the density function is a sufficient condition). Sometimes fewer exclusion arises together with a lower expectation of the valuation conditional on exclusion, such as in the uniform example when c is sufficiently low. In these situations, the proposed (albeit suboptimal) bundle mechanism clearly dominates the optimal separate provision mechanisms. In other situations, however, the expectation of the valuation conditional on exclusion may be higher under the proposed bundle mechanism.

[13] Note, by definition, $p_m^A\left[1 - F_A\left(p_m^A\right)\right] \geq p_m\left[1 - F_A(p_m)\right] > p_m\left[1 - F(p_m)\right]$.

[14] For $p < 1/2$, $F_A(p) = F^b(2p) = 2p^2$. Hence the solution to $\max p\left[1 - F_A(p)\right] = p\left(1 - 2p^2\right)$ is given by $p_m^A = 1/\sqrt{6} \approx 0.408$, and the maximum value is $2/\left(3\sqrt{6}\right) \approx 0.272$.

Then the welfare comparison of the bundle and separate provision mechanisms depend on the trade-off between probability of exclusion and the conditional valuation given exclusion. Intuitively, the higher is the provision cost, the closer the average cost price is to the mean of the distribution, and the more likely that the surplus loss per excluded agent outweighs the gains from fewer exclusion. For this reason, bundling is advantageous for relatively low costs. It is also important to point out that, fewer exclusion (at the sum of separate provision prices) under the bundle mechanism implies a higher revenue, thus it is possible that public goods that could not be provided under the separate provision mechanisms would be provided with probability one using a bundle mechanism. This means that pure bundle mechanism also improves welfare when costs are moderately high. Therefore, the capability of pure bundle mechanism to improve welfare can be non-monotonic in the provision cost c.

The basic intuition for the result that pure bundle mechanism may dominate the optimal separate provision mechanism is that the average valuation of two independent random variables is less dispersed around its mean than the individual random variable. We qualified the above intuition with some technical conditions of symmetric log-concavity.

We close the discussion in this section with an example to demonstrate the necessity of such technical conditions. Suppose that, with probability $\alpha \in (0,1)$, θ_i^j is uniform $[0,2]$ and with probability $1-\alpha$, θ_i^j takes value 1 for sure. This distribution satisfies all conditions in Proposition 3 except for log-concavity. The CDF of θ_i^j is

$$F(\theta_i^j) = \begin{cases} \frac{\alpha}{2}\theta_i^j & \text{for } \theta_i^j < 1 \\ (1-\alpha) + \frac{\alpha}{2}\theta_i^j & \theta_i^j \geq 1 \end{cases}.$$

The CDF of $\left(\theta_i^1 + \theta_i^2\right)/2$ is discontinuous at 1, and the probability that it takes exact value of 1 (its mean) is $(1-\alpha)^2$, smaller than the probability that θ_i^j takes on value 1 which is $1-\alpha$. Write F^A as the CDF of the average. Since the jump at the discontinuity is smaller for the average and the density is symmetric we conclude that there exists $\epsilon > 0$ such that $F^A(y) > F(y)$ for any $y \in (1-\epsilon, 1)$. Thus it is possible that, if the two goods are bundled at a price in this interval, there would be more exclusions with bundling than under separate provision. Also note that since there is a range of values for y that work we may smooth the example with a continuous CDF.

3.4 Example 2: Bundling of a Large Number of Independent Goods

In the previous subsection, we presented a sufficient technical condition for the bundle mechanism to dominate the optimal separate provision mechanisms. An alternative approach is to consider a large number of public goods. Intuitively, when there are a large number of goods with independent valuations, by the law of large number, the average valuation can be made arbitrarily close to the mean and thus less dispersed around the mean than the valuation distribution of any signal good.

Suppose there are an infinite sequence of "potential public goods" indexed by $j = 1, 2, ...$ Let θ_i^j denote agent i's valuation for good j, and assume that for each j, θ_i^j is identically and independently distributed across agents and for each i, θ_i^j is stochastically independent across goods. Finally, assume that there exist 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.[15] For this subsection, we will not impose symmetry in either costs or valuation distributions, thus Proposition 2 will not be used.

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? We consider two cases of public goods. First, if $\mathrm{E}\theta_i^j > \lim_{n\to\infty} C^j(n)/n = c^j$, then there exists N such that $C^j(n) \leq nc^j + n(\mathrm{E}\theta_i^j - c^j)/2$ for every $n \geq N$. Applying Chebyshev's inequality, we have

$$\begin{aligned}
\Pr\left[\sum_{i=1}^{n} \theta_i^j \leq C^j(n)\right] &\leq \Pr\left[\sum_{i=1}^{n} \theta_i^j \leq \frac{n(\mathrm{E}\theta_i^j + c^j)}{2}\right] = \Pr\left[\sum_{i=1}^{n} \theta_i^j - n\mathrm{E}\theta_i^j \leq -\frac{n(\mathrm{E}\theta_i^j - c^j)}{2}\right] \qquad (12) \\
&\leq \Pr\left[\left|\sum_{i=1}^{n} \theta_i^j - n\mathrm{E}\theta_i^j\right| \geq \frac{n(\mathrm{E}\theta_i^j - c^j)}{2}\right] = \frac{4\mathrm{Var}\left(\sum_i \theta_i^j\right)}{n^2(\mathrm{E}\theta_i^j - c^j)^2} \leq \frac{4\sigma^2}{n^2(\mathrm{E}\theta_i^j - c^j)^2}.
\end{aligned}$$

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 $1 - \varepsilon$. Second, if $\mathrm{E}\theta_i^j < c^j$, then a similar argument establishes that the first best provision probability converges to zero as the number of agents goes to infinity. Since only the first case is interesting, so we will now assume that there exists $\delta > 0$ such that $\mathrm{E}\theta_i^j - c^j \geq \delta$ for all j for the remainder of this section.[16]

We now show that when the number of public goods is sufficiently large, bundle mechanisms can approximate the ex post efficient rule. Let m be the number of public goods and we will study what happens when $m \to \infty$. Let $\epsilon > 0$ and consider an anonymous mechanism where all m public goods are provided for sure ($\rho^j(\theta) = 1$ for all $\theta \in \Theta$ and all j) and where the inclusion and transfer

[15] This is guaranteed if there exists an interval $[a, b]$ such that $\Theta_i^j \subset [a, b]$ for every j.

[16] If $\mathrm{E}\theta_i^j > c^j$ for some goods, but the inequality is revererserd 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.

rules are given by,

$$\eta^j(\theta) = \begin{cases} 1 & \text{if } \sum_j \theta_i^j \geq \sum_j \mathrm{E}\theta_i^j - \varepsilon m \\ 0 & \text{if } \sum_j \theta_i^j < \sum_j \mathrm{E}\theta_i^j - \varepsilon m \end{cases} \quad j \in \{1, ..., m\} \tag{13}$$

$$t(\theta) = \begin{cases} \sum_j \mathrm{E}\theta_i^j - \varepsilon m & \text{if } \sum_j \theta_i^j \geq \sum_j \mathrm{E}\theta_i^j - \varepsilon m \\ 0 & \text{if } \sum_j \theta_i^j < \sum_j \mathrm{E}\theta_i^j - \varepsilon m \end{cases}$$

In words, whether an agent gets access to a good depends only on her announced valuation for the whole bundle: if her total announced valuations exceeds the threshold $\sum_j \mathrm{E}\theta_i^j - \varepsilon m$, then she gets access to the bundle and pays a user fee of $\sum_j \mathrm{E}\theta_i^j - \varepsilon$; otherwise she is excluded from the bundle and pays nothing. Truth-telling is a dominant strategy for this mechanism and participation constraints are satisfied. Thus we only need to check the budget balance constraint (6).

By the assumption of a uniform upper bound σ^2 for the variances of $\theta_i^1, ...\theta_i^m$, we can use Chebyshev's' inequality to get

$$\Pr\left[\sum_j \theta_i^j - \sum_j \mathrm{E}\theta_i^j < -\varepsilon m\right] \leq \frac{\mathrm{Var}\left(\sum_j \theta_i^j\right)}{\varepsilon^2 m^2} \leq \frac{m\sigma^2}{\varepsilon^2 m^2} = \frac{\sigma^2}{\varepsilon^2 m}. \tag{14}$$

Since the right hand side converges to zero as $m \to \infty$, there exists $\bar{m} < \infty$ for every $\varepsilon > 0$ such that, for every $m \geq \bar{m}$,

$$\Pr\left[\sum_{j=1}^m \theta_i^j - \sum_{j=1}^m \mathrm{E}\theta_i^j < -\varepsilon m\right] \leq \varepsilon. \tag{15}$$

The interpretation of (15) is that the probability that an agent is excluded to consume the bundle in mechanism (13) can be made arbitrarily small as m is sufficiently large. Thus, for every $\varepsilon > 0$, there exists $\bar{m}$ such that, for all $m \geq \bar{m}$, the expected user fee collected in the proposed bundle mechanism, $\mathrm{E}nt(\theta)$, satisfies

$$\mathrm{E}nt(\theta) \geq n(1-\varepsilon)\left(\sum_{j=1}^m \mathrm{E}\theta_i^j - \varepsilon m\right). \tag{16}$$

Since we are considering the case in which there exists N such that for all $n \geq N$, $C^j(n) \leq n(\mathrm{E}\theta_i^j + c^j)/2$ for every j, we know that, for all $m \geq \bar{m}, n \geq N$,

$$\mathrm{E}nt(\theta) - \sum_{j=1}^m C^j(n) \geq n(1-\varepsilon)\left(\sum_{j=1}^m \mathrm{E}\theta_i^j - \varepsilon m\right) - n\sum_{j=1}^m (\mathrm{E}\theta_i^j + c^j)/2. \tag{17}$$

Let $\delta > 0$ be the uniform bound of $\mathrm{E}\theta_i^j - c^j$ ($\mathrm{E}\theta_i^j - c^j \geq \delta$ for all j), and μ be the uniform bound of $\mathrm{E}\theta_i^j$ ($\mathrm{E}\theta_i^j \leq \mu$ for all j), we have

$$\begin{aligned} & n(1-\varepsilon)\left(\sum_{j=1}^{m} \mathrm{E}\theta_i^j - \varepsilon m\right) - n\sum_{j=1}^{m}(\mathrm{E}\theta_i^j + c^j)/2 \\ = \; & \frac{n\left[\sum_{j=1}^{m}(\mathrm{E}\theta_i^j - c^j)\right]}{2} - n\varepsilon\left[\sum_{j=1}^{m}\mathrm{E}\theta_i^j + (1-\varepsilon)m\right] \\ \geq \; & nm\left\{\frac{\delta}{2} - \varepsilon\left[\mu + (1-\varepsilon)\right]\right\} \end{aligned}$$

Thus if we pick $\epsilon < \delta/\left[2(1+\mu)\right]$, inequality (17) implies that when m and n are sufficiently large, the proposed bundle mechanism satisfies budget balance constraint. Moreover, as inequality (15) shows, the probability that an agent is excluded can be made arbitrarily small (i.e. approaching the ex post efficient rule of "never exclude") when m is sufficiently large. We thus conclude that the proposed bundle mechanism can approximate arbitrarily well the outcome of the first best efficient mechanism as m and n are both sufficiently large.

However, under the optimal separate provision mechanism, the problem collapses to a special case of a model considered in Norman [11], and it is known from his results that the probability of exclusion is bounded away from zero. The example thus illustrates that bundling the goods together may improve economic efficiency.

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, the valuations of individual goods become irrelevant and the mechanism designer only need to provide consumers to reveal the average valuation of the bundle. But, the distribution over the average valuation collapses into a mass point as the number of goods in the bundle is taken to infinity, so the informational problem essentially disappears. Unlike the case with private goods, efficiency is not compromised when all agents are provided access to the public goods even when their valuations to particular goods may be below the per capita costs.

4 The Model with Binary Valuations

So far, we have demonstrated the welfare improvement potential of pure bundle mechanism over optimal separate provision mechanisms under two sets of conditions: (1). there are finite number of public goods but the valuation distributions are symmetric and log-concave; (2). the number of public goods is large. However, the optimal mechanism with multiple public goods may not involve pure bundling. To characterize the optimal mechanism, we consider a simple discrete version of our general model for tractability.

Suppose that there are two public goods and the valuation of each public good is a binary random variable and independent. For tractability, we assume that the two public goods are symmetric in all respects. More explicitly, the valuation for good j can either be "high" ($\theta_i^j = h$) or "low" ($\theta_i^j = l$). Thus the type space for an individual is $\Theta_i = \{(h,h), (h,l), (l,h), (l,l)\}$. For notational brevity we will henceforth write $\theta_i = hh$ instead of (h,h), $\theta_i = hl$ instead of (h,l), and analogously for other valuations. In the baseline model we also assume that $\alpha = \Pr[\theta_i^1 = h] = \Pr[\theta_i^2 = h] \in (0,1)$, implying that the probability distribution F over Θ_i is:[17]

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

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, which together with the symmetry on the demand side will allow us to restrict attention to symmetric optimal mechanisms. The per capita costs are also kept constant, which does simplify the notation, but would be easy to relax.

