Optimal Provision of Multiple Excludable Public Goods*

Hanming Fang† Peter Norman‡

This Version: January 2008

Abstract

This paper studies the optimal provision mechanism for multiple excludable public goods when agents' valuations are private information. For a parametric class of problems with binary valuations, we demonstrate that the optimal mechanism involves bundling if a regularity condition, akin to a hazard rate condition, on the distribution of valuations is satisfied. Bundling alleviates the free riding problem in large economies in two ways: first, it may increase the asymptotic provision probability of socially efficient public goods from zero to one; second, it decreases the extent of use exclusions. If the regularity condition is violated, then the optimal solution replicates the separate provision outcome.

Keywords: Public Goods Provision; Bundling; Exclusion

JEL Classification Number: H41.

*This paper substantially generalizes and supersedes Fang and Norman [8], a manuscript that was circulated under the title "An Efficiency Rationale for Bundling of Public Goods." We thank Mark Armstrong, Ted Bergstrom, Martin Hellwig, Larry Samuelson, Stephen Morris and seminar participants at many seminars and conferences for comments and helpful discussions. The usual disclaimer applies.

†Department of Economics, Duke University, 213 Social Sciences Building, Box 90097, Durham, NC 27708-0097; and the NBER. Email: `hanming.fang@duke.edu`.

‡Department of Economics, UNC-Chapel Hill, 107 Gardner Hall, Chapel Hill, NC 27599-3305. Email: `normanp@email.unc.edu`.

1 Introduction

This paper studies the optimal provision mechanism for multiple excludable public goods. We briefly consider a somewhat more general setup where we obtain some characterization results, but most of the paper focuses on a parametric version of the model where valuations are binary. In the binary valuation case, we demonstrate that there is a considerable degree of bundling in the optimal solution if a regularity condition, akin to a hazard rate condition, on the distribution of valuations is satisfied. If the regularity condition is violated, which happens when valuations are too strongly positively correlated across goods, the optimal solution replicates the separate provision outcome.

To motivate the importance of a better understanding of bundling of *non-rival* goods, we note that many goods that are provided in bundles are close to fully non-rival. The most striking example is the access to electronic libraries, for which the typical contractual arrangement is a site license that allows access to every issue of every journal in the electronic library. Another example is cable TV, where the basic pricing scheme consists of a limited number of available packages. Other examples include computer software and digital music files. For several of these cases, the pros and cons of bundling *for the consumer* have been frequently debated by the media, legal scholars, and in the courtroom. Still, there is no normative benchmark that explicitly considers the non-rival nature of these goods in the economics literature.

We consider a model with M excludable public goods, meaning that all goods are fully non-rival, but consumers can be excluded from usage. Each consumer is characterized by a valuation for each good, and the willingness to pay for a subset of goods is the sum of the individual good valuations. In addition, the cost of provision for each good is independent of which other goods are provided. Under these separability assumptions, the first best benchmark is to provide good j if and only if the sum of valuations for good j over all consumers exceeds its provision cost and to exclude no consumer from usage. Under perfect information, there is thus no role for either bundling or use exclusions. However, when preferences are private information, consumers must be given appropriate incentives to truthfully reveal their willingness to pay. Together with self-financing and participation constraints, it is then impossible to implement the (non-bundling) perfect information social optimum. Bundling is then potentially useful because it improves possibilities to extract surplus from the consumers, which will then relax a binding constraint on the problem.

The first part of the paper considers a relatively general setup, and characterizes the form of optimal provision mechanisms in symmetric environments. We then apply these results in the special case with binary valuations for which we obtain an exact characterization of the constrained efficient mechanism. To make the problem tractable, we impose symmetry conditions on costs and type distributions in addition to the restriction that valuations take on only two values for each of the M excludable public goods.

The solution to the problem is rather striking. When the economy is large in the sense that the number of consumers goes to infinity, the optimal mechanism will either provide all goods with probability close to one, or provide all goods with probability close to zero. Which of the two scenarios applies depends on whether or not a *monopolist profit maximizer* that provides the goods for sure could break even. If a regularity condition, akin to a hazard rate condition, on the distribution of valuations is satisfied, the optimal mechanism also prescribes a very simple rule for user access to the public goods once they are provided. All consumers will fall into one of three groups: the first group, consisting of those whose numbers of goods for which they have high valuations strictly exceed a threshold, will be given access to the grand bundle consisting of all goods; the second group, consisting of those whose number of goods for which they have high valuations are strictly lower than the threshold, will be given access to only those goods for which they have a high valuation; the third group, consisting of those with exactly the threshold-level number of high valuation goods, will always be given access to their high valuation goods, and will also be given access to their low valuation goods with some probability. Note that the third group of consumers for which some randomization is applied will be quite rare when there are many public goods, i.e., as M gets large.

For the special case with two goods we solve the problem also for the case where the regularity condition discussed above is violated. With two goods, the regularity condition is violated when the valuations are too positively correlated. This invalidates the standard approach to solve screening problems using only downwards adjacent constraint, as there are simply too few "mixed types" (those with one high and one low valuation) to justify giving them a better treatment than the type with two low valuations. We show that the optimal solution in fact is identical to the solution when both goods must be provided separately. It is interesting to relate this to the seminal contribution by Adams and Yellen [1]. Their explanation of bundling was that the monopolist seller knows more about the willingness to pay for the bundle than for the components provided that there is negative correlation. It has been shown that the relevant comparison is the willingness to pay for the components versus the willingness to pay for the average, implying that essentially the same explanation also works when valuations are independent.[1] Our result here shows that this logic falls apart when valuations are too positively correlated.

The analytical tractability for the multidimensional mechanism design problem in our paper comes from exploiting some important similarities to unidimensional problems. In particular, in the unidimensional case, it is known that maximizing social surplus subject to budget and participation constraints leads to a Lagrangian characterization which can be interpreted as a compromise between profit and welfare maximization (see Hellwig [10] and Norman [17]). In our model we cannot collapse the constraints to a single integral constraint, but we are able to use the optimality conditions to link the values of the multipliers associated with various constraints so that the optimal solution can be understood analogously.

[1] See, e.g., McAfee, McMillan and Whinston [15] and Fang and Norman [9].

The remainder of the paper is structured as follows. Section 2 presents the model and some characterization results. Section 3 introduces the special case when valuations are binary and derives the optimal mechanism in the regular case. In Section 4 we use some special cases to better interpret the characterization in Section 3, in particular, we characterize the optimal mechanism for the two good case when the regularity condition is violated. Section 5 contains a brief discussion of the relevance of our analysis with respect to concrete anti-trust issues. Appendix A contains all the proofs of results in Section 3.[2]

2 The Model

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

2.1 The Environment

There are M *excludable* public goods, labeled by $j \in \mathcal{J} = \{1, ..., M\}$ and n consumers, indexed by $i \in \mathcal{I} = \{1, ..., n\}$. Each public good is indivisible, and the cost of providing good j, denoted $C^j(n)$, is independent of which of the other goods are provided.[3] Since n is the number of *consumers* in the economy, *not* the number of *user*s, all goods are fully non-rival. The rationale for indexing cost by n is to be able to analyze large economies without making the public goods a "free lunch" in the limit. We therefore allow for the existence of $c^j > 0$ such that $\lim_{n\to\infty} C^j(n)/n = c^j > 0$. There is no need to give this assumption any economic interpretation. It is best viewed as a way to ensure that the provision problem remains "significant" with a large number of agents.

Consumer i is described by a valuation for each good $j \in \mathcal{J}$, so that her type is given by a vector $\theta_i = \left(\theta_i^1, ..., \theta_i^M\right) \in \Theta \subset R^M$. Agent i has preferences represented by the utility function,

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

where $\mathtt{I}_i^j$ is a dummy variable taking value 1 when i consumes good j and 0 otherwise, and t_i is the quantity of the numeraire good transferred from i to the mechanism designer. Preferences over lotteries are of expected utility form.

[2] An appendix available at the authors' websites contains the proofs for the more technical results in Section 2.

[3] Armstrong [2] [3] studied the profit maximization problem for multiproduct monopolist.

The type θ_i is private information to the agent. While we allow valuations *across goods* to be correlated for the individual, it is essential that we assume independence *across agents.* We denote by F the joint cumulative distribution over θ_i. For brevity of notation, we let $\theta \equiv (\theta_1, ..., \theta_n) \in \Theta^n$, which will be referred to as a *type profile.* In the usual fashion, we let $\theta_{-i} = (\theta_1, .., \theta_{i-1}, \theta_{i+1}, ...\theta_n)$.

2.2 Randomized Direct Mechanisms

An outcome in our environment has three components: (1). Which goods, if any, should be provided; (2). Who are to be given access to the goods that are provided; and (3). How to share the costs. The set of feasible *pure* outcomes is thus

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

By the revelation principle, we restrict attention to direct mechanisms for which truth-telling is a Bayesian Nash equilibrium. A pure direct mechanism is a map from Θ^n to A. We represent a randomized mechanism in analogy with the representation of mixed strategies in Aumann [4]. That is, let $\Xi \equiv [0,1]$, and think of $\vartheta \in \Xi$ as the outcome of a fictitious lottery, where, without loss of generality, ϑ is uniformly distributed and independent of θ. A *random direct mechanism* is then a measurable mapping $\mathcal{G} : \Theta^n \times \Xi \to A$. A conceptual advantage of this representation is that it allows for a useful decomposition. That is, we may write $\mathcal{G}$ as a $(2M+1)$-tuple, $\mathcal{G} = (\{\zeta^j\}_{j\in\mathcal{J}}, \{\omega^j\}_{j\in\mathcal{J}}, \tau)$ where,

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

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

2.3 The Design Problem

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

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

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

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

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

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

A mechanism is *incentive feasible* if it satisfies (IC), (BB) and (IR). Utility is transferable, implying that constrained ex ante Pareto efficient allocations may be characterized by solving a utilitarian planning problem, where a fictitious social planner seeks to maximize total surplus in the economy, subject to the constraints (IC), (BB) and (IR).[5] Thus a mechanism is thus *constrained efficient* if it maximizes

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

over all incentive feasible mechanisms.[6]

It is ex post efficient to provide good j if and only if $\sum_{i\in\mathcal{I}}\theta_i^j\geq C^j(n)$, and to never exclude any agent from usage, which is the same rule as the first best rule for a single public good. This is implementable if and only if a *non-excludable* public good can be efficiently provided under (IC), (BB) and (IR). (See Mailath and Postlewaite [12].) Our setup is thus a second best problem.

[4] As shown in Borgers and Norman [7] it is without loss of generality to consider a resource constraint in ex ante form. Given independence and two or more agents, transfers can be adjusted so as to satisfy ex post budget balance without changing the interim expected payoff for any individual.

[5] The qualifier *ex ante* is crucial for the equivalence. See Ledyard and Palfrey [13] for a characterization of *interim* efficiency.

[6] All these constraints are noncontroversial if the design problem is interpreted as a private bargaining agreement. If the goods are government provided, the participation constraints (IR) may seem questionable. One defense in this context is that the participation constraint is a reduced form of an environment where agents may vote with their feet. Another defense is to view this as a reduced form for inequality aversion of the planner. See Hellwig [11].