An ex post optimal mechanism is to provide good j if and only if $\sum_{i=1}^n \theta_i^j > (\geq) C^j(n) = cn$. If $h \leq c$ "never provide" is thus ex post optimal, which can be trivially implemented. Furthermore, if $l \geq c$ "always provide" is ex post optimal and can be implemented by charging a constant tax equal to c. We therefore maintain the assumption that $l < c < h$ in order to keep the problem interesting.

4.1 Asymptotic Provision Probabilities under the Optimal Separate Provision Mechanisms

In this section, we establish as a benchmark the asymptotic provision probabilities of the two public goods when the provision problem for each public good is considered in isolation. As we argued in Section 3.1, we need only to consider simple anonymous symmetric mechanisms.

To emphasize that the solution depends on the size of the economy we denote a separate provision mechanism for the provision of good j in an economy of size n by $(\rho_n^j, \eta_n^j, t_n^j)$, where

[17] While independence across agents is absolutely crucial for the analysis, independence across goods is not. Everything would go though with rather minor modifications with a probability distribution of the form $\left\{\sigma(hh), \frac{\sigma(m)}{2}, \frac{\sigma(m)}{2}, \sigma(ll)\right\}$, where $\sigma(m)$ is the probability of a "mixed type". That is, as long as the symmetry is kept we can handle positive or negative correlation between the valuation for the goods rather easily.

$\rho_n^j : \{1, ..., n\} \to [0,1]$ and $\rho_n^j(m)$ denotes the probability of provision if m agents announce a high valuation for good j; $\eta_n^j \in [0,1]$ is the inclusion probability for type l and $t_n^j = (t_n^j(h), t_n^j(l))$ are the transfers. In principle it is also possible to exclude agents of type h, but this tightens the downwards incentive constraint for type h and is an option that will never be used, so we immediately build that in to the mechanism to simplify notation. Arguments similar to Propositions 2 and 3 in Norman [11] can be used to get a tight characterization of the asymptotic provision and inclusion rules under the optimal separate provision mechanisms:

Proposition 4 ***(Norman [11])*** *Consider a sequence of economies of size $\{n\}_{n=1}^{\infty}$. Then,*

1. *if $\alpha h < c, \lim_{n\to\infty} \mathrm{E}\rho_n^j(m) = 0$ for any sequence of feasible mechanisms $\left\{\rho_n^j, \eta_n^j, t_n^j\right\}$;*
2. *if $\alpha h > c, \lim_{n\to\infty} \mathrm{E}\rho_n^j(m) = 1$ for any sequence of constrained optimal mechanisms $\left\{\hat{\rho}_n^j, \hat{\eta}_n^j, \hat{t}_n^j\right\}$. Moreover, the sequence $\left\{\rho_n^j, \eta_n^j, t_n^j\right\}$ satisfies:*

$$\begin{aligned}
\lim_{n\to\infty} \eta_n^j &= \frac{\alpha h - c}{\alpha h - l}, \\
\lim_{n\to\infty} t_n^j(l) &= \frac{\alpha h - c}{\alpha h - l} l, \\
\lim_{n\to\infty} t_n^j(h) &= \left[1 - \frac{\alpha h - c}{\alpha h - l}\right] h + \frac{\alpha h - c}{\alpha h - l} l.
\end{aligned}$$

The formal proof is omitted here, instead we provide an heuristic explanation for the result.[18] Since the effect on the provision probability from any individual announcement is negligible in a large economy, the incentive constraint for a type-h agent is roughly that

$$\mathrm{E}\widehat{\rho}_n^j(m)\, h - \widehat{t}_n^j(h) \geq \mathrm{E}\widehat{\rho}_n^j(m)\, \widehat{\eta}_n^j h - \widehat{t}_n^j(l).$$

and the participation constraint for the low type dictates that $\widehat{t}_n^j(l) = \mathrm{E}\widehat{\rho}_n^j(m)\, \widehat{\eta}_n^j l$. Because the incentive constraint for the high type binds in the optimal mechanism, budget balance then requires that, approximately,

$$\begin{aligned}
\mathrm{E}\widehat{\rho}_n^j(m)\, c &= \alpha \widehat{t}_n^j(h) + (1-\alpha)\, \widehat{t}_n^j(l) \approx \alpha \left[\widehat{t}_n^j(l) + \mathrm{E}\widehat{\rho}_n^j(m)\, h \left(1 - \widehat{\eta}_n^j\right)\right] + (1-\alpha)\, \widehat{t}_n^j(l) \\
&= \widehat{t}_n^j(l) + \alpha \mathrm{E}\widehat{\rho}_n^j(m)\, h \left(1 - \widehat{\eta}_n^j\right) \\
&= \mathrm{E}\widehat{\rho}_n^j(m)\, \widehat{\eta}_n^j l + \mathrm{E}\widehat{\rho}_n^j(m)\, \alpha h \left(1 - \widehat{\eta}_n^j\right). \qquad (18)
\end{aligned}$$

[18] Details available on request from the authors.

Hence, $\widehat{\eta}_n^j \approx (\alpha h - c) / (\alpha h - l)$ follows from (18) since in the limit it is valid to ignore the effects from being pivotal. Note that, by inspecting (18), if $\alpha h < c$, (since by assumption, $l < c$ as well), then $\lim_{n\to\infty} \mathrm{E}\widehat{\rho}_n^j (m) = 0$. Otherwise the budget balance constraint must be violated for large n. On the other hand, if instead $\alpha h > c$, it is feasible to provide for sure (for any n) with the transfers specified in Proposition 4, and inclusion probability $\widehat{\eta}_n^j = (\alpha h - c) / (\alpha h - l)$. Conditional on this inclusion probability, the ex post efficient rule is to provide public good j whenever

$$\begin{aligned} & mh + (n-m)\,\widehat{\eta}_n^j l \geq cn, \\ \iff \quad & \frac{m}{n}h + \frac{n-m}{n}\widehat{\eta}_n^j l \geq c. \end{aligned}$$

An application of Chebyshev's inequality guarantees that

$$\operatorname{plim}\left(\frac{m}{n}h + \frac{n-m}{n}\widehat{\eta}_n^j l\right) = \alpha h + (1-\alpha)\,\frac{\alpha h - c}{\alpha h - l} l > \alpha h > c.$$

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

4.2 Example 3: Improvement when Bundling is Allowed

Asymptotically, the optimal separate provision mechanisms characterized in Propositions 4 may not be efficient. First of all, the asymptotic provision probability is zero when $\alpha h < c$ while efficiency requires that the public good be provided whenever $\alpha h + (1-\alpha) l > c$; second, when $\alpha h > c$, there is still inefficiency due to positive probability of exclusion of low valuation agents, even though the public good is provided asymptotically with probability 1. Before we characterize the optimal provision mechanism with bundling, we first provide an example to demonstrate that improvement can be indeed be achieved via bundling by showing a particular incentive compatible, balanced-budget voluntary mechanism (that may not necessarily be optimal) can improve upon the mechanisms without bundling.

Suppose that $\alpha h + (1-\alpha) l > c$. Consider the following provision mechanism with bundling:

- $t_{hh} = t_{hl} = t_{lh} = 2c / \left(2\alpha - \alpha^2\right), t_{ll} = 0$;
- $\eta_{hh} = \eta_{hl} = \eta_{lh} = 1, \eta_{ll} = 0$.
- $\rho^1 (m) = \rho^2 (m) = 1$ for all $m \in \{0, 1, ..., n\}$, i.e. always provide the public goods.

Notice that the (ex ante) budget balance constraint holds for any n by construction of the mechanism since

$$\begin{aligned}&\alpha^2 t_{hh} + (1-\alpha)\,\alpha t_{hl} + \alpha(1-\alpha)t_{lh} + (1-\alpha)^2\, t_{ll}\\ = \;&\left[\alpha^2 + (1-\alpha)\,\alpha + \alpha(1-\alpha)\right]\frac{2c}{2\alpha-\alpha^2} = 2c.\end{aligned}$$

It can also be verified that all the incentive compatibility constraints are satisfied for any n if $h + l - 2c/\left(2\alpha - \alpha^2\right) \geq 0$.

Now, suppose that the valuations of each public good satisfies that $l < c$, and $\alpha h < c$, then by Proposition 4, we know that the provision probability without bundling converges to zero for each good. However, if $h + l \geq 2c/\left(2\alpha - \alpha^2\right)$, the mechanism proposed above will provide both public goods with probability one. It is also easy to show that there exists configurations of l, h, c such that $l < c, \alpha h < c$ and $h + l \geq 2c/\left(2\alpha - \alpha^2\right)$. We summarize the above discussions as follows:

Claim. *Fix any $c > 0, \alpha \in (0,1)$. There exists $h > l$ such that the provision probability under the optimal separate provision mechanisms is zero but they are provided with probability one under the proposed bundle mechanism.*

The range of values of h and l for any $c > 0$, and $\alpha \in (0,1)$ for which the above-proposed bundling mechanism outperforms the optimal mechanism without bundling is depicted in Figure 2. The intuition for the improvement of bundling mechanism is as follows. In the revenue maximizing mechanism without bundling, only high valuation types are included in the public good provision, thus a fraction α^2 of the agents are included in both goods, and a fraction $2\alpha(1-\alpha)$ agents are included on one and only one good, and the remainder agents are excluded from both goods. In the proposed bundling mechanism, all agents are included in both public goods except that type-ll agents are excluded from both. Thus in the bundling mechanism, more agents contribute since only $(1-\alpha)^2$ consumers are excluded, even though the contribution is smaller per agent.

4.3 The Full Design Problem with Binary Valuations

We will now solve the design problem to maximize social surplus (8) subject to the incentive compatibility constraints in (5), the feasibility constraint (6) and the participation constraints (7).

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

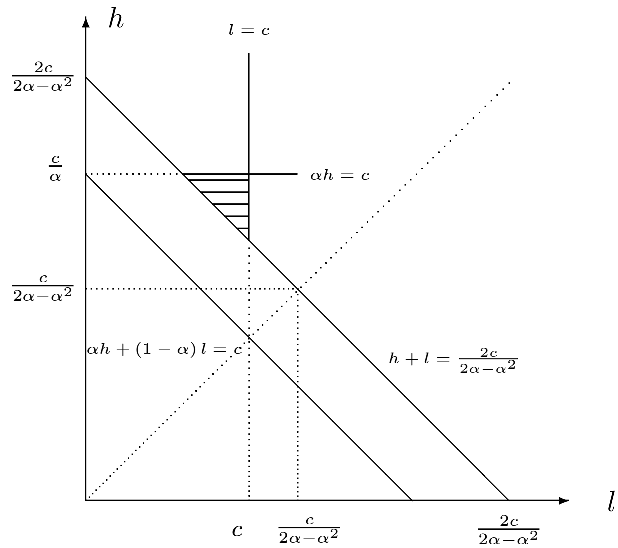

Figure 2: The Bundling Mechanism Outperforms Optimal Non-bundling Mechanism in the Shaded Region.

denote the number of agents announcing different types, and let

$$\mathcal{X}_n = \left\{ x \in \{0, ..., n\}^4 : x_{hh} + x_{hl} + x_{lh} + x_{ll} = n \right\}. \tag{19}$$

be the set of possible values $x(\theta)$ can take when there are n agents. The anonymity of the mechanism implies that the provision rule depends only on the *number of agents* who announce different valuation combinations. That is, with some abuse of notation, the class of mechanisms we consider is

$$M = \left(\left\{ \rho^j, \eta^j \right\}_{j=1,2}, t \right), j \in \{1, 2\}, \tag{20}$$

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

$$\rho^1 (x_{hh}, x_{hl}, x_{lh}, x_{ll}) \ = \ \rho^2 (x_{hh}, x_{lh}, x_{hl}, x_{ll}), \tag{21a}$$

$$\eta^1_{hh} \ = \ \eta^2_{hh}, \eta^1_{hl} \ = \ \eta^2_{lh}, \eta^1_{lh} \ = \ \eta^2_{hl}, \eta^1_{ll} \ = \ \eta^2_{ll}, t_{hl} \ = \ t_{lh}. \tag{21b}$$

To ease notation, we define

$$s^j(x, \eta) = \left(\eta^j_{hh} x_{hh} + \eta^j_{hl} x_{hl} \right) h + \left(\eta^j_{lh} x_{lh} + \eta^j_{ll} x_{ll} \right) l - cn$$

as the surplus generated if good $j = 1, 2$ is provided in state x (given truth-telling) if the inclusion probability is η^j.