2.4 Simple Anonymous Mechanisms

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

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

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

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

There are a number of simplifications in (5) relative to (3). First, inclusion and transfer rules are the same for all agents; second, conditional on θ, the provision probability $\rho^j(\theta)$ is stochastically independent from all other provision probabilities, and all inclusion probabilities; third, the inclusion and transfer rules for any agent i are independent of the realization of θ_{-i}; and fourth, all agents are treated symmetrically in terms of the transfer and inclusion rules. But (5) allows provision rules to treat agents asymmetrically. We therefore need a definition to express what it means for the name of an agent to be irrelevant:

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

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

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

The idea is roughly that risk neutral agents care only about the perceived probability of consuming each good and the expected transfer. Therefore, there is nothing to gain from conditioning transfers and inclusion probabilities on θ_{-i}, or by making inclusion and provision rules conditionally dependent. Mechanisms of the form (5) are therefore sufficient. Moreover, given an incentive feasible mechanism, permuting the roles of the agents leaves the surplus unchanged and all constraints satisfied. An anonymous incentive feasible mechanism that generates the same surplus as the initial mechanism can therefore be obtained by averaging over the $n!$ permuted mechanisms.[7]

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

2.5 Symmetric Treatment of the Goods

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

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

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

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

In defining a symmetric mechanism, the *same permutation of goods must be applied for all agents.* As an example, suppose that there are two agents and two goods, and that the valuation for each good is either h or l. In this case $\Theta = \{(h,h), (h,l), (l,h), (l,l)\}$. Consider the type profile $\theta = (\theta_1, \theta_2) = ((h,l), (l,h)) \in \Theta^2$. Applying the only non-identity permutation of the goods, i.e., $P(1) = 2$ and $P(2) = 1$, to all agents generates a type profile $\theta^P = \left(\theta_1^P, \theta_2^P\right) = ((l,h), (h,l))$. Definition 3 requires that the allocations for type profile $((l,h), (h,l))$ is the same as the allocation for $((h,l), (l,h))$ with goods relabeled, and that transfers are unchanged.[8] The result is:

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

The idea is similar to that of Proposition 1, except that it is the identities of the goods that are permuted. Consider the case with two goods, and suppose that the two goods are treated

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

asymmetrically. Reversing the role of the goods, an alternative mechanism that generates the same surplus is obtained. Averaging over the original and the reversed mechanism creates a symmetric mechanism where surplus is unchanged.[9] Incentive feasibility of the new mechanism follows from incentive feasibility of the original mechanism. Proposition 2 generalizes this procedure by permuting the goods ($M!$ possibilities) and creating a symmetric mechanism by averaging over these permuted mechanisms.

3 The Case with Binary Valuations

In the remainder of the paper we assume that the valuation for each good j is either high or low, so that $\theta_i^j \in \{l, h\}$ for each $i \in \mathcal{I}$ and $j \in \mathcal{J}$. The probability of any type $\theta_i \in \Theta \equiv \{l, h\}^M$ is denoted $\beta_i(\theta_i)$, and using independence $\beta(\theta) = \times_{i=1}^n \beta_i(\theta_i)$ and $\beta_{-i}(\theta_{-i}) = \times_{i' \neq i} \beta_{i'}(\theta_{i'})$ denote the probabilities of type profile θ and θ_{-i} respectively.[10] Moreover, suppose that $C^j(n) = cn$ for all $j \in \mathcal{J}$.[11]

For any $\theta_i \in \Theta = \{l, h\}^M$, write $m(\theta_i) \in \{0, ..., M\}$ as the number of goods for which $\theta_i^j = h$, i.e.,

$$m(\theta_i) = \# \left\{ j \in \mathcal{J} : \theta_i^j = h \right\}.$$

Given any $m \in \{0, ..., M\}$ there are $\frac{M!}{m!(M-m)!}$ types of $\theta_i \in \Theta$ such that $\theta_i^j = h$ for exactly m goods. Maintaining the assumption that θ_i is an exchangeable random variable, we therefore have that the probability that an agent has a high valuation for exactly m goods, denoted by β_m, is given by

$$\beta_m = \frac{M!}{m!(M-m)!} \beta_i(\theta_i), \tag{6}$$

where θ_i is any type with m high valuations. This formulation allows the willingness to pay between goods to be correlated. In the simplest case where valuations for different goods are i.i.d., for any θ_i such that $m(\theta_i) = m$, $\beta_i(\theta_i) = \alpha^m (1-\alpha)^{M-m}$ and $\beta_m = \frac{M!}{m!(M-m)!} \alpha^m (1-\alpha)^{M-m}$, where α is the probability that $\theta_i^j = h$.

Appealing to Proposition 1, we only consider mechanisms in the form of (5), implying that it is without loss of generality to consider:[12]

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

[10] Note that the probability of type θ_i, $\beta_i(\theta_i)$ does *not* depend on i. The subscript i is used in β_i so that we can use β to denote the probability of the valuation profile $\theta = (\theta_1, ..., \theta_n)$.

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

[12] Note that the dimensionality of the problem could be reduced further by using Propositions 1 and 2, but it is notationally convenient to impose these additional symmetry restrictions at a later stage. As such, we are indexing η_i^j and t_i by i, even though we know from Propositions 1 and 2 that they do not need to depend on i in the optimal mechanism.

$$\max_{\{\rho,\eta,t\}} \sum_{\theta\in\Theta^n} \beta(\theta) \sum_{j=1}^{M} \rho^j(\theta) \left[\sum_{i=1}^{n} \eta_i^j(\theta_i)\,\theta_i^j - cn\right] \tag{7}$$

$$0 \leq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i^j - t_i(\theta_i) - \left[\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j\left(\theta_{-i},\theta_i'\right)\eta_i^j\left(\theta_i'\right)\theta_i^j - t_i\left(\theta_i'\right)\right]$$

$$\text{for every } \theta_i,\theta_i' \in \Theta\times\Theta \tag{8}$$

$$0 \leq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i},\mathbf{l})\,\eta_i^j(\mathbf{l})\,l - t_i(\mathbf{l}) \tag{9}$$

$$0 \leq \sum_{i=1}^{n}\sum_{\theta_i\in\Theta} \beta_i(\theta_i)\,t_i(\theta_i) - \sum_{\theta\in\Theta^n} \beta(\theta) \sum_{j=1}^{M} \rho^j(\theta)\,cn \tag{10}$$

$$\rho^j(\theta) \in [0,1], \quad \eta_i^j(\theta_i) \in [0,1],$$

where Θ_{-i} denotes the set of possible profiles for all agents except i. In words, the planner maximizes the social surplus subject to the incentive compatibility constraints (8), the participation constraint for type $\mathbf{l} = (l, ..., l)$ (9), and the resource constraint (10).

To simplify our discussions in the analysis below, we use the following terminology:

Definition 4 *An incentive constraint (8) is referred to as:*

1. *a downwards incentive constraint if* $\theta_i^j \geq \theta_i'^j$ *for all* $j \in \mathcal{J}$;
2. *an upwards incentive constraint if* $\theta_i^j \leq \theta_i'^j$ *for all* $j \in \mathcal{J}$;
3. *a diagonal incentive constraint if there exists* $j,k \in \mathcal{J}$ *so that* $\theta_i^j > \theta_i'^j$ *and* $\theta_i^k < \theta_i'^k$.

Incentive constraints that rule out deviations where a single coordinate is misrepresented are henceforth called *adjacent*. If the jth coordinate in θ_i is changed from h to l we write $\theta_i|l_j$ and if the kth coordinate in θ_i is changed from l to h we write $\theta_i|h_k$.

3.1 The Relaxed Optimization Problem

Guided by intuition based on unidimensional mechanism design problems, we will now formulate a *relaxed problem* where all incentive constraints in (8) except the downwards adjacent incentive constraints are removed. The relaxed problem is fully described below in (11). Analogous to the standard approach in unidimensional problems, we will show in Lemma 12 below that this provides a valid solution to the full problem when the solution to the relaxed problem is monotonic in the following sense:

Definition 5 *Mechanism* (ρ,η,t) *is monotonic if* $\eta_i^j(\theta_i) \leq \eta_i^j\left(\theta_i'\right)$ *and* $\rho^j(\theta_{-i},\theta_i) \leq \rho^j\left(\theta_{-i},\theta_i'\right)$ *whenever* $m(\theta_i) \leq m\left(\theta_i'\right)$ *and* $\theta_i^j \leq \theta_i'^j$.

In Lemma 13 below we will provide sufficient conditions in terms of primitives for the solution to the relaxed problem to be monotonic.

The relaxed optimization problem where the full set of incentive constraints is replaced with the downwards adjacent incentive constraints may be written as:

$$\max_{\{\rho,\eta,t\}} \sum_{\theta\in\Theta^n} \beta(\theta) \sum_{j=1}^{M} \rho^j(\theta) \left[\sum_{i=1}^{n} \eta_i^j(\theta_i)\,\theta_i^j - cn\right] \tag{11}$$

$$\text{s.t. } 0 \leq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i^j - t_i(\theta_i) - \left[\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i},\theta_i|l_k)\,\eta_i^j(\theta_i|l_k)\,\theta_i^j - t_i(\theta_i|l_k)\right]$$

$$\text{for every } k \text{ such that } \theta_i^k = h \qquad [m(\theta_i) \text{ constraints for every } i \text{ and } \theta_i] \tag{12}$$

$$0 \leq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i},\mathbf{l})\,\eta_i^j(\mathbf{l})\,l - t_i(\mathbf{l}) \tag{13}$$

$$0 \leq \sum_{i=1}^{n}\sum_{\theta_i\in\Theta} \beta_i(\theta_i)\,t_i(\theta_i) - \sum_{\theta\in\Theta^n} \beta(\theta) \sum_{j=1}^{M} \rho^j(\theta)\,cn \tag{14}$$

$$\rho^j(\theta) \in [0,1], \qquad \eta_i^j(\theta_i) \in [0,1].$$

We will next state a sequence of intermediate results that will be used to characterize the solution to (11) and, eventually, the full problem in (7).

3.2 Relating the Multipliers

First, a standard argument based on compactifying the constraint set assures that there are solutions to the relaxed optimization problem:

Lemma 1 *There exists at least one optimal solution to (11).*

Next, we use the general symmetry result in Proposition 2 to show that the value of a multiplier for any of the downwards adjacent incentive constraint (12) depends only on the number of goods for which the consumer has a high valuation.

Lemma 2 *It is without loss of generality to assume that for every $m \in \{1,...,M\}$ there exists some $\lambda(m) \geq 0$ such that $\lambda(m)$ is the multiplier associated with every constraint (12) applicable for every $\theta_i \in \Theta$ with $m(\theta_i) = m$.*

Hence, $\lambda(m)$ denotes the multiplier for all downwards adjacent incentive constraints for types with $m \geq 1$ high valuations. We also let $\lambda(0)$ denote the participation constraint for type $\mathbf{l} = (l,...,l)$ and let Λ denote the value of the multiplier for the resource constraint (14).

We now show that the multipliers $\lambda(m), m \in \{0,...,M\}$ and Λ are closely linked. Let θ_i be a type for which agent i has a high valuation for $m \in \{0,....,M-1\}$ goods. There are m different ways to replace a coordinate in θ_i that corresponds to a high valuation. Also, there are $M-m$ types of θ_i' that can be "turned into" type θ_i by replacing a single high valuation with a low valuation.

As a result of these observations, we have that the optimality conditions to the program (11) with respect to $t_i(\theta_i)$ are given by

$$-\lambda(m)m+\lambda(m+1)(M-m)+\Lambda\beta_i(\theta_i)=0, \quad (15)$$

whereas the optimality condition for $\theta_i=\mathbf{h}=(h,....,h)$ is

$$\lambda(M)+\Lambda\beta_i(\mathbf{h})=0 \quad (16)$$

Using the identity $\beta_m=\frac{M!}{m!(M-m)!}\beta_i(\theta_i)$, and the difference equation defined by (15) and (16), we obtain:

Lemma 3 *For every $m\in\{0,...,M\}$ the value of $\lambda(m)$ is related to Λ in accordance with*

$$\begin{aligned} m\lambda(m) &= \frac{m!(M-m)!}{M!}\Lambda\sum_{j=m}^{M}\beta_j \quad (17)\\ &= \frac{m!(M-m)!}{M!}\Lambda\Pr[m(\theta_i)\geq m], \end{aligned}$$

where β_j is defined in (6).

A rough intuition for Lemma 3 is that if the designer could extract an extra unit of surplus from types with m high valuations without upsetting any constraints, then an extra unit could be obtained from all higher types as well because only the difference in the transfers are relevant for the (adjacent downwards) incentive constraints. This explains why the value of the multipliers are proportional to the probability that the number of valuations exceeds m.

3.3 Inclusion Rules

Next, we will show that the solution has the intuitively plausible (but non-obvious) property that exclusions are only used for goods for which the agent has low valuations.