Now we provide more details of the constraints that a mechanism $m \in M$ must satisfy. First, we need to be explicit about the probability distribution of the random variable x. For any $x = (x_{hh}, x_{hl}, x_{lh}, x_{ll}) \in \mathcal{X}_n$, the probability mass of x, denoted by $\mathbf{a}_n(x)$ is:

$$\mathbf{a}_n(x) = \frac{n!}{x_{hh}!x_{hl}!x_{lh}!x_{ll}!} \left(\alpha^2\right)^{x_{hh}} [\alpha(1-\alpha)]^{x_{hl}} [\alpha(1-\alpha)]^{x_{lh}} \left[(1-\alpha)^2\right]^{x_{ll}}. \tag{23}$$

Now we write down the incentive compatibility constraints. In principle, there are a total of twelve incentive constraints to be satisfied. However, in our environment, types are naturally ordered as hh being the "highest type", hl and lh being "middle types" and ll being the "lowest type". We therefore 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 hh is better off announcing her true type than type hl and hl is better off announcing her true type than ll, then there are no incentives for hh to announce ll. Together with the symmetry of the mechanism (21), we have two distinct incentive constraints constraints:

$$\begin{aligned} 2\eta^1_{hh} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll})\,h - t_{hh} \;\geq\; \\ \eta^1_{hl} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}+1, x_{lh}, x_{ll})h \\ +\eta^1_{lh} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}+1, x_{ll})h - t_{hl}, \end{aligned} \tag{24a}$$

$$\begin{aligned} \eta^1_{hl} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}+1, x_{lh}, x_{ll})h \\ +\eta^1_{lh} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}+1, x_{ll})l - t_{hl} \;\geq\; \\ \eta^1_{ll} \textstyle\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\,(h+l) - t_{ll}, \end{aligned} \tag{24b}$$

where (24a) states that type-hh agents do not have incentives to mis-report as type hl; and (24b) states that type-hl agents do not have incentives to mis-report as type ll.

Next, 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.[19] Hence, using symmetry (21), we can write the only relevant participation constraint as

$$2\eta^1_{ll} \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)l - t_{ll} \;\geq\; 0. \tag{25}$$

[19]The argument is that, by incentive compatibility, all higher types are better off than pretending to be type ll. Since the payoff from pretending to be ll is higher than the payoff for a (truth-telling) type ll, interim indivdual rationality follows for any other type.

Finally, the budget balance constraint can be simplified considerably due to the simple transfer schemes and the constant per capita costs. That is,

$$\mathrm{E}\left(\sum_i \tau_i(\theta) - \sum_{j=1,2} \rho^j(x) C^j(n)\right) = \sum_i \mathrm{E}\tau_i(\theta) - \sum_{j=1,2} \mathrm{E}\rho^j(x) C^j(n)$$
$$= n\left[\alpha^2 t_{hh} + \alpha(1-\alpha)[t_{hl} + t_{lh}] + (1-\alpha)^2 t_{ll} - \sum_{j=1,2} \mathrm{E}\rho^j(x) c\right].$$

We may thus express the budget balance constraint (6) in per capita form, using symmetry (21) and the explicit expression for $\mathrm{E}\rho^j(x)$, as

$$\alpha^2 t_{hh} + \alpha(1-\alpha) 2t_{hl} + (1-\alpha)^2 t_{ll} - 2c \sum_{x \in \mathcal{X}_n} \mathbf{a}_n(x) \rho^1(x) \geq 0. \tag{26}$$

Using again the symmetry (21), we can thus express the relaxed programming problem as:[20]

$$\max_{\{\rho^1, \eta^1, t\}} 2 \sum_{x \in \mathcal{X}_n} \mathbf{a}_n(x) \rho^1(x) \left[\frac{\left(\eta^1_{hh} x_{hh} + \eta^1_{hl} x_{hl}\right) h + \left(\eta^1_{lh} x_{lh} + \eta^1_{ll} x_{ll}\right) l}{n} - c\right] \tag{27}$$

s.t. (24a)-(24b), (25) and (26),

$$\eta^1_{\theta_i} \geq 0, 1 - \eta^1_{\theta_i} \geq 0 \text{ for each } \theta_i \in \Theta_i, \tag{28}$$

$$\rho^1(x) \geq 0, 1 - \rho^1(x) \geq 0 \text{ for each } x \in \mathcal{X}_n, \tag{29}$$

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

Lemma 3 *There exists at least one optimal solution to (27).*

The proof is standard by first compactifying the constraint set and then applying Weierstrass Theorem.

5 The Optimal Mechanism

It can be shown that Slater's' condition for constraint qualification holds, so the Kuhn-Tucker conditions are necessary for an optimum. Since we know that a solution to (27) exists, these first order conditions therefore provide a characterization of the optimal mechanism, provided that the constraints that we ignored when formulating (27) are satisfied at the candidate solution.

[20]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.

Constraint	**Multiplier**
IC (24a)	λ_{hh}
IC (24b)	λ_{hl}
IR (25)	λ_{ll}
BB (26)	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 1: Notation of multipliers.

5.1 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. Table 1 lists our notations for associated multipliers to optimization problem (27).

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

$$\begin{aligned} &(\text{w.r.t. } t_{hh}) && -\lambda_{hh} + \Lambda\alpha^2 = 0, \\ &(\text{w.r.t. } t_{hl}) && \lambda_{hh} + \lambda_{hl} + \Lambda 2\alpha(1-\alpha) = 0, \\ &(\text{w.r.t. } t_{ll}) && \lambda_{hl} - \lambda_{ll} + \Lambda(1-\alpha)^2 = 0. \end{aligned} \tag{30}$$

We thus immediately conclude:

Lemma 4 *In any solution to (27) the multipliers* $(\lambda_{hh}, \lambda_{hl}, \lambda_{ll}, \Lambda)$ *satisfy:* $\lambda_{hh} = \alpha^2\Lambda, \lambda_{hl} = (2\alpha - \alpha^2)\Lambda$, *and* $\lambda_{ll} = \Lambda$.

5.2 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

$$\begin{aligned} G(\Phi) &\equiv (1-\Phi)2l + \Phi\left[\frac{2\alpha - \alpha^2}{\alpha(1-\alpha)}l - \frac{\alpha^2}{\alpha(1-\alpha)}h\right], \\ H(\Phi) &\equiv (1-\Phi)2l + \Phi\left[\frac{2}{(1-\alpha)^2}l - \frac{2\alpha - \alpha^2}{(1-\alpha)^2}(h+l)\right]. \end{aligned} \tag{31}$$

The result is:

Lemma 5 *Let* $M = (\rho^1, \rho^2, \eta^1, \eta^2, t)$ *be a symmetric solution to (27) and let* $\Phi = \Lambda/(1+\Lambda)$, *where* Λ *is the associated multiplier on the resource constraint. Also, suppose that* $E\left[\rho^j(x)|\theta_i\right] > 0$ *for all* $\theta_i \in \Theta_i$ *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. $\eta^1_{ll} = \eta^2_{ll} = \begin{cases} 1 & \text{if } H(\Phi) > 0 \\ y \in [0,1] & \text{if } H(\Phi) = 0 \\ 0 & \text{if } H(\Phi) < 0. \end{cases}$

The formal is proof is relegated to the appendix. Here we provide an interpretation of this result. Note that $\Phi = \Lambda/(1+\Lambda) \in [0,1]$, and $G(\Phi) \geq 0$ if and only if

$$(1-\Phi)\overbrace{\alpha(1-\alpha)2l}^{\text{Term 1}} + \Phi\overbrace{\left[\alpha(2-\alpha)l - \alpha^2 h\right]}^{\text{Term 2}} \geq 0. \tag{32}$$

for all Φ. To understand term 2 in expression (32), consider two candidate inclusion rules. The first candidate is $\eta^1_{lh} = \eta^2_{hl} = \eta^1_{ll} = \eta^2_{ll} = 0$, which states that agent i 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 $\eta^1_{lh} = \eta^2_{hl} = 1$ and $\eta^1_{ll} = \eta^2_{ll} = 0$, which states that 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 $\left[\alpha^2 + 2\alpha(1-\alpha)\right](h+l) = \alpha(2-\alpha)(h+l)$. Thus, the marginal revenue effect of increasing η^1_{lh} and η^2_{hl} from 0 to 1 is given by

$$\alpha(2-\alpha)(h+l) - \alpha 2h = \alpha(2-\alpha)l - \alpha^2 h,$$

which is term 2 in expression (32). Term 1 in expression (32), $2\alpha(1-\alpha)l$, 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. Thus we show that $G(\Phi)$ is essentially a weighted average of the optimality conditions for an unconstrained social planner and a profit maximizing provider, where the weight on term 2 – 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 $\eta_{lh}^1 = \eta_{hl}^2 = 1$. On the other hand, if term 2 is negative, i.e. if $\alpha(2-\alpha)l < \alpha^2 h$, then some algebra on expression (32) shows that $G(\Phi) \geq 0$ if

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

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, $\eta_{lh}^1 = \eta_{hl}^2 = 1$ will be optimal only when Λ, or the shadow price of resources is sufficiently low. To summarize, item 2 of Lemma 5 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

$$(1-\Phi)(1-\alpha)^2 2l + \Phi[2l - \alpha(2-\alpha)(h+l)] \geq 0.$$

The term $(1-\alpha)^2 2l$ is the gain in social surplus when η_{ll}^1 and η_{ll}^2 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 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 $\eta_{ll}^1 = \eta_{ll}^2 = 1$ is optimal. Otherwise, $H(\Phi) \geq 0$ if and only if

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

Note

$$\frac{\Phi_{lh}^*}{\Phi_{ll}^*} = \frac{2-\alpha}{1-\alpha} > 1.$$

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 $\eta_{ll}^1 = \eta_{ll}^2 > 0$, then we know that $\Phi \leq \Phi_{ll}^* < \Phi_{lh}^*$, thus $\eta_{lh}^1 = \eta_{hl}^2 = 1$; symmetrically, if $\eta_{lh}^1 = \eta_{hl}^2 < 1$, then we know $\Phi \geq \Phi_{lh}^* > \Phi_{ll}^*$, then $\eta_{ll}^1 = \eta_{ll}^2 = 0$. We summarize the above discussion as:

Lemma 6 *Suppose that $E[\rho^j(x)|\theta_i] > 0$ for all $\theta_i \in \Theta_i$ and $j = 1,2$. Let $\Phi = \Lambda/(1+\Lambda)$. 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 we have not yet determined the value of Λ in the optimal mechanism, we have a rather simple characterization of the optimal inclusions as a function of the still unknown multiplier on the resource constraint.

5.3 Optimal Provision Rules

Now we analyze the optimal provision rules $\rho^j(x)$. Take the first order condition with respect to $\rho^1(x)$ in maximization problem (27), we obtain

$$\begin{aligned}
&2\mathbf{a}_n(x)\left[\tfrac{\left(\eta^1_{hh}x_{hh}+\eta^1_{hl}x_{hl}\right)h+\left(\eta^1_{lh}x_{lh}+\eta^1_{ll}x_{ll}\right)l}{n}-c\right]+\lambda_{hh}\left[2\eta^1_{hh}\mathbf{a}_{n-1}\left(x_{hh}-1,x_{hl},x_{lh},x_{ll}\right)h\right]\\
&-\lambda_{hh}\left[\eta^1_{hl}\mathbf{a}_{n-1}\left(x_{hh},x_{hl}-1,x_{lh},x_{ll}\right)h-\eta^1_{hl}\mathbf{a}_{n-1}\left(x_{hh},x_{hl},x_{lh}-1,x_{ll}\right)h\right]\\
&+\lambda_{hl}\left[\eta^1_{hl}\mathbf{a}_{n-1}\left(x_{hh},x_{hl}-1,x_{lh},x_{ll}\right)h+\eta^1_{lh}\mathbf{a}_{n-1}\left(x_{hh},x_{hl},x_{lh}-1,x_{ll}\right)l\right]\\
&-\lambda_{hl}\left[\eta^1_{ll}\mathbf{a}_{n-1}\left(x_{hh},x_{hl},x_{lh},x_{ll}-1\right)(h+l)\right]+\lambda_{ll}2\eta^1_{ll}\mathbf{a}_{n-1}\left(x_{hh},x_{hl},x_{lh},x_{ll}-1\right)l\\
&-\Lambda\mathbf{a}_n(x)\,2c+\gamma(x)-\phi(x)=0,
\end{aligned} \tag{33}$$

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

$$\begin{aligned}
\mathbf{a}_{n-1}\left(x_{hh}-1,x_{hl},x_{lh},x_{ll}\right)&=\tfrac{\mathbf{a}_n(x)}{\alpha^2}\tfrac{x_{hh}}{n}, &\qquad \mathbf{a}_{n-1}\left(x_{hh},x_{hl}-1,x_{lh},x_{ll}\right)&=\tfrac{\mathbf{a}_n(x)}{\alpha(1-\alpha)}\tfrac{x_{hl}}{n}\\
\mathbf{a}_{n-1}\left(x_{hh},x_{hl},x_{lh}-1,x_{ll}\right)&=\tfrac{a(x)}{\alpha(1-\alpha)}\tfrac{x_{lh}}{n}, &\qquad \mathbf{a}_{n-1}\left(x_{hh},x_{hl},x_{lh},x_{ll}-1\right)&=\tfrac{1}{(1-\alpha)^2}\tfrac{x_{ll}}{n},
\end{aligned} \tag{34}$$

exploiting the relationships between multipliers in Lemma 4, and substituting $\eta^1_{hh}=\eta^1_{hl}=1$ due to Lemma 5, we can simplify (33) to

$$\begin{aligned}
&2\left[\tfrac{\left(x_{hh}+x_{hl}\right)h+\left(\eta^1_{lh}x_{lh}+\eta^1_{ll}x_{ll}\right)l}{n}-c\right](1-\Phi)\\
&+\alpha^2\Phi\left[2\tfrac{1}{\alpha^2}\tfrac{x_{hh}}{n}h-\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{hl}}{n}h-\eta^1_{lh}\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{lh}}{n}h\right]\\
&+\alpha(2-\alpha)\Phi\left[\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{hl}}{n}h+\eta^1_{lh}\tfrac{1}{\alpha(1-\alpha)}\tfrac{x_{lh}}{n}l-\eta^1_{ll}\tfrac{1}{(1-\alpha)^2}\tfrac{x_{ll}}{n}(h+l)\right]\\
&+\Phi2\eta^1_{ll}\tfrac{1}{(1-\alpha)^2}\tfrac{x_{ll}}{n}l-\Phi2c+\tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)}=0,
\end{aligned} \tag{35}$$

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).