Lemma 4 *Suppose that $\theta_i\in\Theta$ such that $\theta_i^j=h$ and that $\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^j(\theta_i,\theta_{-i})>0$. Then, $\eta_i^j(\theta_i)=1$ in any optimal solution to (11).*[13]

At first blush, Lemma 4 may appear like a "no distortion at the top result." However, this is not the case because it says that a consumer should be given access to her high valuation goods *irrespective of* her total number of goods with high valuations. Instead, the result is best understood in terms of the relationship between the multipliers in (15). Providing a high valuation good to a type with m high valuations relaxes the downwards adjacent incentive constraints (12) for every

[13] The condition $\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^j(\theta_i,\theta_{-i})>0$ is needed because the inclusion rule has no effect on either the objective function or the constraints when the conditional probability of provision is zero, but, one may of course set $\eta_i^j(\theta_i)=1$ also in this case.

type with exactly m high valuations. There are m such incentive constraints, so the increased utility from truth-telling for these types is $\lambda(m)\,mh$. However, giving access to the high valuation good j does make it more tempting for a type with $m+1$ high valuations to announce only m high valuations. There are $M-m$ ways to change one coordinate into a lie with m high valuations, so the decreased utility from announcing a type with one of the other coordinates changed from h to l is $-\lambda(m+1)(M-m)\,h$. It is then immediate from (15) that the positive effect from making truth-telling more appealing always dominates, thus all consumers always get access to the high valuation goods.

In order to simplify the discussion of the inclusion rules for good j with $\theta_i^j = l$, it is useful to define

$$G_m(\Phi) = (1-\Phi)(M-m)\,l\beta_m + \Phi\left[\beta_m (M-m)\,l - (h-l)\sum_{j=m+1}^{M}\beta_j\right]. \tag{18}$$

The inclusion rule for low valuation goods is characterized by:

Lemma 5 *Suppose that $\theta_i \in \Theta$ with $m(\theta_i) = m$. Suppose that $\theta_i^j = l$ for some $j \in \mathcal{J}$ and that $\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\,\rho^j(\theta_i,\theta_{-i}) > 0$. Then,*

$$\eta_i^j(\theta_i) = \eta(m) \equiv \begin{cases} 0 & \text{if } G_m(\Phi) < 0 \\ z \in [0,1] & \text{if } G_m(\Phi) = 0 \\ 1 & \text{if } G_m(\Phi) > 0, \end{cases} \tag{19}$$

in any optimal solution to (11) where $\Phi = \frac{\Lambda}{1+\Lambda}$.

To interpret (18) and Lemma 5, we first notice that $(M-m)\,l\beta_m$ is the (positive) effect on *social welfare* if the inclusion rule changes so as to give access to the low valuation goods for all consumers with exactly m high valuations. To understand the second term in (18), consider the expected revenue for the provider if she gives consumers with m or more high valuations full access to all the goods, and the remaining consumers access to only their high valuation goods. Under the above access policy, the provider can charge $mh + (M-m)\,l$ to the group of consumers who consume all goods, and charge h per good to the remaining consumers who only obtain access to their high valuation goods. Assuming that all goods would be provided for sure, the expected revenue under the above access and pricing policy is then be given by

$$\begin{aligned} R(m) &= \Pr[m(\theta_i) \geq m]\,[mh + (M-m)\,l] + h\sum_{j=1}^{m-1}\beta_j j \\ &= \sum_{j=m}^{M}\beta_j\,[mh + (M-m)\,l] + h\sum_{j=1}^{m-1}\beta_j j. \end{aligned} \tag{20}$$

In the absence of a provision decision, a profit maximizer would thus simply pick m to maximize $R(m)$. We note that

$$\begin{aligned} R(m) - R(m+1) &= \left\{ \sum_{j=m}^{M} \beta_j \left[mh + (M-m) l \right] + h \sum_{j=1}^{m-1} \beta_j j \right\} - \left\{ \sum_{j=m+1}^{M} \beta_j \left[(m+1) h + (M-m-1) l \right] + h \sum_{j=1}^{m} \beta_j j \right\} \\ &= \beta_m \left[mh + (M-m) l \right] + \sum_{j=m+1}^{M} \beta_j \left[mh + (M-m) l \right] - \sum_{j=m+1}^{M} \beta_j \left[(m+1) h + (M-m-1) l \right] - h\beta_m m \\ &= \beta_m (M-m) l - (h-l) \sum_{j=m+1}^{M} \beta_j . \end{aligned}$$

Thus the second term in (18) is the (positive or negative) effect on *profits* from allowing a consumer with m high valuations to consume also her low valuation goods. We conclude that we may interpret the constrained welfare optimization as maximizing a *weighted average of welfare and profits*, with weights on profits increasing with the multiplier Λ associated with the resource constraint (14). This property has been shown in unidimensional settings (see Hellwig [10] and Norman [17]), where incentive feasibility can be characterized by a single constraint using methods from Myerson [16], but to the best of our knowledge, no analogous result directly applicable for our multidimensional setting has been shown in the literature.

3.4 Provision Rules

We are now in a position to characterize the optimal provision rule. To do this transparently, it is useful to first denote $H^j(\theta, m)$ as the number of agents with a high valuation for good j and m high valuations in total under the type profile $\theta = (\theta_1, ..., \theta_n)$. Symmetrically, we let $L^j(\theta, m)$ denote the the number of agents with a low valuation for good j and m high valuations in total under type profile θ. The solution to the relaxed problem can now be characterized rather sharply in terms of the multiplier of the resource constraint only:

Lemma 6 *The provision rule for good j in the optimal solution to (11) satisfies*

$$\rho^j(\theta) = \begin{cases} 0 & \text{if } \sum_{m=1}^{M} H^j(\theta, m) h + \sum_{m=0}^{M} L^j(\theta, m) \frac{1}{\beta_m (M-m)} \max\{0, G_m(\Phi)\} - cn < 0 \\ z \in [0,1] & \text{if } \sum_{m=1}^{M} H^j(\theta, m) h + \sum_{m=0}^{M} L^j(\theta, m) \frac{1}{\beta_m (M-m)} \max\{0, G_m(\Phi)\} - cn = 0 \\ 1 & \text{if } \sum_{m=1}^{M} H^j(\theta, m) h + \sum_{m=0}^{M} L^j(\theta, m) \frac{1}{\beta_m (M-m)} \max\{0, G_m(\Phi)\} - cn > 0. \end{cases} \quad (21)$$

Just like the optimal inclusion rule, (21) may be understood as a compromise between welfare maximization and profit maximization. To see this, first note the fact that all high valuation agents consume the good for sure conditional on provision implies that $\sum_{m=1}^{M} H^j(\theta, m) h$ is the surplus from all consumers with a high valuation for good j. Next, recall that $G_m(\Phi)$ had the interpretation as the effect from giving an agent with m high valuations access to *all* the goods for which the agent has a low valuation. However, in order to get the surplus from the low valuation agents in the same units as $\sum_{m=1}^{M} H^j(\theta, m) h$ in (21), we need to rescale $G_m(\Phi)$ by $\frac{1}{\beta_m (M-m)}$. The reason for $\frac{1}{\beta_m}$ is as follows. The inclusion rules involves a trade-off between the possibility that an agent has exactly m high valuations (with probability β_m) and the probability that the number of high valuations exceed m (with probability $\sum_{j=m+1}^{M} \beta_j$). In contrast, the natural unit for the welfare created by providing good j is for a fixed θ, implying that we need to scale up $G_m(\Phi)$ by $\frac{1}{\beta_m}$. In addition, as all goods are treated symmetrically, the marginal effect of changing the inclusion rule is to provide access to $M-m$ goods. In contrast, providing a good only gives the agents a single extra good on the margin. This difference is what accounts for the scaling factor $\frac{1}{M-m}$ in (21). Hence, the first two terms may be thought of as the effect from provision on a combination of social welfare and revenue, whereas the final term obviously is the associated cost of provision.

3.5 Sufficient Conditions for the Solution to the Relaxed Problem (11) To Solve the Full Problem (7)

We will now discuss when the solution to the relaxed problem (11) also solves the full problem (7). Many steps in this analysis are similar to Matthews and Moore [14], but the multidimensional nature of our environment leads to some important differences.

If the mechanism is monotonic in the sense of Definition 5, we show that, the diagonal constraints are irrelevant (Lemma 7) and that all downwards (upwards) constraints are implied by the downwards adjacent (upwards adjacent) constraints (Lemma 8 and 9 respectively):

Lemma 7 *Suppose that (ρ, η, t) is monotonic and that all downwards and upwards incentive constraints hold. Then, all incentive constraints in (7) are satisfied.*

Lemma 8 *Suppose that (ρ, η, t) is monotonic and that all downwards adjacent incentive constraints hold. Then, all downwards constraints in (7) are satisfied.*

Lemma 9 *Suppose that (ρ, η, t) is monotonic and that all upwards adjacent incentive constraints hold. Then, all upwards constraints in (7) are satisfied.*

A qualitative difference in the way downwards and upwards adjacent constraints are dealt with is that the latter may possibly be violated at the optimal solution to (11). However, should the downwards adjacent constraint bind, then it follows that the upwards adjacent constraints are all satisfied:

Lemma 10 *Suppose that (ρ, η, t) is such that the downwards adjacent constraints in (11) bind. Then all upwards adjacent incentive constraints in (7) are satisfied.*

Combing the results above, we have that the solution to the relaxed problem is also a solution to the full problem provided that the downwards adjacent incentive constraints bind and that the solution to the relaxed problem is monotonic. We therefore need conditions for when this is true in the optimal solution to the relaxed problem (11). As a first step we establish that monotonicity guarantees that the constraints bind.

Lemma 11 *Let (ρ, η, t) be an optimal solution to (11). If (ρ, η, t) is monotonic and is **not*** ex post *efficient, then every downwards adjacent incentive constraint and the participation constraint for type $\mathbf{l} = (l, ..., l)$ bind.*

The basic idea behind Lemma 11 is obvious: if there is slack in a constraint, one can increase inclusions or the probability of provision. However, the actual proof is a little bit more subtle because the inclusion and provision probabilities may already be at their upper bound for the particular θ_i where there is slack. To deal with this possibility we instead extract some additional cash from type θ_i. An inductive argument shows that this additional cash can be used to improve the allocation unless the solution is already *ex post* efficient.

We can now combine these preliminary results and provide a characterization of when the solution to the relaxed problem solves the full problem.

Lemma 12 *Let (ρ, η, t) be an optimal solution to (11). If (ρ, η, t) is monotonic and is **not*** ex post *efficient, then (ρ, η, t) is also an optimal solution to the full problem (7).*

Finally, we derive useful sufficient conditions for the conditions in Lemma 12 to be fulfilled:

Lemma 13 *Let (ρ, η, t) be a solution to (11). Then,*

- *(ρ, η, t) is monotonic if $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ is strictly decreasing in m on $\{0, ..., M-1\}$;*
- *Moreover, a sufficient condition for $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ to be strictly decreasing in m on $\{0, ..., M-1\}$ is that the valuations for any goods j and $j' \neq j$ are independent.*

The condition is almost, but not quite, a hazard rate condition, where the term $M - m$ makes the condition diverge from a simple hazard rate condition. Intuitively, the explanation of the "almost hazard rate condition" is that we are either providing access to all goods or only to the goods for which the agent has high valuations. Allowing access to all goods is better in terms of social surplus, but may reduce the revenue raised. More precisely, with probability β_m a consumers has exactly m high valuations, and such a consumer will be willing to pay $(M - m)\, l$ more if he is given access to the low valuation goods. On the other hand, with probability $\sum_{j=m+1}^{M}\beta_j$ the number of high valuations exceed m, and for such a consumer the revenue is reduced by $(h - l)$. In general, $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ may be non-monotonic, because there are no restrictions on the behavior of $\beta_m (M - m)$. Independence is one way to rule out too extreme fluctuations in this term, essentially because it makes the tails of the distribution over m very thin.