We can collect terms in (35) and simplify it as

$$\begin{aligned}&2\tfrac{x_{hh}}{n}h + 2\tfrac{x_{hl}}{n}h - 2c + \tfrac{x_{lh}}{n}\eta^1_{lh}\overbrace{\left\{(1-\Phi)\,2l + \Phi\left[\frac{\alpha(2-\alpha)}{\alpha(1-\alpha)}l - \frac{\alpha^2}{\alpha(1-\alpha)}h\right]\right\}}^{G(\Phi)} \\ &+\tfrac{x_{lh}}{n}\eta^1_{ll}\overbrace{\left\{(1-\Phi)\,2l + \Phi\left[\frac{2}{(1-\alpha)^2}l - \frac{\alpha(2-\alpha)}{(1-\alpha)^2}\frac{x_{ll}}{n}(h+l)\right]\right\}}^{H(\Phi)} + \tfrac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)} = 0.\end{aligned} \tag{36}$$

¿From Lemma 5, we know that $\eta^1_{lh} = 0$ if $G(\Phi) < 0$ and $\eta^1_{ll} = 0$ if $H(\Phi) < 0$. Therefore, if we define

$$Q^1\left(\frac{x}{n},\Phi\right) \equiv \frac{x_{hh}}{n}h + \frac{x_{hl}}{n}h + \frac{x_{lh}}{n}\frac{\max\{0,G(\Phi)\}}{2} + \frac{x_{ll}}{n}\frac{\max\{0,H(\Phi)\}}{2} - c. \tag{37}$$

(36) can be further simplified as

$$2Q^1\left(\frac{x}{n},\Phi\right) + \frac{\gamma(x)-\phi(x)}{\mathbf{a}_n(x)} = 0.$$

Lemma 7 *Let $M = (\rho^1, \eta^1, t)$ be an optimal solution to (27) and $\Phi = \Lambda/(1+\Lambda)$ where Λ is the multiplier associated with the constraint (26) at the optimal solution. Then,*

1. *$\rho^1(x) = 1$ whenever $Q^1(x/n,\Phi) > 0$,*

2. *$\rho^1(x) = 0$ whenever $Q^1(x/n,\Phi) < 0$.*

Analogously, given the multiplier Λ associated with the balanced-budget constraint at the optimal solution, the optimal provision rule for good 2 is determined by the value of a function

$$Q^2\left(\frac{x}{n},\Phi\right) \equiv \frac{x_{hh}}{n}h + \frac{x_{lh}}{n}h + \frac{x_{hl}}{n}\max\{0,G(\Phi)\} + \frac{x_{ll}}{n}\max\{0,H(\Phi)\} - c,$$

such that $\rho^2(x) = 1$ whenever $Q^2(x/n,\Phi) > 0$ and $\rho^2(x) = 0$ whenever $Q^2(x/n,\Phi) < 0$.

To summarize, we have characterized the optimal inclusion and provision rules for any given value of the Lagrange multiplier Λ associated with the balanced-budget constraint. Such characterization provides some partial information regarding the asymptotic provision probability in the optimal mechanism with bundling. For example, the above characterization tells us that $\alpha h > c$ is a sufficient but not necessary condition for the provision probability to converge to one.[21] In contrast,

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

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. Note that

$$\begin{aligned} \lim_{n\to\infty} Q^1\left(\frac{x}{n}, \Phi_n\right) &= Q^1(\mu, \Phi) \\ &= \alpha h + \alpha(1-\alpha)\max\{0, G(\Phi)\} + (1-\alpha)^2 \max\{0, H(\Phi)\} - c \end{aligned} \tag{38}$$

where $\Phi = \lim_{n\to\infty} \Phi_n$. Thus a sufficient condition for $Q^1(\mu, \Phi) > 0$ (and hence probability one provision of public good 1 asymptotically) is $\alpha h > c$. A similar conclusion can be obtained for public good 2.

5.4 Asymptotic Results

In this section, we provide a full characterization of the asymptotic provision probability in an optimal mechanism. We index the mechanisms by the size of the economy by $\left\{\rho_n^j, \eta_n^j, t_n\right\}_{j=1}^2$, 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 we have used our result from Lemma 5 that 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*

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

where

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

Bundling\Exclusion	No Exclusion	Exclusion
No Bundling	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite [6])	$\mathrm{E}\rho_n^{j*} \to 0,$ if $\alpha h < c$ $\mathrm{E}\rho_n^{j*} \to 1,$ if $\alpha h > c$ (Norman [11])
Bundling Allowed	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite [6])	$\mathrm{E}\rho_n^{j*} \to 0,$ 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 2: The Asymptotic Provision Probability under Different Bundling and Exclusion Possibilities.

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

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

5.5 Summary

In Sections 4 and 5, we studied the optimal bundling mechanism in an environment with n agents and two public goods. Each agent's valuations for the public goods are independent and take values h and l with probability α and $1-\alpha$ respectively. Suppose that the per capita provision cost of each public good is $c \in (l,h)$. Table 2 summarizes our results and the existing results in the literature with different possibilities of bundling and exclusion. The efficiency rationale for the bundling in the provision of excludable public goods is clearly demonstrated in large economies.

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 binary valuation parametric class of examples, 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. All these result are related to the idea that bundling reduces the variance in the distribution of valuations.

While we solved for the optimal mechanism for the provision of multiple public goods in a very specific example, the intuition highlighted in this paper that bundling of unrelated public goods can play an important role in the efficient mechanism is more general. Under reasonable conditions, bundling decreases the set of agents excluded from the public goods and it allows the mechanism designer to collect more revenues, which in turn implies that public goods may be more likely provided.

There are two interesting directions where we can extend this paper. First, can we characterize the optimal mechanism for the provision of multiple public goods when the valuation distributions are more general? Second, will a social planner sometimes find it optimal to bundle the provision of private goods with a public goods to alleviate the free-riding problem in the 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] Aumann, Robert. J., "Mixed and Behavioral Strategies in Infinite Extensive Games," *Annals of Mathematical Studies,* 52, 1964, 627-650.

[3] Bewley, Truman F., "A Critique of Tiebout's Theory of Local Public Expenditures." Econometrica, Vol 49, 1981, 713-740.

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

[5] 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.

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

[7] 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.

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

[9] Myerson, R. and M.A. Satterthwaite, "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory,* **29,** 1983, 265-281. 619-632.

[10] Nalebuff, Barry, "Bundling," Yale ICF Working Paper No. 99-14, 1999.

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

[12] Proschan, Frank, "Peakedness of Distributions of Convex Combinations," *Annals of Mathematical Statistics,* 36(6), December 1965, 1703-1706.

A Appendix: Proofs of Main Results

Proof of Proposition 1. Proposition 1 follows from the following two claims:

Claim 1 *For any incentive feasible mechanism $\mathcal{M}$ of the form (3), there exist an incentive feasible mechanism*

$$\left(\left(\rho^j, \eta_1^j, ..., \eta_n^j\right)_{j=1,2}, t_1, ..., t_n\right) \tag{A1}$$

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

Proof. Pick an arbitrary mechanism $\mathcal{M} = (\zeta^1, \zeta^2, \omega^1, \omega^2, \tau)$ and let $k \in [0,1]$. For $j = 1,2$ and $i \in \mathcal{I}$, define the functions $\rho^j : \Theta \to [0,1]$, $\eta_i^j : \Theta_i \to [0,1]$, $t_i : \Theta_i \to R$ by

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

Consider agent i of type $\theta_i \in \Theta_i$ who announces $\hat{\theta}_i \in \Theta_i$. Suppose $\int_{\Theta_{-i}}\int_0^1 \zeta^j\left(\hat{\theta}_i, \theta_{-i}, x\right) dxd\mathbf{F}(\theta_{-i}) = 0$. In this case, we know that $\mathrm{E}_{-i}\left[\zeta^j(\hat{\theta}_i, \theta_{-i}, x)\omega_i^j(\hat{\theta}_i, \theta_{-i}, x)\right] = 0$ since $\omega_i^j(\hat{\theta}_i, \theta_{-i}, x)$ is bounded between 0 and 1. Thus agent i's expected payoff is $-t_i(\hat{\theta}_i) = -\int_{\Theta_{-i}} \tau\left(\hat{\theta}_i, \theta_{-i}\right) d\theta_{-i}$ under both the original and the simple mechanism. In the other case where $\int_{\Theta_{-i}}\int_0^1 \zeta^j\left(\hat{\theta}_i, \theta_{-i}, x\right) dxd\mathbf{F}(\theta_{-i}) > 0$, agent i's utility given the mechanism $\left(\left(\rho^j, \eta_1^j, ..., \eta_n^j\right)_{j=1,2}, t_1, ..., t_n\right)$ is

$$\begin{aligned}
& \mathrm{E}_{-i}\left[\sum_{j=1,2} \rho^j\left(\hat{\theta}_i, \theta_{-i}\right)\eta_i^j\left(\hat{\theta}_i\right)\theta_i - t_i\left(\hat{\theta}_i\right)\right] \\
=& \sum_{j=1,2} \mathrm{E}_{-i}\zeta^j\left(\hat{\theta}_i, \theta_{-i}, x\right)\frac{\mathrm{E}_{-i}\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, x\right)\zeta^j\left(\hat{\theta}_i, \theta_{-i}, x\right)}{\mathrm{E}_{-i}\zeta^j\left(\hat{\theta}_i, \theta_{-i}, x\right)}\theta_i - \mathrm{E}_{-i}\tau\left(\hat{\theta}_i, \theta_{-i}\right) \\
=& \sum_{j=1,2} \mathrm{E}_{-i}\omega_i^j\left(\hat{\theta}_i, \theta_{-i}, x\right)\zeta^j\left(\hat{\theta}_i, \theta_{-i}, x\right)\theta_i - \mathrm{E}_{-i}\tau\left(\hat{\theta}_i, \theta_{-i}\right) \\
=& \mathrm{E}_{-i}\left[\sum_{j=1,2} \omega_i^j\left(\hat{\theta}_i, \theta_{-i}, x\right)\zeta^j\left(\hat{\theta}_i, \theta_{-i}, x\right)\theta_i - \tau\left(\hat{\theta}_i, \theta_{-i}\right)\right],
\end{aligned}$$

which is the utility that agent i of type θ_i obtains if she announces type $\hat{\theta}_i$ under the original mechanism $\mathcal{M} = (\zeta^1, \zeta^2, \omega^1, \omega^2, \tau)$. Hence, all incentive and participation constraints continue to hold when we change from $\mathcal{M}$ to the simplified mechanism. Moreover, since

$$\mathrm{E}_{-i}\left[\sum_{j=1,2} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i\right] = \mathrm{E}_{-i}\left[\sum_{j=1,2} \zeta^j(\theta,x)\,\omega_i^j(\theta,x)\,\theta_i\right] \text{ for every } \theta_i,$$

it follows by integration over Θ_i and summation over i that

$$\mathrm{E}\left[\sum_{i\in\mathcal{I}}\sum_{j=1,2} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i\right] = \mathrm{E}\left[\sum_{i\in\mathcal{I}}\sum_{j=1,2} \zeta^j(\theta,x)\,\omega_i^j(\theta,x)\,\theta_i\right].$$

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

$$\sum_{j=1,2}\mathrm{E}\rho^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_i^j(\theta_i)\,\theta_i - C^j(n)\right] = \sum_{j=1,2}\mathrm{E}\zeta^j(\theta,x)\left[\sum_{i\in\mathcal{I}}\omega_i^j(\theta,x)\,\theta_i - C^j(n)\right].$$

Hence, the original and the simplified mechanisms generate the same social surplus. Finally, $t_i(\theta_i) = \mathrm{E}_{-i}\tau(\theta)$ implies that $\sum_{i\in\mathcal{I}}\mathrm{E}t_i(\theta_i) = \sum_{i\in\mathcal{I}}\mathrm{E}\tau(\theta)$, thus the balanced-budget constraint is also unaffected. We conclude that the simplified mechanism constructed in (A1) generates the same social surplus and satisfies all the constraints if the original mechanism $\mathcal{M}$ is incentive feasible. ■

Claim 2 *For every simplified incentive feasible mechanism of the form (A1), there exists an anonymous simple incentive feasible mechanism $\widetilde{M}$ of the form (9) that generates the same social surplus.*

Proof. Suppose that a simplified mechanism $M = \left(\left(\rho^j, \eta_1^j, ..., \eta_n^j\right)_{j=1,2}, t_1, ..., t_n\right)$ of the form *(A1)* is incentive feasible. For any given $\theta \in \Theta$, let $P_k(\theta) \in \Theta$ denote the k-th permutation of θ; $\mathcal{P}(\theta) = \{P_1(\theta), ..., P_{n!}(\theta)\}$ the set of all possible permutations; and, for $i \in \mathcal{I}$, let $P_k^i(\theta) \in \Theta_i$ denote the agent i's type in permutation $P_k(\theta)$; and $\chi_k^i \in \mathcal{I}$ denote the index of the agent who has i's valuation θ_i in permutation $P_k(\theta)$.[22] For each $k \in \{1, ..., n!\}$, construct $M_k = \left(\left(\rho_k^j, \eta_{k1}^j, ..., \eta_{kn}^j\right)_{j=1,2}, t_{k1}, ..., t_{kn}\right)$ from the simplified mechanism M as

$$\begin{aligned}
\rho_k^j(\theta) &= \rho^j(P_k(\theta)) \quad \forall\, \theta\in\Theta,\ j=1,2,\\
\eta_{ki}^j(\theta_i) &= \eta_{\chi_k^i}^j(\theta_i) \quad \forall\, \theta_i\in\Theta_i,\ j=1,2, i\in\mathcal{I},\\
t_{ki}(\theta_i) &= t_{\chi_k^i}(\theta_i) \quad \forall\, \theta_i\in\Theta_i, , i\in\mathcal{I}.
\end{aligned}$$

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

Now let mechanism $\widetilde{M} = \left(\left(\widetilde{\rho}^j, \widetilde{\eta}_1^j, ..., \widetilde{\eta}_n^j\right)_{j=1,2}, \widetilde{t}_1, ..., \widetilde{t}_n\right)$ be

$$\begin{aligned}
\widetilde{\rho}^j(\theta) &= \frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta) \quad \forall\, \theta \in \Theta,\ j = 1,2 \\
\widetilde{\eta}_i^j(\theta_i) &= \frac{\sum_{k=1}^{n!} \mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]} \quad \forall\, \theta_i \in \Theta_i, i \in \mathcal{I},\ j = 1,2 \\
\widetilde{t}_i(\theta_i) &= \frac{1}{n!}\sum_{k=1}^{n!} t_{ki}(\theta_i) \quad \forall\, \theta_i \in \Theta_i, i \in \mathcal{I}.
\end{aligned}$$

By construction, mechanism $\widetilde{M}$ is anonymous since $\widetilde{\rho}^j(\theta) = \widetilde{\rho}^j(\theta')$ if θ' is a permutation of θ. Now we argue that $\widetilde{\eta}_i^j(\theta_i)$ is the same for all $i \in \mathcal{I}$. Consider agent i and i', and suppose that $\theta_i = \theta_{i'}$. It is to see that, by construction, the sets $\left\{\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)\right\}_{k=1}^{n!}$ and $\left\{\mathrm{E}_{-i'}\left[\rho_k^j(\theta)\right]\eta_{ki'}^j(\theta_{i'})\right\}_{k=1}^{n!}$ are identical; moreover, $\mathrm{E}_{-i}\left[\widetilde{\rho}^j(\theta)\right] = \mathrm{E}_{-i'}\left[\widetilde{\rho}^j(\theta)\right]$. Therefore $\widetilde{\eta}_i^j(\theta_i) = \widetilde{\eta}_{i'}^j(\theta_{i'})$ for all $i, i' \in \mathcal{I}$ if $\theta_i = \theta_{i'}$. Finally, it is clear that $\widetilde{t}_i(\theta_i) = \widetilde{t}_{i'}(\theta_{i'})$ for all $i, i' \in \mathcal{I}$ if $\theta_i = \theta_{i'}$. Thus there exists $\widetilde{\eta}^j(\theta_i), \widetilde{t}$ such that $\left(\widetilde{\eta}_i^j(\cdot), \widetilde{t}_i(\cdot)\right) = \left(\widetilde{\eta}^j(\cdot), \widetilde{t}(\cdot)\right)$ for all $i \in \mathcal{I}$. Therefore mechanism $\widetilde{M}$ is of the form (9).

Now we show that $\widetilde{M}$ is incentive feasible and maximizes the social planner's expected surplus. First, since M and M_k are identical except for the permutation of the agents, we have, for $k = 1, ..., n!$,

$$\sum_{j=1,2}\mathrm{E}\left\{\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_{ki}^j(\theta_i)\theta_i^j - C^j(n)\right]\right\} = \sum_{j=1,2}\mathrm{E}\left\{\rho^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_i^j(\theta_i)\theta_i^j - C^j(n)\right]\right\}. \tag{A2}$$

Hence,

$$\begin{aligned}
&\sum_{j=1,2}\mathrm{E}\left\{\widetilde{\rho}^j(\theta)\left[\sum_{i\in\mathcal{I}}\widetilde{\eta}^j(\theta_i)\theta_i^j - C^j(n)\right]\right\} \\
=\ & \sum_{j=1,2}\mathrm{E}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}}\frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]}\theta_i^j - C^j(n)\right]\right\} \\
=\ & \sum_{j=1,2}\sum_{i\in\mathcal{I}}\mathrm{E}_{\theta_i}\mathrm{E}_{-i}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\left[\frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]}\theta_i^j\right]\right\} - \mathrm{E}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]C^j(n) \\
=\ & \sum_{j=1,2}\sum_{i\in\mathcal{I}}\mathrm{E}_{\theta_i}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)\theta_i^j\right\} - \mathrm{E}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]C^j(n) \\
=\ & \sum_{j=1,2}\sum_{i\in\mathcal{I}}\mathrm{E}\left\{\frac{1}{n!}\sum_{k=1}^{n!}\left[\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\theta_i^j\right]\right\} - \mathrm{E}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]C^j(n)
\end{aligned}$$

$$= \frac{1}{n!}\sum_{k=1}^{n!}\sum_{j=1,2}\mathrm{E}\left\{\rho_k^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_{ki}^j(\theta_i)\theta_i^j - C^j(n)\right]\right\}$$
$$= \sum_{j=1,2}\mathrm{E}\left\{\rho^j(\theta)\left[\sum_{i\in\mathcal{I}}\eta_i^j(\theta_i)\theta_i^j - C^j(n)\right]\right\},$$

where the last equality follows from (A2). Hence the social surplus generated by the simple anonymous mechanism $\widetilde{M}$ is identical to that by original mechanism M.

Now we show that the constructed mechanism $\widetilde{M}$ is incentive feasible. First, since the agents' valuations are drawn from identical distributions and M_k and M only differ in the index of the agents, we have $\mathrm{E}\rho_k^j(\theta) = \mathrm{E}\rho^j(\theta)$ 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. Thus

$$\mathrm{E}\sum_{i\in\mathcal{I}}\widetilde{t}_i(\theta_i) - \sum_{j=1,2}\mathrm{E}\widetilde{\rho}^j(\theta)C^j(n)$$
$$= \mathrm{E}\sum_{i\in\mathcal{I}}\frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\theta_i) - \sum_{j=1,2}\mathrm{E}\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)C^j(n)$$
$$= \mathrm{E}\sum_{i\in\mathcal{I}}t_i(\theta_i) - \sum_{j=1,2}\mathrm{E}\rho^j(\theta)C^j(n),$$

hence the simple mechanism $\widetilde{M}$ is budget balanced if the original mechanism M is. Second, we note that incentive compatibility holds for any permuted mechanism, i.e.,

$$\mathrm{E}_{-i}\sum_{j=1,2}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\theta_i^j - t_{ki}(\theta_i) \geq \mathrm{E}_{-i}\sum_{j=1,2}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j - t_{ki}(\widehat{\theta}_i,\theta_{-i}) \quad \text{(A3)}$$
$$\forall i \in \mathcal{I}, \text{ and } \theta_i,\widehat{\theta}_i\in\Theta_i.$$

Hence,

$$\mathrm{E}_{-i}\sum_{j=1,2}\widetilde{\rho}^j(\theta)\widetilde{\eta}^j(\theta_i)\theta_i^j - \widetilde{t}(\theta_i)$$
$$= \mathrm{E}_{-i}\sum_{j=1,2}\left[\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\theta)\right]\frac{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]\eta_{ki}^j(\theta_i)}{\sum_{k=1}^{n!}\mathrm{E}_{-i}\left[\rho_k^j(\theta)\right]}\theta_i^j - \frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\theta_i)$$
$$= \mathrm{E}_{-i}\sum_{j=1,2}\frac{1}{n!}\sum_{k=1}^{n!}\mathrm{E}_{-i}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\theta_i^j - \frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\theta_i)$$
$$= \frac{1}{n!}\sum_{k=1}^{n!}\left[\mathrm{E}_{-i}\sum_{j=1,2}\rho_k^j(\theta)\eta_{ki}^j(\theta_i)\theta_i^j - t_{ki}(\theta_i)\right]$$

$$
\begin{aligned}
&\geq \frac{1}{n!}\sum_{k=1}^{n!}\left[\mathrm{E}_{-i}\sum_{j=1,2}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j - t_{ki}(\widehat{\theta}_i,\theta_{-i})\right] \\
&= \mathrm{E}_{-i}\sum_{j=1,2}\frac{1}{n!}\sum_{k=1}^{n!}\rho_k^j(\widehat{\theta}_i,\theta_{-i})\eta_{ki}^j(\widehat{\theta}_i,\theta_{-i})\theta_i^j - \frac{1}{n!}\sum_{k=1}^{n!}t_{ki}(\widehat{\theta}_i,\theta_{-i}) \\
&= \sum_{j=1,2}\mathrm{E}_{-i}\widetilde{\rho}^j(\widehat{\theta}_i,\theta_{-i})\widetilde{\eta}_i^j\left(\widehat{\theta}_i\right)\theta_i^j - \widetilde{t}(\widehat{\theta}_i),
\end{aligned}
$$

where the inequality follows from (A3). Hence $\widetilde{M}$ is incentive compatible. Finally, $\widetilde{M}$ also satisfies the participation constraints hold because, as is clear from the third line in the above calculation, all the permutation mechanisms satisfy participation constraints. ■

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

Proof. Consider a simple anonymous incentive feasible mechanism $\widetilde{M} = \left\{\widetilde{\rho}^j, \widetilde{\eta}^j, \widetilde{t}\right\}_{j=1}^2$. Let $\widehat{M}$ be the "mirror image" of $\widetilde{M}$ in the sense that the roles of goods 1 and 2 are reversed. Specifically, for every $\theta = (\theta^1, \theta^2)$ let

$$
\begin{aligned}
&\widehat{\rho}^1\left(\theta^1,\theta^2\right) = \widetilde{\rho}^2\left(\theta^2,\theta^1\right) \text{ and } \widehat{\rho}^2\left(\theta^1,\theta^2\right) = \widetilde{\rho}^1\left(\theta^2,\theta^1\right), \\
&\widehat{\eta}^1(\theta_i^1,\theta_i^2) = \widetilde{\eta}^2(\theta_i^2,\theta_i^1) \text{ and } \widehat{\eta}^2(\theta_i^1,\theta_i^2) = \widetilde{\eta}^1(\theta_i^2,\theta_i^1), \\
&\widehat{t}(\theta_i^1,\theta_i^2) = \widetilde{t}(\theta_i^2,\theta_i^1).
\end{aligned} \tag{A4}
$$