If the solution to the relaxed problem (11) is monotonic (which, from Lemma 13, will be the case if $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ is strictly decreasing in m), we can show that the inclusion rule for low valuation goods as characterized in Lemma 5 can take a sharp threshold rule:

Lemma 14 *Let* (ρ, η, t) *be a solution to (11) and suppose that* $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ *is strictly decreasing in* m *on* $\{0, ..., M-1\}$ *. Then, there exists some* $\widetilde{m}$ *such that:*

1. $\eta_i^j(\theta_i) = \eta(m) = 0$ *for every* θ_i *with* $\theta_i^j = l$ *if* $m(\theta_i) < \widetilde{m}$*;*

2. $\eta_i^j(\theta_i) = \eta(m) = 1$ *for every* θ_i *with* $\theta_i^j = l$ *if* $m(\theta_i) > \widetilde{m}$*.*

3.6 The Main Result

Now we provide our main result regarding the limit of the sequences of optimal solutions to (11) as the number of agents n goes out of bounds. We will henceforth index mechanisms by the number of agents n when needed.

The rationale for considering such sequences is that it allows us to *take limits of exact solutions to our optimization problems*, to obtain a more easily interpretable characterization of the solutions when n is large. The key advantage of considering a large population limit is that it allows us to use a version of the "Paradox of Voting" – for large n, it is almost as if the provision rule is constant in type announcements – which tremendously simplifies the description of the optimal mechanism. Formally, we can use the Central Limit Theorem to establish:

Lemma 15 *Let* (ρ_n, η_n, t_n) *be a solution to the constrained welfare problem (11). Then,* $\mathrm{E}\left[\rho_n^j(\theta)\,|\theta_i'\right] - \mathrm{E}\left[\rho_n^j(\theta)\,|\theta_i''\right] \to 0$ *as* $n \to \infty$ *for any* j *and any pair* $\theta_i', \theta_i'' \in \Theta$*.*

Lemma 15 immediately implies that $\mathrm{E}\left[\rho_n^j(\theta)\,|\theta_i'\right] - \mathrm{E}\left[\rho_n^j(\theta)\right] \to 0$. Hence, all conditional probabilities appearing in the incentive constraints may be approximated by the *ex ante* probability of provision, which greatly simplifies the analysis. Our main result is about the provision probabilities and inclusion rules in the limit as n goes to infinity is given as follows:

Proposition 3 *Suppose that* $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ *is strictly decreasing on* $\{0, ..., M-1\}$ *and let* $R(m)$ *be defined as in (20). Then,*

1. $\lim_{n\to\infty} E\rho_n^j(\theta) = 0$ *for every* j *if* $\max_m R(m) - cM < 0$ *in any sequence of feasible solutions to the full problem (7)*;

2. $\lim_{n\to\infty} E\rho_n^j(\theta) = 1$ *for every* j *if* $\max_m R(m) - cM > 0$ *in any sequence of optimal solutions to the full problem (7)*;

3. *When* $\max_m R(m) - cM > 0$, *let* m^* *be the smallest* m *such that* $R(m) - cM > 0$*. Then there exists* $N < \infty$ *such that if* $n \geq N$,

   (a) *for every type* θ_i *with* $m(\theta_i) \geq m^*$, $\eta_n^j(\theta_i) = 1$ *for every good* j*;*

(b) *for every type* θ_i *with* $m(\theta_i) = m^* - 1$, $\eta_n^j(\theta_i) = 1$ *for every good* j *such that* $\theta_i^j = h$, *and*

$$\eta_n^j(\theta_i) \to \frac{R(m^*) - cM}{R(m^*) - R(m^* - 1)}$$

for ever good j *such* $\theta_i^j = l$;

(c) *for every type* θ_i *with* $m(\theta_i) \leq m^* - 2$, $\eta_n^j(\theta_i) = 1$ *for every good* j *such that* $\theta_i^j = h$ *and* $\eta_n^j(\theta_i) = 0$ *for every good* j *such that* $\theta_i^j = l$.

The proof of Proposition 3 requires quite a bit of technical work. Still, the key idea is rather simple. Once we have established that the solution must have a threshold characterization (Lemma 14), and that the influence of any individual agent's announcement vanishes as the number of agents goes to infinity, it is quite clear that we want to set the threshold as low as possible, subject to the constraint that provision must be self-financing. The condition on whether $\max_m R(m) - cM$ is positive thus will dictate whether there will be sufficient budget to provide all the public goods. Most work in the proof goes into establishing that a large economy is approximately like an economy where the provision decisions are made *ex ante*, not conditioning on θ. The intuition for this is that the larger the economy is, the more certain we can be that the economy is close to the expected.

4 Special Cases and Implications

4.1 One Good or Independent Provision

When there are no complementarities in preferences and production costs, it is clear that the analysis of providing a single excludable public good is the same as the analysis of the case where M goods are provided by independent agencies. Also, the case when valuations for different goods are perfectly correlated is also equivalent to the case with a single good.

Let $\alpha = \sum_{m=1}^{M} \frac{m}{M} \beta_m$ denote the (marginal) probability that a consumer has a high willingness to pay for any particular good. With only a single good, the condition that $\frac{1}{M-m} \frac{\sum_{j=m+1}^{M} \beta_j}{\beta_m}$ is decreasing in m is satisfied by default, thus Proposition 3 is always applicable. Furthermore,

$$R(m) = \begin{cases} l & \text{if } m = 0 \\ \alpha h & \text{if } m = 1. \end{cases}$$

Hence, Proposition 3 says that $\lim_{n\to\infty} E\rho_n(\theta) = 0$ if $\max\{l, \alpha h\} - c < 0$; and $\lim_{n\to\infty} E\rho_n(\theta) = 1$ if $\max\{l, \alpha h\} - c < 0$.

The case with $l \geq c$ is trivial: if the low valuation exceeds the cost of provision, it is first best efficient to always provide and never exclude any consumers from usage, which can be implemented by a uniform user fee of c. When $l < c < \alpha h$, we know from Proposition 3 that the probability for access for type l goods converges to

$$\frac{R(1) - c}{R(1) - R(0)} = \frac{\alpha h - c}{\alpha h - l} \in (0, 1) \tag{22}$$

as n goes to infinity. Moreover, $G_1(\Phi)$ as defined in (18) is equal to zero, thus the provision rule characterized in (21) simplifies to

$$\rho^j(\theta) = \begin{cases} 0 & \text{if } H^j(\theta, 1)h - cn < 0 \\ z \in [0,1] & \text{if } H^j(\theta, 1)h - cn = 0 \\ 1 & \text{if } H^j(\theta, 1)h - cn > 0, \end{cases} \tag{23}$$

which interestingly is exactly the same provision rule as the one that would be optimal for a profit-maximizing monopolist; thus for the one good case, the welfare loss associated with a for-profit monopolist relative to the constrained social optimum is only due to over-exclusion, not in under-provision. With more than one good, there is typically under-provision by a for-profit monopolistic provider for finite n.

4.2 Two Goods

Here we describe the results for the two-good case, which was the focus of Fang and Norman [8]. The two-good case is also interesting because we can also characterize the optimal mechanism when the regularity condition is violated (the characterization for that case is provided in the next subsection).

Note that with two goods, $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ is strictly decreasing in m on $\{0,1\}$ if and only if

$$\frac{\beta_2}{\beta_1} < \frac{1}{2}\frac{\beta_1+\beta_2}{\beta_0} = \frac{1}{2}\frac{1-\beta_0}{\beta_0} \iff \frac{\beta_1}{2} > \frac{\beta_0\beta_2}{1-\beta_0}. \tag{24}$$

Furthermore,

$$R(m) = \begin{cases} 2l & \text{for } m = 0 \\ (\beta_1+\beta_2)(h+l) & \text{for } m = 1 \\ 2\alpha h & \text{for } m = 2, \end{cases}$$

where $\alpha = \frac{\beta_1}{2} + \beta_2$. Ruling out the trivial case of $l \geq c$, we see from Proposition 3 that there are three possibilities (see also Proposition 5 in Fang and Norman [8]):

1. If $\max\{(\beta_1+\beta_2)(h+l), 2\alpha h\} < 2c$, then $\lim_{n\to\infty} E\rho_n^j(\theta) = 0$ for $j \in \{1,2\}$;

2. $(\beta_1+\beta_2)(h+l) > 2c$, then $\lim_{n\to\infty} E\rho_n^j(\theta) = 1$ for $j \in \{1,2\}$; all consumers with at least 1 high valuation good get access to both goods, and those with only low valuations get access to each good with probability
$$\frac{R(1)-2c}{R(1)-R(0)} = \frac{(\beta_1+\beta_2)(h+l)-2c}{(\beta_1+\beta_2)(h+l)-2l} \in (0,1);$$

3. $2\alpha h > 2c > (\beta_1+\beta_2)(h+l)$, then $\lim_{n\to\infty} E\rho_n^j(\theta) = 1$ for $j \in \{1,2\}$; all consumers get access to their high valuation goods, and those with 2 low valuations do not get any access at all, but those with one high valuation get access to their low valuation good with probability
$$\frac{R(2)-2c}{R(2)-R(1)} = \frac{2\alpha h - 2c}{2\alpha h - (\beta_1+\beta_2)(h+l)} \in (0,1). \tag{25}$$

It is worth emphasizing that in the optimal *joint* provision mechanism, both goods will be provided with probability one asymptotically if $\max\{(\beta_1+\beta_2)(h+l), 2\alpha h\} > c$, in stark contrast to the optimal *separate* provision mechanism characterized in Subsection 4.1 where a good is provided asymptotically if $\alpha h > c$. It is clear that there exits a non-empty parameter region where $(\beta_1+\beta_2)(h+l) > 2c > 2\alpha h$ where we get asymptotic non-provision if goods are provided separately, but the optimal bundling mechanism provides both goods for sure.

The increased provision probability for efficient public goods under a bundling mechanism relative to the separate provision mechanism is only one channel through which bundling may increase efficiency. The optimal bundling mechanism also leads to strict efficiency gains relative the non-bundling mechanism by increasing the probability of inclusion for low-valuation agents, even in cases when the goods can also be provided without bundling. To see this, suppose that $\alpha h > c$ so that both public goods will be asymptotically provided with probability one with or without bundling. From (22), we know that under the best separate provision mechanism, the probability for access to a low valuation agent is $(\alpha h - c)/(\alpha h - l)$. In contrast, (25) implies that the *ex ante* probability for access conditional on a low valuation for the case where $2c > (\beta_1+\beta_2)(h+l)$ is

$$\underbrace{\frac{\beta_1}{\beta_0+\beta_1}}_{\substack{\text{prob of mixed type}\\ \text{given low valuation}}} \quad \underbrace{\frac{2\alpha h - 2c}{2\alpha h - (\beta_1+\beta_2)(h+l)}}_{\text{from (25)}}. \tag{26}$$

Some algebra shows that (26) is larger than $\frac{\alpha h - c}{\alpha h - l}$ whenever $\frac{\beta_1}{2} > \frac{\beta_0\beta_2}{1-\beta_0}$, which is precisely the condition under which Proposition 3 is applicable. Fewer consumers are thus excluded in the optimal bundling mechanism. A similar calculation applies to the case with $(\beta_1+\beta_2)(h+l) > 2c$.

4.3 The Case with a Binding Monotonicity Constraint

The case with two goods also provides a useful setup for investigating the case when the regularity condition on $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ fails. As shown in (24) this reduces to the condition that $\frac{\beta_1}{2} \le \frac{\beta_0\beta_2}{1-\beta_0}$, which may be interpreted as saying that the valuations are (sufficiently strongly) positively correlated. In the Appendix we prove that the asymptotic characterization for this case is:

Proposition 4 *Assume that* $\frac{\beta_1}{2} \le \frac{\beta_0\beta_2}{1-\beta_0}$ *and* $l < c$. *Then:*

1. $\lim_{n\to\infty} E\rho_n^j(\theta) = 0$ *for every* j *if* $\alpha h < c$ *for any sequence* $\{\rho_n, \eta_n, t_n\}$ *of feasible mechanisms.*

2. $\lim_{n\to\infty} E\rho_n^j(\theta) = 1$ *for every* j *if* $\alpha h > c$ *for any sequence* $\{\rho_n, \eta_n, t_n\}$ *of optimal mechanisms. Moreover, all consumers get access to the high valuation goods and*

$$\eta_n^j(\theta_i) \to \frac{\alpha h - c}{\alpha h - l} \in (0,1)$$

as $n \to \infty$ *for every* θ_i *with* $\theta_i^j = l$.