We now show that mechanism $\widehat{M}$ is also a simple anonymous incentive feasible mechanism that generates the same social surplus as mechanism $\widetilde{M}$. Notice that

$$
\begin{aligned}
&\mathrm{E}\widehat{\rho}^1\left(\theta^1,\theta^2\right)\widehat{\eta}^1(\theta_i^1,\theta_i^2)\theta_i^1 \;=\; \mathrm{E}\widetilde{\rho}^2\left(\theta^2,\theta^1\right)\widetilde{\eta}^2(\theta_i^2,\theta_i^1)\theta_i^1 \\
&= \int_{\theta^1\in(\Theta^1)^n}\left[\int_{\theta^2\in(\Theta^2)^n}\widetilde{\rho}^2\left(\theta^2,\theta^1\right)\widetilde{\eta}^2(\theta_i^2,\theta_i^1)\theta_i^1\Pi_{k=1}^n dF^2(\theta_k^2|\theta_k^1)\right]\Pi_{k=1}^n dF^1(\theta_k^1) \\
&= \int_{\theta^2\in(\Theta^2)^n}\left[\int_{\theta^1\in(\Theta^1)^n}\widetilde{\rho}^2\left(\theta^1,\theta^2\right)\widetilde{\eta}^2(\theta_i^1,\theta_i^2)\theta_i^2\Pi_{k=1}^n dF^2(\theta_k^1|\theta_k^2)\right]\Pi_{k=1}^n dF^1(\theta_k^2) \\
&= \int_{\theta^2\in(\Theta^2)^n}\left[\int_{\theta^1\in(\Theta^1)^n}\widetilde{\rho}^2\left(\theta^1,\theta^2\right)\widetilde{\eta}^2(\theta_i^1,\theta_i^2)\theta_i^2\Pi_{k=1}^n dF^1(\theta_k^1|\theta_k^2)\right]\Pi_{k=1}^n dF^2(\theta_k^2) \\
&= \mathrm{E}\widetilde{\rho}^2\left(\theta^1,\theta^2\right)\widetilde{\eta}^2(\theta_i^1,\theta_i^2)\theta_i^2,
\end{aligned} \tag{A5}
$$

where the third equality comes from the fact that labeling of the integrand variables is arbitrary, and the fourth from the assumption that $F^1(\cdot|v) = F^2(\cdot|v)$. Similarly,

$$
\begin{aligned}
&\mathrm{E}\widehat{\rho}^2\left(\theta^1,\theta^2\right)\widehat{\eta}^2(\theta_i^1,\theta_i^2)\theta_i^2 = \mathrm{E}\widetilde{\rho}^1\left(\theta^1,\theta^2\right)\widetilde{\eta}^1(\theta_i^1,\theta_i^2)\theta_i^1 \\
&\mathrm{E}\widehat{t}(\theta_i^1,\theta_i^2) = \mathrm{E}\widetilde{t}(\theta_i^1,\theta_i^2)
\end{aligned} \tag{A6}
$$

Together, (A5) and (A6) imply that the ex ante utility for each $i \in \mathcal{I}$ is the same under mechanism $\widehat{M}$ as that under mechanism $\widetilde{M}$. Moreover, since $C^1(n) = C^2(n)$ and $\mathrm{E}\widehat{\rho}^1\left(\theta^1,\theta^2\right) = \mathrm{E}\widetilde{\rho}^2\left(\theta^1,\theta^2\right)$

and $\mathrm{E}\widehat{\rho}^2\left(\theta^1,\theta^2\right)=\mathrm{E}\widetilde{\rho}^1\left(\theta^1,\theta^2\right)$, the expected provision costs are the same, thus $\widehat{M}$ and $\widetilde{M}$ generate the same social surplus. Since provision costs are the same and $\mathrm{E}\widehat{t}(\theta_i^1,\theta_i^2)=\mathrm{E}\widetilde{t}(\theta_i^1,\theta_i^2)$, $\widehat{M}$ satisfies balanced-budget constraint (6). We finally show that $\widehat{M}$ satisfies participation and incentive compatibility constraints. Calculations similar to that in (A5) show that

$$\begin{aligned}\mathrm{E}_{-i}\widehat{\rho}^1\left(\theta_i'^1,\theta_{-i}^1,\theta_i'^2,\theta_{-i}^2\right)&=\mathrm{E}_{-i}\widetilde{\rho}^2\left(\theta_i'^2,\theta_{-i}^1,\theta_i'^1,\theta_{-i}^2\right),\\ \mathrm{E}_{-i}\widehat{\rho}^2\left(\theta_i'^1\theta_{-i}^1,\theta_i'^2,\theta_{-i}^2\right)&=\mathrm{E}_{-i}\widetilde{\rho}^1\left(\theta_i'^2,\theta_{-i}^1,\theta_i'^1,\theta_{-i}^2\right).\end{aligned}$$

Write $U(\theta_i,\theta_i';\widehat{M})$ for the expected utility of type θ_i agent from announcing θ_i' under mechanism $\widehat{M}$. We have

$$\begin{aligned}&U\left(\theta_i^1,\theta_i^2,\theta_i'^1,\theta_i'^2;\widehat{M}\right)=\mathrm{E}_{-i}\sum_{j=1,2}\widehat{\rho}^j\left(\theta_i'^1,\theta_{-i}^1,\theta_i'^2,\theta_{-i}^2\right)\widehat{\eta}^j\left(\theta_i'^1,\theta_i'^2\right)\theta_i^j-\widehat{t}(\theta_i^1,\theta_i^2)\\ =\;&\mathrm{E}_{-i}\widetilde{\rho}^1\left(\theta_i'^2,\theta_{-i}^1,\theta_i'^1,\theta_{-i}^2\right)\widetilde{\eta}^1\left(\theta_i'^2,\theta_i'^1\right)\theta_i^2+\mathrm{E}_{-i}\widetilde{\rho}^2\left(\theta_i'^2,\theta_{-i}^1,\theta_i'^1,\theta_{-i}^2\right)\widetilde{\eta}^2\left(\theta_i'^2,\theta_i'^1\right)\theta_i^1-\widetilde{t}(\theta_i^2,\theta_i^1)\\ =\;&U\left(\theta_i^2,\theta_i^1,\theta_i'^2,\theta_i'^1,M\right)\end{aligned}$$

Thus participation constraint holds for type $\theta_i=\left(\theta_i^1,\theta_i^2\right)$ consumer under mechanism $\widehat{M}$ if and only the participation constraint of type $\left(\theta_i^2,\theta_i^1\right)$ consumder is satisfied under mechanism $\widetilde{M}$. Same is true for the incentive compatibility constraint. We conclude that $\widehat{M}$ is incentive feasible and generates the same social surplus as $\widetilde{M}$.

Now construct a new mechanism M where

$$\begin{aligned}\rho^j(\theta)&=\frac{1}{2}\widetilde{\rho}^j(\theta)+\frac{1}{2}\widehat{\rho}^j(\theta),\\ \eta^j(\theta_i)&=\frac{\widetilde{\eta}^j(\theta)\,\mathrm{E}_{-i}\widetilde{\rho}^j(\theta)+\widehat{\eta}^j(\theta)\,\mathrm{E}_{-i}\widehat{\rho}^j(\theta)}{\mathrm{E}_{-i}\widetilde{\rho}^j(\theta)+\mathrm{E}_{-i}\widehat{\rho}^j(\theta)},\\ t(\theta_i)&=\frac{1}{2}\widetilde{t}(\theta_i)+\frac{1}{2}\widehat{t}(\theta_i).\end{aligned}$$

By construction of mechanim $\widehat{M}$, mechanism M is symmetric. Arguments similar to those above show that mechanism M satisfies budget balanced, incentive compatible and participation constraints. We now show that mechanism M generates the same social surplus as mechanism $\widetilde{M}$. We also notice that

$$\begin{aligned}&\mathrm{E}\rho^1(\theta)\,\eta^1(\theta_i^1,\theta_i^2)\theta_i^1\\ =\;&\mathrm{E}\left[\left(\frac{1}{2}\widetilde{\rho}^1(\theta)+\frac{1}{2}\widehat{\rho}^1(\theta)\right)\frac{\widetilde{\eta}^1(\theta_i^1,\theta_i^2)\mathrm{E}_{-i}\widetilde{\rho}^1(\theta)+\widehat{\eta}^1(\theta_i^1,\theta_i^2)\mathrm{E}_{-i}\widehat{\rho}^1(\theta)}{\mathrm{E}_{-i}\widetilde{\rho}^1(\theta)+\mathrm{E}_{-i}\widehat{\rho}^1(\theta)}\theta_i^1\right]\\ =\;&\mathrm{E}_{\theta_i}\left[\left(\frac{1}{2}\mathrm{E}_{-i}\widetilde{\rho}^1(\theta)+\frac{1}{2}\mathrm{E}_{-i}\widehat{\rho}^1(\theta)\right)\frac{\widetilde{\eta}^1(\theta_i^1,\theta_i^2)\mathrm{E}_{-i}\widetilde{\rho}^1(\theta)+\widehat{\eta}^1(\theta_i^1,\theta_i^2)\mathrm{E}_{-i}\widehat{\rho}^1(\theta)}{\mathrm{E}_{-i}\widetilde{\rho}^1(\theta)+\mathrm{E}_{-i}\widehat{\rho}^1(\theta)}\theta_i^1\right]\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\mathrm{E}\widetilde{\eta}^1(\theta_i^1,\theta_i^2)\widetilde{\rho}^1(\theta)\,\theta_i^1 + \frac{1}{2}\mathrm{E}\widehat{\eta}^1(\theta_i^1,\theta_i^2)\widehat{\rho}^1(\theta)\,\theta_i^1 \\
&= \frac{1}{2}\mathrm{E}\widetilde{\eta}^1(\theta_i^1,\theta_i^2)\widetilde{\rho}^1(\theta)\,\theta_i^1 + \frac{1}{2}\mathrm{E}\widetilde{\eta}^2(\theta_i^1,\theta_i^2)\widetilde{\rho}^2(\theta)\,\theta_i^2,
\end{aligned}
$$

where the last equality follows from (A5). Symmetrically

$$
\mathrm{E}\rho^2(\theta)\,\eta^2(\theta_i^1,\theta_i^2)\theta_i^2 = \frac{1}{2}\mathrm{E}\widetilde{\eta}^2(\theta_i^1,\theta_i^2)\widetilde{\rho}^2(\theta) + \frac{1}{2}\mathrm{E}\widetilde{\eta}^1(\theta_i^1,\theta_i^2)\widetilde{\rho}^1(\theta)\,\theta_i^1,
$$

thus,

$$
\sum_{j=1,2}\mathrm{E}\rho^j(\theta)\,\eta^j(\theta_i^1,\theta_i^2)\theta_i^j = \sum_{j=1,2}\mathrm{E}\widetilde{\rho}^j(\theta)\,\widetilde{\eta}^j(\theta_i^1,\theta_i^2)\theta_i^j.
$$

Moreover,

$$
\begin{aligned}
\mathrm{E}\sum_{j=1}^{2}\rho^j(\theta)\,C^j(n) &= \mathrm{E}\left[\frac{1}{2}\widetilde{\rho}^1(\theta)+\frac{1}{2}\widehat{\rho}^1(\theta)\right]C^1(n) + \mathrm{E}\left[\frac{1}{2}\widetilde{\rho}^2(\theta)+\frac{1}{2}\widehat{\rho}^2(\theta)\right]C^2(n) \\
&= \mathrm{E}\left[\frac{1}{2}\widetilde{\rho}^1(\theta)+\frac{1}{2}\widetilde{\rho}^2(\theta)+\frac{1}{2}\widetilde{\rho}^2(\theta)+\frac{1}{2}\widetilde{\rho}^1(\theta)\right]C^1(n) \\
&= \mathrm{E}\sum_{j=1}^{2}\widetilde{\rho}^j(\theta)\,C^j(n),
\end{aligned}
$$

where the equalities come from (A4) and the assumption that $C^1(n) = C^2(n)$. Hence the expected provison costs are unchanged. Thus, the total expected social surplus generated by mechanism M is the same as in $\widetilde{M}$. ■

Proof of Lemma 1.

Proof. Pick any $0 \le \epsilon \le (\overline{\theta} - \underline{\theta})/2$. Log-concavity and symmetry implies that

$$
\begin{aligned}
\ln f(\widetilde{\theta}) &= \ln f\left(\frac{1}{2}\left(\widetilde{\theta}-\epsilon\right)+\frac{1}{2}\left(\widetilde{\theta}+\epsilon\right)\right) \ge \frac{1}{2}\ln f(\widetilde{\theta}-\epsilon) + \frac{1}{2}\ln f(\widetilde{\theta}+\epsilon) \\
&= \ln f(\widetilde{\theta}-\epsilon) = \ln f(\widetilde{\theta}+\epsilon),
\end{aligned}
$$

thus $f(\widetilde{\theta}) \ge f(\theta)$ for all $\theta \in [\underline{\theta},\overline{\theta}]$. Hence, $f(\widetilde{\theta}) \ge 1/\left(\overline{\theta}-\underline{\theta}\right)$, since otherwise, $\int_{\underline{\theta}}^{\overline{\theta}} f(\theta)\,d\theta \le \int_{\underline{\theta}}^{\overline{\theta}} f(\widetilde{\theta})d\theta = f(\widetilde{\theta})\left(\overline{\theta}-\underline{\theta}\right) < 1$, a contradiction. We can also check that log-concavity of f implies that $p[1-F(p)]$ is strictly single-peaked over $[\underline{\theta},\overline{\theta}]$. Now,

$$
\frac{d}{dp}\,p[1-F(p)]\big|_{p=\widetilde{\theta}} = \frac{1}{2} - \widetilde{\theta}f(\widetilde{\theta}) \le \frac{1}{2} - \frac{\widetilde{\theta}}{\overline{\theta}-\underline{\theta}} = \frac{1}{2} - \frac{1}{2}\frac{\overline{\theta}+\underline{\theta}}{\overline{\theta}-\underline{\theta}} \le 0,
$$

thus together with the single peakedness, we know that $p[1-F(p)]$ is decreasing in p for all $p \ge \widetilde{\theta}$. Hence $p_m \le \widetilde{\theta}$. Finally, if $f\left(\widetilde{\theta}\right) > 1/\left(\overline{\theta}-\underline{\theta}\right)$ or $\underline{\theta} > 0$, $d\left\{p[1-F(p)]\right\}/dp$ is strictly negative when evaluated at $p = \widetilde{\theta}$. Thus $p_m < \widetilde{\theta}$. ■

Proof of Lemma 5.