The interesting aspect of Proposition 4 is that the solution is identical to the case when bundling is not allowed. To understand why, recall that asymptotic provision or non-provision is related to whether the maximal revenue for a monopolistic provider of the goods – if provided – exceeds the costs. The revenue maximizing selling strategy for a monopolist, if both public goods are provided, is either to sell goods separately at price h, or sell the goods as a bundle at price $h+l$, or to charge l for each good. These selling strategies generate a revenue of $2\alpha h$, $(\beta_1+\beta_2)(h+l)$, and $2l$ respectively. Since we are already assuming $l<c$, the question is thus whether $\max\{(\beta_1+\beta_2)(h+l), 2\alpha h\}$ exceeds $2c$. For the first part of Proposition 4, if $\alpha h<c$ and $l<c$ are both satisfied, we have that

$$\begin{aligned}(\beta_1+\beta_2)(h+l) &< (\beta_1+\beta_2)(h+c) < (\beta_1+\beta_2)\left(\frac{1}{\alpha}+1\right)c \\ &= \left[2+\frac{(1-\beta_0)\left(\frac{1}{2}\beta_1+\beta_2\right)-\beta_2}{\frac{1}{2}\beta_1+\beta_2}\right]c<2c\end{aligned}$$

when $\frac{1}{2}\beta_1 \leq \frac{\beta_0\beta_2}{1-\beta_0}$. This calculation shows that it is impossible to provide the goods with probability 1 if $\alpha h<c$, and the idea for why this translates into $E\rho_n^j(\theta)\to 0$ is the same as for the previous analysis.

For the second part of Proposition 4, consider the case when either separate provision or bundling provide sufficient revenue to cover the cost of provision. In the solution characterized by Proposition 4, the asymptotic *ex ante* probability of getting access to low valuation good j is

$$\left(\beta_0+\frac{1}{2}\beta_1\right)\frac{\alpha h-c}{\alpha h-l}.$$

If, instead, the mechanism which is optimal for the case with $\frac{\beta_1}{2}>\frac{\beta_0\beta_2}{1-\beta_0}$ is used, the *ex ante* probability of getting access to low valuation good is

$$\beta_0\frac{(\beta_1+\beta_2)(h+l)-2c}{(\beta_1+\beta_2)(h+l)-2l}.$$

Some algebra along the lines discussed in connection with (26) shows that the *ex ante* probability of getting access and therefore also the social surplus is actually smaller using the bundling mechanism in this case.

5 Conclusion and Discussion

This paper studies the optimal provision mechanism for multiple excludable public goods when agents' valuations are private information. For a parametric class of problems with M goods whose valuations take binary values, we fully characterize the optimal mechanism and demonstrate that it involves bundling if a regularity condition, akin to a hazard rate condition, on the distribution of valuations is satisfied. Bundling alleviates the free riding problem in large economies in two ways:

first, it may increase the asymptotic provision probability of socially efficient public goods from zero to one; second, it decreases the extent of use exclusions. For the case of two goods, we also show that if the regularity condition is violated, then the optimal solution replicates the separate provision outcome.

Our Model as a Positive Theory of Bundling. Bundled discounts are common in many markets. In fact, some have argued that almost any commodity is best viewed as a bundle of characteristics.[14] However, once we take a narrower perspective of only considering bundles of commodities that are viable "stand-alone goods" in the sense that a consumer has a substantial willingness to pay for each component even if no other goods in the bundle are consumed, it is arguable that commonly bundled commodities mostly have non-rival properties, or, more generally, goods with a substantial fixed costs. Software, music, electronic libraries, and TV programming are obvious examples, and these kinds of goods are growing more important. As such, our paper may be interpreted as a positive theory of bundling of goods with no-rival properties.

It is worth noting that in this paper we focused on the harder problem of maximizing social welfare instead of maximizing profits; but the profit-maximizing selling strategy is a by-product of our analysis. It is easy to see that the only change in the asymptotic characterization of the profit-maximizing selling mechanism is that the threshold number of high valuation goods a consumer needs in order to get access to her low valuation goods will increase. Or, put differently, the price of the grand bundle will be higher than in Proposition 3 because a profit-maximizing monopolist will seek as high a profit as possible rather than to only break even. The rest of Proposition 3 is unchanged for the profit-maximization problem.

The key difference between our model of bundling for non-rival goods and the standard bundling model for private goods with constant unit costs is that the decision for whether to provide a good becomes non-trivial. This in turn makes it necessary to go beyond the standard setup with a single consumer. As a result the characterization of the optimal selling mechanism in our setup is more complicated when there is a finite number of consumers.

A Regulatory Benchmark for Bundling of Goods with Large Fixed Costs. Bundling may under some circumstances violate current U.S. anti-trust legislation; indeed the legality and desirability of bundling of software (a non-rival good) have been a critical point of debate in several recent cases. While economists have been involved in this discussion, to the best of our knowledge there is not yet a clear-cut normative benchmark in the economics literature. The analysis in this paper is a small step towards the development of a useful normative benchmark for bundling in natural monopoly situations.

[14] For example, a car can be considered as a "bundle" consisting of a chassis, an engine, cup holders, a stereo system, air conditioning etc.

Though our model is highly stylized, it has enough flexibility to generate a non-trivial trade-off. On the one side, it is a possible that bundling is *required* for the monopolist to break even. In this case the profit-maximizing outcome with bundling is better for the consumer than the welfare maximizing outcome without bundling; thus a requirement to "unbundle" is strictly worse for the consumers, regardless of other budget-balancing remedies that may be combined with the decision to unbundle.[15] On the other side, the fact that bundling makes it easier for the monopolist to extract consumer surplus can make consumers worse off for the obvious reasons.

Our model still makes many restrictive assumptions on the cost and preference side. In particular, while the assumption that preferences exhibit no complementarities highlights the role of bundling as a screening instrument, it is counter-factual in many cases. Moreover, arguments based on alleged complementarities are often used to justify bundling. However, it should be rather clear that in itself, complementarities are not sufficient to provide a rationale for bundling: if many consumers like to consume a particular computer program jointly with their computer, they may choose to do so regardless of whether the computer program is bundled with the computer.[16] Even though the argument that "bundling is good for the customer because the customers like the goods together" is suspect, complementarities in preferences may matter if combined with some other potential justification for bundling, such as the considerations about fixed costs of the current paper. We intend to study this question in future research.

References

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

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

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

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

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

[15] This line of reasoning was an important part of the motivation in the decision by the Office of Fair Trading [18] in the U.K. on alleged anti-competitive mixed bundling by the British Sky Broadcasting Limited.

[16] Indeed, for computers bought on the internet, the buyer already has a long array of options on how to customize the computer, and it would be easy to allow the consumer to opt out of the standard software bundle (a choice which is usually not available).

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

[7] Borgers, Tilman and Peter Norman (2007). "A Note on Budget Balance under Interim Participation Constraints: The Case of Independent Types." Forthcoming, *Economic Theory.*

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

[9] Fang, Hanming and Peter Norman (2006). "To Bundle or Not to Bundle." *RAND Journal of Economics,* Vol. 37, No. 4, 946-963.

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

[11] Hellwig, Martin F. (2005). "A Utilitarian Approach to the Provision and Pricing of Excludable Public Goods." *Journal of Public Economics*, Vol. 89, 1981-2003.

[12] Mailath, George and Andrew Postlewaite (1990). "Asymmetric Information Bargaining Problems with Many Agents." *Review of Economic Studies,* Vol. 57, 351-368.

[13] John O. Ledyard and Thomas R. Palfrey (2007). "A General Characterization of Interim Efficient Mechanisms for Independent Linear Environments." *Journal of Economic Theory,* Vol. 133, No. 1, 441-466.

[14] Matthews, Steven, and John Moore (1987). "Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening." *Econometrica*, Vol. 55, No. 2, 441-467.

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

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

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

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

A Appendix: Proofs of Results in Section 3

Proof of Lemma 1.

Proof. The only variables that are not automatically in a compact set are the transfers. However, $t_i(\mathbf{l}) \leq Ml < Mh$ and $t_i(\theta_i) - t_i(\theta_i|l_k) \leq m(\theta_i)h + [M - m(\theta_i)]l < Mh$. Recursive application of (12) therefore implies that we may bound $t_i(\theta_i)$ from above by M^2h. Since this is also an upper bound for the difference between $t_i(\theta_i)$ and $t_i(\theta_i|l_k)$, it follows from (14) that we may bound $t_i(\theta_i)$ from below by $-M^2h$. Hence, existence of a solution to (11) follows from Weierstrass maximum theorem. ∎

Proof of Lemma 2.

Proof. Let $m(\theta_i) = m\left(\widehat{\theta}_i\right) = m$, let $\lambda_i\left(\theta_i, \theta_i'\right)$ denote the multiplier associated with one of the $m-1$ downwards adjacent constraints for type θ_i and let $\lambda_i\left(\widehat{\theta}_i, \widehat{\theta}_i'\right)$ denote the multiplier associated with one of the $m-1$ downwards adjacent constraints for type $\widehat{\theta}_i$. Proposition 2 ensures that provision and inclusion rules are symmetric and by use of strong duality in linear programming we find that it is without loss to assume that $\lambda_i\left(\theta_i, \theta_i'\right) = \lambda_i\left(\widehat{\theta}_i, \widehat{\theta}_i'\right)$. ∎

Proof of Lemma 3.

Proof. Fix m and let $\theta_i \in \Theta$ with $m(\theta_i) = m$. Consider the incentive compatibility conditions involving θ_i :

$$0 \leq \underbrace{\sum_{\theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta)\, \eta_i^j(\theta_i)\, \theta_i^j - t_i(\theta_i)}_{\text{Term } A} - \underbrace{\sum_{\theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i}, \theta_i|l_k)\, \eta_i^j(\theta_i|l_k)\, \theta_i^j - t_i(\theta_i|l_k)}_{\text{Term } B} \quad \text{(A1)}$$

We note that $t_i(\theta_i)$ enters in Term A for m conditions: there are m ways to manipulate a single coordinate so as to announce a profile with one h replaced with a l. From Lemma *2,* the multiplier associated with each of these conditions is $\lambda(m)$. In addition, $t_i(\theta_i)$ enters in Term B for $M - m$ downwards incentive constraints for types with $m+1$ high valuations. Finally, it also enters the resource constraint (14). The first order condition with respect to $t_i(\theta_i)$ can thus be written as

$$-m\lambda(m) + (M-m)\lambda(m+1) + \Lambda\beta_i(\theta_i) = -m\lambda(m) + (M-m)\lambda(m+1) + \Lambda\frac{m!\,(M-m)!}{M!}\beta_m = 0, \quad \text{(A2)}$$

where we used (6) for the first equality.

For $m = M$, condition (A2) reads $M\lambda(M) = \Lambda\beta_M$, implying that Lemma 3 is also true for $m = M$. Now suppose that $\lambda(m)$ is given by the expression in (17) for some $m \leq M$. The optimality condition (A2) with respect to $t_i\left(\theta_i'\right)$ where $m\left(\theta_i'\right) = m-1$ then reads:

$$\begin{aligned}
0 &= -(m-1)\lambda(m-1) + (M-m+1)\lambda(m) + \frac{(m-1)!\,(M-m+1)!}{M!}\beta_{m-1} \\
&= -(m-1)\lambda(m-1) + \frac{(M-m+1)}{m}m\lambda(m) + \frac{(m-1)!\,(M-m+1)!}{M!}\beta_{m-1} \\
&= -(m-1)\lambda(m-1) + \frac{(M-m+1)}{m}\frac{m!\,(M-m)!}{M!}\Lambda\sum_{j=m}^{M}\beta_j + \frac{(m-1)!\,(M-m+1)!}{M!}\beta_{m-1} \\
&= -(m-1)\lambda(m-1) + \frac{(m-1)!\,(M-m+1)!}{M!}\Lambda\sum_{j=m-1}^{M}\beta_j,
\end{aligned}$$

where the third equality follows from induction hypothesis. Thus, (17) holds also for $m-1$. The result follows from induction. ■

Proof of Lemma 4.

Proof. Let j be some good for which $\theta_i^j = h$. The optimality conditions for problem (11) with respect to $\eta_i^j(\theta_i)$ may be written as

$$\begin{aligned}
\sum_{\theta_{-i}\in\Theta_{-i}}\beta(\theta)\rho^j(\theta)h + \lambda(m)m\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^j(\theta)h & \\
-\lambda(m+1)(M-m)\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^j(\theta)h + \gamma_i^j(\theta_i) - \phi_i^j(\theta_i) &= 0 \\
\gamma_i^j(\theta_i)\eta_i^j(\theta_i) = 0 \text{ and } \phi_i^j(\theta_i)\left(1-\eta_i^j(\theta_i)\right) &= 0,
\end{aligned} \tag{A3}$$

where $\gamma_i^j(\theta_i)$ and $\phi_i^j(\theta_i)$ are respectively the multipliers for the constraints $\eta_i^j(\theta_i) \geq 0$ and $1 - \eta_i^j(\theta_i) \geq 0$. Since we are assuming that $\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^j(\theta_i,\theta_{-i}) > 0$ and since

$$\sum_{\theta_{-i}\in\Theta_{-i}}\beta(\theta)\rho^j(\theta)h = \beta_i(\theta_i)\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^j(\theta)h, \tag{A4}$$

we may simplify the first equality in (A3) and write

$$\beta_i(\theta_i)h + \lambda(m)mh - \lambda(m+1)(M-m)h + \gamma_i^j(\theta_i) - \phi_i^j(\theta_i) = 0, \tag{A5}$$

which, by the use of the complementary slackness conditions in (A3), implies that

$$\eta_i^j(\theta_i) = \begin{cases} 1 & \text{if } \beta_i(\theta_i) + \lambda(m)m - \lambda(m+1)(M-m) > 0 \\ x\in[0,1] & \text{if } \beta_i(\theta_i) + \lambda(m)m - \lambda(m+1)(M-m) = 0 \\ 0 & \text{if } \beta_i(\theta_i) + \lambda(m)m - \lambda(m+1)(M-m) < 0. \end{cases}$$

But, from (15) we have that $\lambda(m)m - \lambda(m+1)(M-m) = \Lambda\beta_i(\theta_i)$, implying that

$$\beta_i(\theta_i) + \lambda(m)m - \lambda(m+1)(M-m) = (1+\Lambda)\beta_i(\theta_i) > 0.$$

■

Proof of Lemma 5.

Proof. Let j be some good for which $\theta_i^j = l$. The optimality conditions for problem (11) with respect to $\eta_i^j(\theta_i)$ are:

$$\sum_{\theta_{-i}\in\Theta_{-i}} \beta(\theta)\rho^j(\theta)l + \lambda(m)m \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\rho^j(\theta)l \tag{A6}$$

$$-\lambda(m+1)\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\rho^j(\theta)\left[(M-m-1)l+h\right] + \gamma_i^j(\theta_i) - \phi_i^j(\theta_i) = 0$$

$$\gamma_i^j(\theta_i)\eta_i^j(\theta_i) = 0 \text{ and } \phi_i^j(\theta_i)\left(1-\eta_i^j(\theta_i)\right) = 0,$$

where $\gamma_i^j(\theta_i)$ and $\phi_i^j(\theta_i)$ are the multipliers for the constraints $\eta_i^j(\theta_i) \geq 0$ and $1-\eta_i^j(\theta_i) \geq 0$ respectively. Using (A4) as in the proof of Lemma 4, we may rearrange the first inequality in (A6) as:

$$\beta_i(\theta_i)l + \lambda(m)ml - \lambda(m+1)\left[(M-m-1)l+h\right] + \frac{\gamma_i^j(\theta_i) - \phi_i^j(\theta_i)}{\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\rho^j(\theta)} = 0. \tag{A7}$$

Using the relations between the multipliers we find that

$$\begin{aligned}
& \beta_i(\theta_i)l + \lambda(m)ml - \lambda(m+1)\left[(M-m-1)l+h\right] \\
/(15)/ \quad = \quad & \beta_i(\theta_i)l(1+\Lambda) - \lambda(m+1)(h-l) \\
/\text{Lemma } 3/ \quad = \quad & \beta_i(\theta_i)l(1+\Lambda) - (h-l)\frac{m!\,(M-(m+1))!}{M!}\Lambda\sum_{j=m+1}^{M}\beta_j \\
/(6)/ \quad = \quad & \beta_m\frac{m!\,(M-m)!}{M!}l(1+\Lambda) - (h-l)\frac{m!\,(M-(m+1))!}{M!}\Lambda\sum_{j=m+1}^{M}\beta_j, \\
= \quad & \frac{m!\,(M-(m+1))!}{M!}\left[\beta_m(M-m)l(1+\Lambda) - (h-l)\Lambda\sum_{j=m+1}^{M}\beta_j\right] \\
= \quad & \frac{m!\,(M-(m+1))!}{M!}(1+\Lambda)G_m(\Phi)
\end{aligned}$$

for $\Phi = \frac{\Lambda}{1+\Lambda}$. Substituting into (A7) and using the complementary slackness conditions in (A6), we have

$$\eta_i^j(\theta_i) = \eta(m) = \begin{cases} 0 & \text{if } G_m(\Phi) < 0 \\ z \in [0,1] & \text{if } G_m(\Phi) = 0 \\ 1 & \text{if } G_m(\Phi) > 0, \end{cases}$$

as asserted. ■

Proof of Lemma 6.

Proof. The optimality conditions for $\rho^j(\theta)$ associated with the problem (11) may (with somewhat ambiguous notation) be written as

$$
\begin{aligned}
&\beta(\theta)\left[\sum_{i=1}^{n}\eta_i^j(\theta_i)\theta_i^j-cn\right]+\sum_{m=0}^{M}\lambda(m)\,m\left[H^j(\theta,m)\beta_{-i}(\theta_{-i})h+L^j(\theta,m)\beta_{-i}(\theta_{-i})\eta(m)l\right]\\
&-\sum_{m=0}^{M-1}\lambda(m+1)\left[H^j(\theta,m)\beta_{-i}(\theta_{-i})(M-m)h+L^j(\theta,m)\beta_{-i}(\theta_{-i})\eta(m)\{(M-m)l+(h-l)\}\right]\\
&-\Lambda\beta(\theta)cn+\gamma^j(\theta)-\phi^j(\theta)\;=\;0,
\end{aligned}
$$

together with the complementary slackness conditions.[17] Recall that $H^j(\theta,m)$ (respectively, $L^j(\theta,m)$) is the number of agents that has a high (respectively, low) valuation for good j and m high valuations in total, whereas $\eta(m)$ denotes the probability that an agent gets access to his low valuation goods when having m high valuations (as characterized in Lemma 5), and $\gamma^j(\theta)$ and $\phi^j(\theta)$ are the multipliers for the boundary constraints. After using (A4) and the fact that

$$
\sum_{i=1}^{n}\eta_i^j(\theta_i)\theta_i^j=\sum_{m=0}^{M}H^j(\theta,m)h+\sum_{m=0}^{M}L^j(\theta,m)\eta(m)l,
$$

we can write the condition as

$$
\begin{aligned}
&\sum_{m=0}^{M}H^j(\theta,m)h+\sum_{m=0}^{M}L^j(\theta,m)\eta(m)l+\sum_{m=0}^{M}\lambda(m)\left[H^j(\theta,m)\frac{m}{\beta_i(\theta_i)}h+L^j(\theta,m)\frac{m}{\beta_i(\theta_i)}\eta(m)l\right]\\
&-\sum_{m=0}^{M-1}\lambda(m+1)\left[H^j(\theta,m)\frac{(M-m)}{\beta_i(\theta_i)}h+L^j(\theta,m)\frac{1}{\beta_i(\theta_i)}\eta(m)((M-m)l+(h-l))\right]\\
&-(1+\Lambda)cn+\frac{\gamma^j(\theta)-\phi^j(\theta)}{\beta(\theta)}\;=\;0.
\end{aligned}
$$

Collecting terms, we get:

$$
\begin{aligned}
&\sum_{m=0}^{M}H^j(\theta,m)\left\{h+\frac{h}{\beta_i(\theta_i)}\left[\lambda(m)m-\lambda(m+1)(M-m)\right]\right\}\\
&\sum_{m=0}^{M}L^j(\theta,m)\left\{\eta(m)l+\frac{\eta(m)}{\beta_i(\theta_i)}\left[\lambda(m)ml-\lambda(m+1)\{(M-m)l+(h-l)\}\right]\right\}\\
&-(1+\Lambda)cn+\frac{\gamma^j(\theta)-\phi^j(\theta)}{\beta(\theta)}\;=\;0.
\end{aligned}
$$

Using the difference equation for the multipliers in (15), which implies that $\lambda(m)m-\lambda(m+1)(M-m)=\Lambda\beta_i(\theta_i)$, we can simplify this further [after using (15)] to:

$$
\begin{aligned}
&\sum_{m=0}^{M}H^j(\theta,m)\{h+\Lambda h\}+\sum_{m=0}^{M}L^j(\theta,m)\left\{\eta(m)l+\eta(m)\left[\Lambda l-\frac{\lambda(m+1)}{\beta_i(\theta_i)}(h-l)\right]\right\}\\
&-(1+\Lambda)cn+\frac{\gamma^j(\theta)-\phi^j(\theta)}{\beta(\theta)}\;=\;0.
\end{aligned}
\tag{A8}
$$

[17] The notation is somewhat unsatisfactory in that $\beta_{-i}(\theta_{-i})$ would be more appropriately denoted by $\beta_{-i}(\theta|\theta_i)$ where θ_i would describe a type that would enter in the particular term (and therefore change with each term). Concretely, $H^j(\theta,m)\beta_{-i}(\theta_{-i})$ would correspond to $\beta_{-i}(\theta|\theta_i)$ for θ_i with $\theta_i^j=h$ and a total of m high valuations.

But using (17) and (6), we have that $\lambda(m+1) = \frac{m!(M-m-1)!}{M!}\Lambda\sum_{j=m+1}^{M}\beta_j$, thus

$$\begin{aligned}
\frac{\lambda(m+1)}{\beta_i(\theta_i)} /(17)/ &= \frac{1}{\beta_i(\theta_i)}\frac{m!(M-m-1)!}{M!}\Lambda\sum_{j=m+1}^{M}\beta_j \\
&= \frac{1}{\beta_i(\theta_i)}\frac{m!(M-m)!}{M!(M-m)}\Lambda\sum_{j=m+1}^{M}\beta_j \\
/(6)/ &= \frac{1}{\beta_m(M-m)}\Lambda\sum_{j=m+1}^{M}\beta_j .
\end{aligned} \tag{A9}$$

Hence,

$$\begin{aligned}
\eta(m)l+\eta(m)\left[\Lambda l-\frac{\lambda(m+1)}{\beta_i(\theta_i)}(h-l)\right] &= \eta(m)l+\eta(m)\left[\Lambda l-\frac{1}{\beta_m(M-m)}\Lambda\sum_{j=m+1}^{M}\beta_j(h-l)\right] \\
&= (1+\Lambda)\eta(m)\left\{l-\underbrace{\frac{\Lambda}{1+\Lambda}}_{=\Phi}\left[\frac{1}{\beta_m(M-m)}\sum_{j=m+1}^{M}\beta_j(h-l)\right]\right\} \\
&= (1+\Lambda)\eta(m)\left\{(1-\Phi)l+\Phi\left[l-\frac{1}{\beta_m(M-m)}\sum_{j=m+1}^{M}\beta_j(h-l)\right]\right\} \\
&= \frac{(1+\Lambda)\eta(m)}{\beta_m(M-m)}\left\{(1-\Phi)\beta_m(M-m)l+\Phi\left[\beta_m(M-m)l-\sum_{j=m+1}^{M}\beta_j(h-l)\right]\right\} \\
/(18)/ &= \frac{(1+\Lambda)}{\beta_m(M-m)}\max\{0,G_m(\Phi)\}.
\end{aligned}$$

Substituting this back into (A8) gives us

$$(1+\Lambda)\left[\sum_{m=0}^{M}H^j(\theta,m)h+\sum_{m=0}^{M}L^j(\theta,m)\left\{\frac{1}{\beta_m(M-m)}\max\{0,G_m(\Phi)\}\right\}-cn\right]+\frac{\gamma^j(\theta)-\phi^j(\theta)}{\beta(\theta)}=0,$$

which, together with the complementary slackness conditions, gives the result. ■

Proof of Lemma 7.