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

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

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^1_{hl} reads

$$\begin{aligned}&2\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,h\frac{x_{hl}}{n}-\lambda_{hh}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,h\\+\quad&\lambda_{hl}\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll})\,h+\gamma_{hl}-\phi_{hl}=0.\end{aligned}\tag{A8}$$

Use the multinomial relationship between $\mathbf{a}_n(x)$ and $\mathbf{a}_{n-1}(x)$, we obtain

$$\mathbf{a}_n(x)=\frac{n}{x_{hl}}\alpha(1-\alpha)\,\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll}),\tag{A9}$$

holds for any x such that $x_{hl}\geq 1$. Hence

$$\begin{aligned}&\sum_{x\in\mathcal{X}_n}\mathbf{a}_n(x)\,\rho^1(x)\,h\frac{x_{hl}}{n}\\=\;&\sum_{x\in\mathcal{X}_n:x_{hl}\geq 1}\frac{n}{x_{hl}}\alpha(1-\alpha)\,\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,\rho^1(x)\,h\frac{x_{hl}}{n}\\=\;&\alpha(1-\alpha)\,h\sum_{x\in\mathcal{X}_n:x_{hl}\geq 1}\mathbf{a}_{n-1}(x_{hh},x_{hl}-1,x_{lh},x_{ll})\,\rho^1(x)\\=\;&\alpha(1-\alpha)\,h\sum_{x\in\mathcal{X}_{n-1}}\mathbf{a}_{n-1}(x)\,\rho^1(x_{hh},x_{hl}+1,x_{lh},x_{ll}).\end{aligned}\tag{A10}$$

Substituting (A10) into (A8), we obtain the condition

$$2\alpha(1-\alpha)\,h-\lambda_{hh}h+\lambda_{hl}h+\widehat{\gamma}_{hl}-\widehat{\phi}_{hl}=0,$$

where

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

By hypothesis of the Lemma,

$$\mathrm{E}\left[\rho^1(x)\,|\theta_i = hl\right] = \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}+1, x_{lh}, x_{ll}) > 0.$$

Thus the "rescaled multipliers" are well-defined, weakly positive, and equal to zero if and only if the "original multiplier" is equal to zero. From Lemma 4, we know that $\lambda_{hh} = \alpha^2\Lambda$ and $\lambda_{hl} = \left(2\alpha - \alpha^2\right)\Lambda$, where Λ is the multiplier on the balanced budget constraint. Hence the condition (**??**) can be rewritten as

$$2\alpha(1-\alpha)h - \alpha^2\Lambda h + \alpha(2-\alpha)\Lambda h + \widehat{\gamma}^1_{hl} - \widehat{\phi}^1_{hl} = 2\alpha(1-\alpha)h + 2\alpha\Lambda h + \widehat{\gamma}^1_{hl} - \widehat{\phi}^1_{hl} = 0.$$

Since $2\alpha(1-\alpha)h + 2\alpha\Lambda h > 0$, we conclude that $\widehat{\phi}^1_{hl} > 0$. Hence $\eta^1_{hl} = 1$ for all x by the complementarity slackness condition. By Proposition 2, $\eta^2_{lh} = 1$ follows. Steps 1 and 2 thus proves part (1) of the lemma.

[Step 3] Consider the optimality conditions for η^1_{lh}. To economize on derivations, we immediately observe that

$$\sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\,\rho^1(x_{hh}, x_{hl}, x_{lh}+1, x_{ll}) = \sum_{x\in\mathcal{X}_n : x_{lh}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\rho^1(x), \tag{A11}$$

and write the optimality condition as

$$\begin{array}{l} 2\sum_{x\in\mathcal{X}_n} \mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{lh}}{n} l - \lambda_{hh}\sum_{x\in\mathcal{X}_n : x_{lh}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\rho^1(x)h \\ +\lambda_{hl}\sum_{x\in\mathcal{X}_n : x_{lh}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\rho^1(x)l + \gamma_{lh} - \phi_{lh} \;=\; 0. \end{array} \tag{A12}$$

Again using a multinomial relationship that, for $x_{lh} \geq 1$,

$$\mathbf{a}_n(x) = \frac{n}{x_{lh}}\alpha(1-\alpha)\,\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll}),$$

we have

$$\begin{aligned}
\sum_{x\in\mathcal{X}_n} \mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{lh}}{n} l &= \sum_{x\in\mathcal{X}_n : x_{lh}\geq 1} \mathbf{a}_n(x)\,\rho^1(x)\,\frac{x_{lh}}{n} l \\
&= \sum_{x\in\mathcal{X}_n : x_{lh}\geq 1} \frac{n}{x_{lh}}\alpha(1-\alpha)\,\mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\,\rho^1(x)\,\frac{x_{lh}}{n} l \\
&= \alpha(1-\alpha)\,l \sum_{x\in\mathcal{X}_n : x_{lh}\geq 1} \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}-1, x_{ll})\,\rho^1(x) \\
&= \alpha(1-\alpha)\,l\mathrm{E}\left[\rho^1(x)\,|\theta_i = lh\right].
\end{aligned}$$

Plug this into (A12) and simplify, we obtain

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

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 (A11), we can write the first order condition for η^1_{ll} as

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

Using a 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 again rewrite the first order condition as

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

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 Proposition 5.

We proceed with a few useful lemmas. The first lemma says that as $n \to \infty$, $Q^1\left(\frac{x}{n}, \Phi_n\right) \to Q^1(\mu, \Phi_n)$ in probability:

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

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

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

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

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

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

we can apply Chebyshev's inequality to obtain

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

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

The second lemma is an application of the Stirling's Lemma:

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. 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

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

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

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

we obtain

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

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

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

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,

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

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

Now let

$$\rho_i^j(\theta_i) = \mathrm{E}\left[\rho^j(x)\,|\theta_i\right] \tag{A13}$$

be agent i's perceived probability that public good j will be provided when agent i annouces 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 $\theta_i, \theta_i' \in \Theta_i$ 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,

$$\overline{\rho}^1 \in \arg\max_{\rho^1:\mathcal{X}_n\to[0,1]} \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\left[\rho^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll}) - \rho^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\right], \tag{A14}$$

It is clear that the solution to (A14) is given by

$$\overline{\rho}^1(x) = \begin{cases} 1 & \text{if } \mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll}) \geq \mathbf{a}_{n-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1) \\ 0 & \text{if } \mathbf{a}_{n-1}(x_{hh}-1, x_{hl}, x_{lh}, x_{ll}) < \mathbf{a}_{-1}(x_{hh}, x_{hl}, x_{lh}, x_{ll}-1). \end{cases} \tag{A15}$$

Using the explicit formula for $\mathbf{a}_{n-1}(x)$, we can express (A15) as

$$\overline{\rho}^1(x) = \begin{cases} 1 & \text{if } \frac{x_{hh}}{\alpha^2} \geq \frac{x_{ll}}{(1-\alpha)^2} \\ 0 & \text{if } \frac{x_{hh}}{\alpha^2} < \frac{x_{ll}}{(1-\alpha)^2}. \end{cases} \tag{A16}$$

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

$$\Pr\left(\frac{m}{n-1} \geq 2\alpha(1-\alpha) + \epsilon\right) \leq \frac{\epsilon}{2} \tag{A17}$$

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 (A16) that, conditonal 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 proability is unaffected by agent i's announcement. Lemma A2 implies that there exists $N' < \infty$ such that

$$\Pr(x_{hh} = \overline{x}_{hh}(m) | m) \leq \frac{\epsilon}{2} \tag{A18}$$

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 $n \geq n^*$,

$$\begin{aligned} &\Pr[n-1-m \leq N'] = \Pr[m \geq (n-1)-N'] \\ &\leq \Pr[m \geq (n-1)-(n-1)[1-2\alpha(1-\alpha)-\epsilon]] \\ &= \Pr\left[\tfrac{m}{n-1} \geq 2\alpha(1-\alpha)+\epsilon\right] \leq \tfrac{\epsilon}{2} \end{aligned} \tag{A19}$$

where the last equality follows from (A17). Hence, for $n \geq n^*$, $n-1-m \leq N'$ with probability of at least $1-\epsilon/2$. Thus, for $n \geq n^*$,

$$\begin{aligned} &\overline{\rho}_i^1(hh) - \overline{\rho}_i^1(ll) = \textstyle\sum_{x \in \mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\left[\overline{\rho}^1(x_{hh}+1, x_{hl}, x_{lh}, x_{ll}) - \overline{\rho}^1(x_{hh}, x_{hl}, x_{lh}, x_{ll}+1)\right] \\ &= \textstyle\sum_{m=0}^{n-1} \Pr(m) \Pr(x_{hh} = \overline{x}_{hh}(m) | m) \\ &= \textstyle\sum_{m=0}^{n-1-N'} \Pr(m) \Pr(x_{hh} = \overline{x}_{hh}(m) | m) + \sum_{m=n-N'}^{n-1} \Pr(m) \Pr(x_{hh} = \overline{x}_{hh}(m) | m) \\ &\leq \textstyle\sum_{m=0}^{n-1-N'} \Pr(m) \frac{\epsilon}{2} + \sum_{m=n-N'}^{n-1} \Pr(m) \\ &= \tfrac{\epsilon}{2} \Pr[n-1-m \geq N'] + \Pr[n-1-m \leq N'] \leq \epsilon \end{aligned} \tag{A20}$$

where the second equality follows from the definition of $\overline{x}_{hh}(m)$; the first inequality follows from (A18); and the last inequality follows from (A19).

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

$$\underline{\rho}^1 \in \arg\min_{\rho:\mathcal{X}_n\to[0,1]} \sum_{x\in\mathcal{X}_{n-1}} \mathbf{a}_{n-1}(x)\left[\rho^1(x_{hh}+1,x_{hl},x_{lh},x_{ll}) - \rho^1(x_{hh},x_{hl},x_{lh},x_{ll}+1)\right], \quad \text{(A21)}$$

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 (A21) is

$$\underline{\rho}^1(x) = \begin{cases} 1 & \text{if } \frac{x_{hh}}{\alpha^2} < \frac{x_{ll}}{(1-\alpha)^2} \\ 0 & \text{if } \frac{x_{hh}}{\alpha^2} \geq \frac{x_{ll}}{(1-\alpha)^2}, \end{cases} \quad \text{(A22)}$$

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

$$\underline{\rho}_i^1(hh) - \underline{\rho}_i^1(ll) = -\sum_{m=0}^{n-1} \Pr(m)\Pr(x_{hh} = \overline{x}_{hh}(m)|\, m) \geq -\epsilon. \quad \text{(A23)}$$

It follows from (A20) and (A23) that, for any conceivable provision rule,

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

■

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 $\theta_i \in \Theta_i$ 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 $n \geq N, \theta_i, \theta_i' \in \Theta_i$. Then

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

(Proof of Proposition 5). 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 (38), 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 7 and A1. Suppose instead that $\alpha(2-\alpha)(h+l) > 2c \geq 2\alpha h$. Then,

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

Observe that

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

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 7 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 (A13) to re-write the incentive compatibility constraint (24b), after using the characterization of inclusion rule in Lemma 6, as

$$\rho_i^1(hl) h + \rho_i^1(lh) \eta_{lh}^1 l - t_{hl} \geq \rho_i^1(ll) \eta_{ll}^1 (h+l) - t_{ll} \geq \rho_i^1(ll) \eta_{ll}^1 (h-l), \tag{A24}$$

where the second inequality comes from the participation constraint (25). Pick an arbitrary $\epsilon > 0$. Then, by Lemma A4, there exists finite N such that for every $n \geq N_1$ and each $\theta_i \in \Theta_i$, for $j = 1,2$,

$$\left| \rho_i^j(\theta_i) - \mathrm{E}\rho_n^1(x) \right| < \epsilon_1 \equiv \frac{\epsilon}{3h}. \tag{A25}$$

Substituting (A25) into (A24), we obtain that for all $n \geq N_1$,

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

which implies that

$$\begin{aligned} t_{hl} &\leq \mathrm{E}\rho_n^1(x)\left[h\left(1-\eta_{ll}^1\right)+\left(\eta_{lh}^1+\eta_{ll}^1\right)l\right]+\epsilon_1\left[h+\eta_{lh}^1 l+\eta_{ll}^1(h-l)\right] \\ &< \mathrm{E}\rho_n^1(x)\left[h\left(1-\eta_{ll}^1\right)+\left(\eta_{lh}^1+\eta_{ll}^1\right)l\right]+\underbrace{3h\epsilon_1}_{\epsilon}. \end{aligned} \tag{A26}$$

Similarly, incentive constraints (24a) can be rewritten as:

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

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

$$\begin{aligned} t_{hh} &< 2\left[\mathrm{E}\rho_n^1(x)\right]h-\mathrm{E}\rho_n^1(x)\left(1+\eta_{lh}^1\right)h+t_{hl}+\epsilon \\ &= \mathrm{E}\rho_n^1(x)\left(1-\eta_{lh}^1\right)h+t_{hl}+\epsilon \\ &< \mathrm{E}\rho_n^1(x)\left(1-\eta_{lh}^1\right)h+\mathrm{E}\rho_n^1(x)\left[h\left(1-\eta_{ll}^1\right)+\left(\eta_{lh}^1+\eta_{ll}^1\right)l\right]+\epsilon \\ &= \mathrm{E}\rho_n^1(x)\left[\left(2-\eta_{ll}^1-\eta_{lh}^1\right)h+\left(\eta_{lh}^1+\eta_{ll}^1\right)l\right]+\epsilon. \end{aligned} \tag{A27}$$

Finally, from the participation constraint (25), there exists N_3 such that for all $n > N_3$,

$$t_{ll} < 2\mathrm{E}\rho_n^1(x)\,\eta_{ll}^1 l+\epsilon. \tag{A28}$$

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 (A26)-(**??**), we can bound the total expected tax revenue as follows:

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

Note that

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

Therefore,

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

which implies that

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

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.