Proof. The solution to (11) violates an incentive constraint in (7) if there exists θ_i,θ_i' such that

$$\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\sum_{j=1}^{M}\rho^j(\theta)\eta_i^j(\theta_i)\theta_i^j-t_i(\theta_i)<\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\sum_{j=1}^{M}\rho^j\left(\theta_{-i},\theta_i'\right)\eta_i^j\left(\theta_i'\right)\theta_i^j-t_i\left(\theta_i'\right). \tag{A10}$$

Since θ_i is exchangeable and costs are identical for all goods, we can apply Proposition 2 (in conjunction with Lemma 4) to conclude that it is without loss of generality to assume that there exists $\left\{\eta(m),P^h(\theta_{-i},m),P^l(\theta_{-i},m),t(m)\right\}_{m=0}^{M}$, such that:

1. $\eta_i^j(\theta_i)=\eta(m)$ for every (θ_i,j) such that $\theta_i^j=l$ for good j and $m(\theta_i)=m$;

2. $\rho^j\left(\theta_{-i},\theta_i'\right)=P^h(\theta_{-i},m)$ for every (θ_i,j) such that $\theta_i^j=h$ for good j and $m(\theta_i)=k$;

3. $\rho^j\left(\theta_{-i},\theta_i'\right) = P^l\left(\theta_{-i},m\right)$ for every $\left(\theta_i,j\right)$ such that $\theta_i^j = l$ for good j and $m\left(\theta_i\right) = k$;

4. $t_i\left(\theta_i\right) = t\left(m\right)$ for every θ_i such that $\theta_i^k = h$ and $m\left(\theta_i\right) = k$.

Consider an arbitrary announcement θ_i' with $m\left(\theta_i'\right) = m'$. Let $r \leq \min\left\{m', m\right\}$ be the number of coordinates such that $\theta_i^j = \theta_i'^j = h$; and let $s \leq \min\left\{M - m', M - m\right\}$ be the number of coordinates such that $\theta_i^j = \theta_i'^j = l$. We can then express the failure of an incentive constraint for the full problem in (A10) as

$$\begin{aligned}
&\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left[P^h\left(\theta_{-i},m\right)mh + P^l\left(\theta_{-i},m\right)\eta\left(m\right)\left(M-m\right)l\right] - t\left(m\right) \\
<\; &\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left\{P^h\left(\theta_{-i},m'\right)\left[rh + \left(m'-r\right)l\right] + P^l\left(\theta_{-i},m'\right)\eta\left(m'\right)\left[sl + \left(M-m'-s\right)h\right]\right\} - t\left(m'\right).
\end{aligned} \tag{A11}$$

We note that if $r < m'$ and $s < M - m'$, then it is possible to announce a type θ_i'' (that differs from θ_i') with $m\left(\theta_i''\right) = m\left(\theta_i'\right) = m'$, but there are $r+1$ coordinates with $\theta_i'^j = \theta_i''^j = h$ and $s+1$ coordinates with $\theta_i'^j = \theta_i''^j = l$. The utility for agent i with type θ_i' from announcing θ_i'' is, using the mechanism described above,

$$\begin{aligned}
&\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left\{P^h\left(\theta_{-i},m'\right)\left[\left(r+1\right)h + \left(m'-r-1\right)l\right] + P^l\left(\theta_{-i},m'\right)\eta\left(m'\right)\left[\left(s+1\right)l + \left(M-m'-s-1\right)h\right]\right\} - t\left(m'\right) \\
=\; &\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left\{P^h\left(\theta_{-i},m'\right)\left[rh + \left(m'-r\right)l\right] + P^l\left(\theta_{-i},m'\right)\eta\left(m'\right)\left[sl + \left(M-m'-s\right)h\right]\right\} - t\left(m'\right) \\
&+ \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left[P^h\left(\theta_{-i},m'\right) - P^l\left(\theta_{-i},m'\right)\eta\left(m'\right)\right]\left(h-l\right) \\
\geq\; &\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left\{P^h\left(\theta_{-i},m'\right)\left[rh + \left(m'-r\right)l\right] + P^l\left(\theta_{-i},m'\right)\eta\left(m'\right)\left[sl + \left(M-m'-s\right)h\right]\right\} - t\left(m'\right)
\end{aligned}$$

where the inequality follows from the assumed monotonicity [i.e. $P^h\left(\theta_{-i},m'\right) \geq P^l\left(\theta_{-i},m'\right)$] and $\eta\left(m'\right) \leq 1$. This is a violation of the upwards incentive constraints, a contradiction to the postulate of the Lemma that all upwards incentive constraints hold. Thus, we conclude that a failure of an incentive constraint implies that either $r = m < m'$ and $s = M - m'$, in which case an upwards incentive constraint fails or $r = m' < m$ and $s = M - m$ in which case a downwards incentive constraint fails. ■

Proof of Lemma 8.

Proof. The proof is by induction. Pick an arbitrary m. Assume that there is some $K < m$ such that a type with m high valuations has no incentives to pretend to be of any type with $k \in \{m-1, ..., K\}$ high valuations. It follows that

$$\begin{aligned}
U\left(m,m\right) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left[P^h\left(\theta_{-i},m\right)mh + P^l\left(\theta_{-i},m\right)\eta\left(m\right)\left(M-m\right)l\right] - t\left(m\right) \\
&\geq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}\left(\theta_{-i}\right)\left[P^h\left(\theta_{-i},k\right)Kh + P^l\left(\theta_{-i},K\right)\eta\left(K\right)\left(M-m\right)l + \left(m-K\right)h\right] - t\left(K\right) \\
&= U\left(m,K\right)
\end{aligned} \tag{A12}$$

is satisfied by hypothesis. By assumption the downwards adjacent incentive constraint for type K holds, implying that

$$\begin{aligned} U(K,K) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},K)Kh + P^l(\theta_{-i},K)\eta(K)(M-K)l\right] - t(K) \\ &\geq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},K-1)(K-1)h + P^l(\theta_{-i},K-1)\eta(K-1)(M-K)l + h\right] - t(K-1) \\ &= U(K,K-1) \end{aligned} \tag{A13}$$

But, the payoff of announcing type $K-1$ for type m is

$$\begin{aligned} &\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},K-1)(K-1)h + P^l(\theta_{-i},K)\eta(K-1)(M-m)l + (m-K+1)h\right] - t(K-1) \\ &= \underbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},K-1)(K-1)h + P^l(\theta_{-i},K)\eta(K-1)(M-K)l + h\right] - t(K-1)}_{\text{RHS in (A13)}} \\ &\quad + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^l(\theta_{-i},K-1)\eta(K-1)(m-K)(h-l) \\ /(\text{A13})/ &\leq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},K)Kh + P^l(\theta_{-i},K)\eta(K)(M-K)l\right] - t(K) \\ &\quad + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^l(\theta_{-i},K-1)\eta(K-1)(m-K)(h-l) \\ &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},K)Kh + P^l(\theta_{-i},K)\eta(K)(M-m)l + (m-K)h\right] - t(K) \\ &\quad - \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^l(\theta_{-i},K)\eta(K)(m-K)(h-l) \\ &\quad + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^l(\theta_{-i},K-1)\eta(K-1)(m-K)(h-l) \\ \Bigg/ \begin{array}{c}(\text{A12}) \text{ and} \\ P^l(\theta_{-i},K-1)\eta(K-1) \\ \leq P^l(\theta_{-i},K)\eta(K)\end{array} \Bigg/ &\leq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},m)mh + P^l(\theta_{-i},m)\eta(m)(M-m)l\right] - t(m), \end{aligned}$$

implying that m has no incentive to mimic $K-1$. By induction it follows that all downwards constraints are satisfied. ∎

<u>*Proof of Lemma 9.*</u>

Proof. The proof is by induction. Let $K > m$ and assume that

$$\begin{aligned} U(m,m) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left\{P^h(\theta_{-i},m)mh + P^l(\theta_{-i},m)\eta(m)(M-m)l\right\} - t(m) \\ &\geq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left\{P^h(\theta_{-i},k)\left[mh + (k-m)l\right] + P^l(\theta_{-i},k)\eta(k)(M-k)l\right\} - t(k) \\ &= U(m,k) \end{aligned}$$

for all $k \in \{m+1,...,K\}$. If $K = M$, all upwards constraint hold by assumption. If $K < M$, the upwards adjacent constraint for type K implies that

$$\begin{aligned} U(K,K) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^h(\theta_{-i},K) Kh + P^l(\theta_{-i},K)\eta(K)(M-K)l - t(K) \qquad \text{(A14)} \\ &\geq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^h(\theta_{-i},K+1)[Kh+l] + P^l(\theta_{-i},K+1)\eta(K+1)(M-K-1)l - t(K+1) \\ &= U(K,K+1). \end{aligned}$$

We then note that

$$\begin{aligned} & U(m,m) - U(m,K+1) \geq U(m,K) - U(m,K+1) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left\{P^h(\theta_{-i},K)[mh+(K-m)l] + P^l(\theta_{-i},K)\eta(K)(M-K)l\right\} - t(K) \\ & - \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^h(\theta_{-i},K+1)[mh+(K+1-m)l] + P^l(\theta_{-i},K+1)\eta(K+1)(M-K-1)l - t(K+1) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left\{P^h(\theta_{-i},K)Kh + P^l(\theta_{-i},K)\eta(K)(M-K)l\right\} - t(K) \\ & + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left\{P^h(\theta_{-i},K)(K-m)(l-h)\right\} \\ & - \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^h(\theta_{-i},K+1)[Kh+l] + P^l(\theta_{-i},K+1)\eta(K+1)(M-K-1)l - t(K+1) \\ & + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^h(\theta_{-i},K+1)(K-m)(h-l) \\ = & \underbrace{U(K,K) - U(K,K+1)}_{\geq 0 \text{ by (A14)}} + \underbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},K+1) - P^h(\theta_{-i},K)\right](K-m)(h-l)}_{\geq 0 \text{ by monotonicity of } P^h}, \end{aligned}$$

implying that type m has no incentive to mis-report as type $K+1$. ■

Proof of Lemma 10.