CASE 2: $\eta^1_{ll} = \eta^2_{ll} = \eta_l \in (0,1), \eta^1_{lh} = \eta^2_{hl} = 1$. We can again use (A26)-(**??**) to bound the total expected tax revenue as follows:

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

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

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

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 $\eta^1_n(lh) = \eta^2_n(hl) = 1$ for all $n \ge N$. Then, taking a subsequence if necessary, we have that $\eta^1_n(lh) = \eta^2_n(hl) < 1$ for all n, which, by Lemma 6, implies that $\eta^j_n(ll) = 0$ for all n in the sequence. The per capita surplus generated by the optimal mechanism M_n in the n^{th} economy in the sequence, denoted by $S(M_n)$, is then

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

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

$$\frac{S(M_n)}{n} \leq 2[\alpha h + \alpha(1-\alpha) l - c] + \varepsilon. \tag{A29}$$

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

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

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

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

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

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

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

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

Now, the expected per capita surplus generates by $\widetilde{M_n}$ is

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

Let $\varepsilon = (1-\alpha)^2 \eta_{ll}^* l > 0$, we know from (A29) that there exists $N < \infty$ such that

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

which implies that mechanisms M_n could not be optimal for $n \geq N$, a contradiction.

Now we have concluded that in the sequence $\{M_n\}$, $\eta_n^1(lh) = \eta_n^2(hl) = 1$ for every $n \geq N$. What is left to show is that $\eta_n^1(ll)$ does converge to η_{ll}^* in the sequence $\{M_n\}$. 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

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

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

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

thus again constradicts the optimality of the mechanism 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

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

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

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

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

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

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

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

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

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

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 analagous to Part 3. Suppose to the contrary that in the sequence of mechanisms $\{M_n\}$, 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 (6) 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

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

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

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

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

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

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 Appendix: Proof of Technical Lemmas

Proof of Lemma 2.

Proof. We can, with no loss of generality, assume that the distribution is symmetric around 0. That is, we can assume $\underline{x}=-\overline{x}$, where $\overline{x}>0$, and let $\widetilde{x}=0$. For $0<\lambda<1/2$ and $t\in[\underline{x},0]$, we define

$$\widetilde{F}(t,\lambda)=\Pr[\lambda x_1+(1-\lambda)x_2\leq t]=\int_{-\overline{x}}^{\overline{x}}F\left(\frac{t-(1-\lambda)u}{\lambda}\right)f(u)\,du$$

where the second inequality uses the symmetry in density, $f(u)=f(-u)$. Taking the derivative with respect to λ we have that

$$\begin{aligned}
\frac{\partial\widetilde{F}(t,\lambda)}{\partial\lambda} &= \frac{1}{\lambda^2}\int_{-\overline{x}}^{\overline{x}}f\left(\frac{t-(1-\lambda)u}{\lambda}\right)(u-t)f(u)\,du \\
&= \frac{1}{\lambda^2}\left[\underbrace{\int_{-\overline{x}}^{t}f\left(\frac{-t+(1-\lambda)u}{\lambda}\right)(u-t)f(u)\,du}_{\equiv T_1}-\underbrace{\int_{t}^{\overline{x}}f\left(\frac{-t+(1-\lambda)u}{\lambda}\right)(u-t)f(u)\,du}_{\equiv T_2}\right].
\end{aligned}$$

Using change of variable $v=t-u$ in the first integral, we get

$$\begin{aligned}
T_1 &= \int_{-\overline{x}}^{t}f\left(\frac{(1-\lambda)(u-t)-\lambda t}{\lambda}\right)(u-t)f(u-t+t)\,du \\
&= -\int_0^{t+\overline{x}}f\left(\frac{-(1-\lambda)v-\lambda t}{\lambda}\right)vf(t-v)\,dv \\
&= -\int_0^{t+\overline{x}}f\left(\frac{(1-\lambda)v+\lambda t}{\lambda}\right)vf(v-t)\,dv,
\end{aligned}$$

where the last equality uses the symmetry of the density. Now, since $t\leq0$, $v-t\geq v$. Thus we can use our notational convention that $f(u)=0$ if $u\notin[-\overline{x},\overline{x}]$ to rewrite T_1 as

$$T_1=-\int_0^{\overline{x}}f\left(\frac{(1-\lambda)v+\lambda t}{\lambda}\right)vf(v-t)\,dv.$$

Similarly, use change of variable $v = u - t$, we have

$$\begin{aligned} T_2 &= \int_t^{\overline{x}} f\left(\frac{(1-\lambda)(u-t)-\lambda t}{\lambda}\right)(u-t) f(u-t+t)\, du \\ &= \int_0^{\overline{x}-t} f\left(\frac{(1-\lambda)v-\lambda t}{\lambda}\right) v f(v+t)\, dv. \end{aligned}$$

Again, note that if $t < 0$, then for all $v \geq \overline{x}, [(1-\lambda)v - \lambda t]/\lambda > (1-\lambda)\overline{x}/\lambda \geq \overline{x}$ since $\lambda \leq 1/2$, thus $f([(1-\lambda)v - \lambda t]/\lambda) = 0$ for all $v \geq \overline{x}$. =Hence we can rewrite T_2 as

$$T_2 = \int_0^{\overline{x}} f\left(\frac{(1-\lambda)v-\lambda t}{\lambda}\right) v f(v+t)\, dv.$$

Hence,

$$\lambda^2 \frac{\partial \widetilde{F}(t,\lambda)}{\partial \lambda} = T_1 + T_2 = \int_0^{\overline{x}} v\left[f(rv-t) f(v+t) - f(rv+t) f(v-t)\right] dv,$$

where $r = (1-\lambda)/\lambda > 1$. Note

$$\begin{aligned} f(rv-t) f(v+t) - f(rv+t) f(v-t) &= -\int_t^{-t} \frac{d}{ds}\left[f(rv-s) f(v+s)\right] ds \\ = \underbrace{\int_t^{-t} \left[f'(rv-s) f(v+s) - f(rv-s) f'(v+s)\right] ds.}_{\equiv T_3} \end{aligned}$$

Consider two cases. In the first case, $rv + t \geq v - t$. Since $t < 0$, this implies that $rv - s \geq v + s$ for any $s < -t$. Since log-concavity of the density f impies

$$\frac{f'(rv-s)}{f(rv-s)} \leq \frac{f'(v+s)}{f(v+s)},$$

thus,

$$f'(rv-s) f(v+s) - f(rv-s) f'(v+s) \leq 0 \text{ for all } s \in (t,-t)$$

Hence in this case $T_3 \leq 0$. In the second case, $rv + t < v - t$. Then let $s^* = (r-1)v/2 < 0$. We can rewrite T_3 as

$$\begin{aligned} T_3 &= \int_t^{t+2s^*} \left[f'(rv-s) f(v+s) - f(rv-s) f'(v+s)\right] ds \\ &\quad + \int_{t+2s^*}^{-t} \left[f'(rv-s) f(v+s) - f(rv-s) f'(v+s)\right] ds, \end{aligned}$$

where we note that $t < t + 2s^* < -t$ under the restriction in this case that $rv + t < v - t$. But,

$$\begin{aligned}
&\int_{t+2s^*}^{-t} \left[f'(rv-s)\,f(v+s) - f(rv-s)\,f'(v+s)\right] ds \\
=&\; -\int_{t+2s^*}^{-t} \left[\frac{d}{ds} f(rv-s)\,f(v+s)\right] ds \\
=&\; f(rv-t-2s^*)\,f(v+t+2s^*) - f(rv+t)\,f(v-t) \\
=&\; f(rv-t-(r-1)v)\,f(v+t+(r-1)v) - f(rv+t)\,f(v-t) \\
=&\; f(v-t)\,f(t+rv) - f(rv+t)\,f(v-t) = 0.
\end{aligned}$$

Hence,

$$T_3 = \int_t^{t+2s^*} \left[f'(rv-s)\,f(v+s) - f(rv-s)\,f'(v+s)\right] ds.$$

Since $s^* < -t$ we have that $t + 2s^* < s^*$. Thus some simple calculation shows that $rv - s > v + s$ for all $s \leq t + 2s^*$.[23] Hence, again using log-concavity, we have

$$f'(rv-s)\,f(v+s) \leq f'(v+s)\,f(rv-s) \text{ for any } s \leq t + 2s^*.$$

Thus $T_3 \leq 0$ in this case as well. We henceforth conclude that for any $t < 0$,

$$\lambda^2 \frac{\partial \widetilde{F}(t,\lambda)}{\partial \lambda} \leq 0.$$

Finally, to show that the inequality is strict, we notice that for $t < 0$ and $v > (\overline{x} + t)/r$ we have that $rv - t > \overline{x} \Rightarrow f(rv - t) = 0$

$$\begin{aligned}
\lambda^2 \frac{\partial \widetilde{F}(t,\lambda)}{\partial \lambda} &= \int_0^{\overline{x}} \underbrace{v\left[f(rv-t)\,f(v+t) - f(rv+t)\,f(v-t)\right]}_{\leq 0 \text{ for all } v} dv \\
&\leq \int_{\frac{\overline{x}+t}{r}}^{\overline{x}+t} v\left[f(rv-t)\,f(v+t) - f(rv+t)\,f(v-t)\right] dv \\
&= -\int_{\frac{\overline{x}+t}{r}}^{\overline{x}+t} v f(rv+t)\,f(v-t)\,dv
\end{aligned}$$

Since $v - t < \overline{x}$ for $v < \overline{x} + t$ and $rv + t < \overline{x}$ for $v \leq \frac{\overline{x}-t}{r}$ and $\frac{\overline{x}+t}{r} < \frac{\overline{x}-t}{r}$ it follows that $f(rv+t)\,f(v-t) > 0$ over the interval $\left[\frac{\overline{x}+t}{r}, \min\left\{\frac{\overline{x}-t}{r}, \overline{x}+t\right\}\right]$. Since

$$\frac{\overline{x}+t}{r} < \min\left\{\frac{\overline{x}-t}{r}, \overline{x}+t\right\} \text{ for any } r > 1 \text{ and } -\overline{x} < t < 0$$

[23] It suffices to show that $rv - (t + 2s^*) > v + (t + 2s^*)$. To see this, note

$$\begin{aligned}
rv - (t + 2s^*) &= rv - [t + (r-1)\,v] = v - t \\
&> v + s^* > v + (t + 2s^*).
\end{aligned}$$

it follows that $\lambda^2 \partial \widetilde{F}(t,\lambda) / \partial \lambda < 0$. Hence

$$F_A(t) = \widetilde{F}(t,0) + \int_0^{\frac{1}{2}} \frac{\partial \widetilde{F}(t,\lambda)}{\partial \lambda} d\lambda = F(t) + \int_0^{\frac{1}{2}} \frac{\partial \widetilde{F}(t,\lambda)}{\partial \lambda} d\lambda < F(t)$$

for every $-\overline{x} < t < 0$, which completes the proof. ■

Proof of Lemma 3.

Proof. For each $x \in X, j = 1,2, \theta_i \in \Theta_i$ we have that $\rho^j(x) \in [0,1], \eta^j_{\theta_i} \in [0,1]$. Next, we note that if $t_{ll} < 0$ and all constraints are satisfied, then a deviation where taxes are changed from t to $t' = (t_{hh}, t_{hl}, t_{lh}, 0)$ and where inclusion and provision rules are unchanged will satisfy all constraints in the relaxed program (27). Similarly, if all constraints hold and $t_{lh} < -l - h$ then the deviation to

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

will satisfy 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

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

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 $t_{\theta_i} > \overline{t} = 2h$ for some θ_i then at least one constraint in (27) must be violated. We therefore conclude that there is no loss in generality to restrict t_{θ_i} to be a number in $[\underline{t}, \overline{t}]$. 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. ■