Proof. Consider types θ_i and θ_i' with $m(\theta_i) = m$ and $m(\theta_i') = m+1$. For ease of notation, define $U(m,m)$ and $U(m+1,m+1)$ as the payoff of truth-telling for type m and $m+1$; and denote the payoff from a type with $m+1$ high valuations to announce a type with m high valuations as $U(m+1,m)$, and the payoff from a type with m high valuations to announce $m+1$ high valuations as $U(m,m+1)$. $U(m+1,m)$ and $U(m,m+1)$ are respectively given as:

$$\begin{aligned} U(m+1,m) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left\{P^h(\theta_{-i},m)mh + P^l(\theta_{-i},m)\eta(m)[(M-m-1)l+h]\right\} - t(m), \\ U(m,m+1) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},m+1)(mh+l) + P^l(\theta_{-i},m+1)\eta(m+1)(M-m-1)l\right] - t(m+1). \end{aligned}$$

We then have that

$$
\begin{aligned}
& U(m,m) - U(m,m+1) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^h(\theta_{-i},m)\, mh + P^l(\theta_{-i},m)\,\eta(m)(M-m)\, l - t(m) \\
& - \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},m+1)[mh+l] + P^l(\theta_{-i},m+1)\,\eta(m+1)(M-m-1)\, l\right] - t(m+1) \\
= & \underbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^h(\theta_{-i},m)\, mh + P^l(\theta_{-i},m)\,\eta(m)\{(M-m-1)\, l + h\} - t(m)}_{=U(m+1,m)} \\
& + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) P^l(\theta_{-i},m)\,\eta(m)(l-h) \\
& \underbrace{- \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},m+1)(m+1)\, h + P^l(\theta_{-i},m+1)\,\eta(m+1)(M-m-1)\, l\right] - t(m+1)}_{=U(m+1,m+1)} \\
& + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[P^h(\theta_{-i},m+1)[h-l]\right] \\
= & \underbrace{U(m+1,m) - U(m+1,m+1)}_{=0 \text{ by Lemma 11}} + \underbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \left\{P^h(\theta_{-i},m+1) - P^l(\theta_{-i},m)\,\eta(m)\right\}[h-l]}_{\geq 0 \text{ as } P^h(\theta_{-i},m+1) - P^l(\theta_{-i},m)\eta(m)\geq 0} \\
\geq & \; 0. \qquad \blacksquare
\end{aligned}
$$

<u>*Proof of Lemma 11.*</u>

Proof. We first prove a useful claim:

Claim A1 *In any solution to (11), $\rho^j(\theta) = 0$ if $\sum_{i=1}^n \eta_i^j(\theta_i)\,\theta_i^j - cn \leq 0$.*

Proof of Claim A1: From Lemma 6, we know that $\rho^j(\theta) > 0$ only if

$$\sum_{m=1}^{M} H^j(\theta,m)\, h + \sum_{m=0}^{M} L^j(\theta,m)\, \frac{1}{\beta_m(M-m)} \max\{0, G_m(\Phi)\} - cn \geq 0.$$

The lemma follows from the fact that

$$\sum_{i=1}^{n} \eta_i^j(\theta_i)\,\theta_i^j - cn = \sum_{m=1}^{M} H^j(\theta,m)\, h + \sum_{m=0}^{M} L^j(\theta,m)\, l - cn,$$

and

$$\frac{1}{\beta_m(M-m)} G_m(\Phi) = l - \Phi\left[(h-l)\sum_{j=m+1}^{M} \beta_j\right] < l. \qquad \blacksquare$$

(Proof of Lemma 11, continued:) Since we assume that (ρ, η, t) is not ex post efficient, it must be the case that $(n-1)h > nc$. Otherwise, the *ex post* optimal mechanism is either to implement if and only if $\theta_i^j = h$ for all i (when $(n-1)h < nc < nh$), or to never provide any good (when

$nh \leq nc$). In each of these two cases, the *ex post* optimal mechanism is trivially implementable under the constraints in (11). We therefore assume that $(n-1)h > nc$ in the remainder of the proof. If a downwards adjacent incentive constraint does not bind, there exists some θ_i such that

$$\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i^j - t_i(\theta_i) > \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i},\theta_i|l_k)\,\eta_i^j(\theta_i|l_k)\,\theta_i^j - t_i(\theta_i|l_k)$$

If $\eta_i^j(\theta_i|l_k) < 1$, we can increase $\eta_i^j(\theta_i|l_k)$ slightly without violating the constraint. Since $(n-1)h > nc$, it follows that $\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\,\rho^j(\theta_{-i},\theta_i|l_k) > 0$ for every j. Hence, the value of the objective function increases, which contradicts the assumption that (ρ,η,t) is an optimum in the first place. Suppose instead that $\eta_i^j(\theta_i|l_k) = 1$. Then, we can find some $\varepsilon > 0$ and raise the tax to $\widetilde{t}_i(\theta_i) = t_i(\theta_i) + \varepsilon$ for θ_i and every θ_i' with $m(\theta_i') = m(\theta_i)$. This raises some extra revenue, implying that the taxes can be lowered slightly for some other types without upsetting the resource constraint. We can therefore proceed inductively as follows. First consider types θ_i' with $m(\theta_i') = m(\theta_i) - 2$. If $\eta_i^j(\theta_i') < 1$ we can increase the surplus by increasing $\eta_i^j(\theta_i')$ for θ_i' with $m(\theta_i') = m(\theta_i) - 2$ and simultaneously decrease $t_i(\theta_i')$ for types with $\theta_i' = m(\theta_i) - 1$ in order to keep the downwards incentive constraint for types with $\theta_i' = m(\theta_i) - 1$ satisfied (as we have slack in resource constraint this can be done). If $\eta_i^j(\theta_i') = 1$ and $m(\theta_i) - 2 \geq 1$ repeat the argument for types θ_i' with $m(\theta_i') = m(\theta_i) - 3$. By Claim A1, it follows that the associated increased access is always desirable from the social welfare viewpoint, so by induction, we conclude that either there exists a feasible allocation with a higher surplus or that $\eta_i^j(\theta_i) = 1$ for all θ_i and j, meaning that there are no active use of exclusions in the solution. However, as we have slack in the resource constraint we can also gain surplus by increasing $\rho^j(\theta)$ for some θ unless the solution to (11) is the *ex post* optimal rule. ■

Proof of Lemma 12.

Proof. By Lemma 7, all "diagonal" incentive constraints hold, and by Lemma 8 all downwards constraints hold. By Lemma 11 we have that all downwards adjacent incentive constraints bind, which, by use of Lemma 10 assures that the upwards adjacent incentive constraints hold. Finally, Lemma 9 guarantees that all upwards incentive constraints hold. Taken together, we have then verified that all incentive constraints in (7) hold if the solution to the relaxed problem is monotonic and different from the first best. ■

Proof of Lemma 13.

Proof. (1). First note that (19) is irrelevant when $m = M$, and it is immediate from (21) that the probability of provision given $(\mathbf{h},\theta_{-i})$ is always weakly higher than under (θ_i,θ_{-i}) for any $\theta_{-i} \in \Theta_{-i}$. Next, from (19) and (21), we know that a sufficient condition for monotonicity is that $\frac{1}{\beta_m(M-m)} \max\{0, G_m(\Phi)\}$ is increasing in m on $\{0, ...M-1\}$. But,

$$\frac{G_m(\Phi)}{\beta_m(M-m)} = l - \Phi\left[\frac{(h-l)}{\beta_m(M-m)} \sum_{j=m+1}^{M} \beta_j\right] \tag{A15}$$

which proves the first part of the Lemma.

(2). If valuations are independent across goods, say the probability that an agent has a high valuation for any good j is α, then the probability that an agent has m high valuations is given by $\beta_m = \frac{M!}{m!(M-m)!}\alpha^m (1-\alpha)^{M-m}$. This implies that

$$\beta_{m+1} = \frac{M!}{(m+1)!(M-m-1)!}\alpha^{m+1}(1-\alpha)^{M-m-1} = \frac{(M-m)\alpha}{(m+1)(1-\alpha)}\beta_m. \tag{A16}$$

Using (A16), we have that

$$\begin{aligned}
\frac{1}{\beta_{m+1}(M-m-1)}\sum_{j=m+2}^{M}\beta_j &= \frac{1}{\beta_{m+1}(M-m-1)}\sum_{j=m+1}^{M-1}\beta_{j+1} \qquad \text{(A17)}\\
&= \frac{(m+1)(1-\alpha)}{(M-m)\alpha(M-m-1)\beta_m}\sum_{j=m+1}^{M-1}\frac{(M-j)\alpha}{(j+1)(1-\alpha)}\beta_j\\
&= \frac{1}{(M-m)\beta_m}\sum_{j=m+1}^{M-1}\frac{(m+1)(M-j)}{(j+1)(M-m-1)}\beta_j\\
/j \geq m+1 \text{ and } \beta_M > 0/ \quad &< \frac{1}{(M-m)\beta_m}\sum_{j=m+1}^{M}\beta_j.
\end{aligned}$$

Since m was arbitrary, we conclude that $\frac{1}{\beta_m(M-m)}\sum_{j=m+1}^{M}\beta_j$ is strictly decreasing in m. ■

Proof of Lemma 14.

Proof. All we need to check is that $G_{m+1}(\Phi) > 0$ is implied by $G_m(\Phi) \geq 0$, and that $G_{m-1}(\Phi) < 0$ is implied by $G_m(\Phi) \leq 0$. Both are immediate from (A15). ■

Proof of Lemma 15.

Proof. To prove this lemma, it turns out to be useful to consider a related auxiliary problem which aims to maximize the average probability of provision (instead of social welfare) under the same constraints as the relaxed problem (11):

$$\max_{\{\rho,\eta,t\}} \sum_{\theta\in\Theta^n}\beta(\theta)\sum_{j=1}^{M}\rho^j(\theta) \tag{A18}$$
$$\text{s.t. } (12),\ (13),\ (14).$$

We first show in Claim A2 below that the characterization of the solution to (A18) is qualitatively similar to the constrained welfare problem (11):

Claim A2 *Let (ρ, η, t) be an optimal solution to (A18) and assume that $\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^j(\theta_i,\theta_{-i}) > 0$. Then,*

1. *conditional on provision, all consumers get access to all their high valuation goods and the inclusion rule for low valuation goods is given by:*

$$\eta_i^j(\theta_i) = \eta(m) \equiv \begin{cases} 0 & \text{if } G_m(1) < 0 \\ z \in [0,1] & \text{if } G_m(1) = 0 \\ 1 & \text{if } G_m(1) > 0; \end{cases} \tag{A19}$$

2. *there exists some $\Lambda \geq 0$ such that the provision rule for good j satisfies*

$$\rho^j(\theta) = \begin{cases} 0 & \text{if } 1 + \Lambda \sum_{m=0}^{M} \left[H^j(\theta,m)\, h + L^j(\theta,m) \frac{\max\{0,G_m(1)\}}{\beta_m(M-m)} - cn \right] < 0 \\ z \in [0,1] & \text{if } 1 + \Lambda \sum_{m=0}^{M} \left[H^j(\theta,m)\, h + L^j(\theta,m) \frac{\max\{0,G_m(1)\}}{\beta_m(M-m)} - cn \right] = 0 \\ 1 & \text{if } 1 + \Lambda \sum_{m=0}^{M} \left[H^j(\theta,m)\, h + L^j(\theta,m) \frac{\max\{0,G_m(1)\}}{\beta_m(M-m)} - cn \right] > 0 \end{cases} \tag{A20}$$

Proof of Claim A2: The derivation of the inclusion rules follow the analysis of the constrained welfare problem step by step and is omitted.

To derive the provision rule, note that the optimality conditions for $\rho^j(\theta)$ associated with the problem (A18) may be written as

$$\begin{aligned} &\beta(\theta) + \sum_{m=0}^{M} \lambda(m)\, m \left[H^j(\theta,m)\, \beta_{-i}(\theta_{-i})\, h + L^j(\theta,m)\, \eta(m)\, l \right] \\ &- \sum_{m=0}^{M-1} \lambda(m+1) \left\{ H^j(\theta,m)\, \beta_{-i}(\theta_{-i})(M-m)\, h + L^j(\theta,m)\, \beta_{-i}(\theta_{-i})\, \eta(m) \left[(M-m)\, l + (h-l) \right] \right\} \\ &- \Lambda \beta(\theta)\, cn + \gamma^j(\theta) - \phi^j(\theta) \;=\; 0, \end{aligned}$$

together with the complementary slackness conditions. By noting that $\frac{\beta_{-i}(\theta_{-i})}{\beta(\theta)} = \frac{1}{\beta_i(\theta_i)}$, we can write the condition as

$$\begin{aligned} &1 + \sum_{m=0}^{M} \lambda(m) \left[H^j(\theta,m) \frac{m}{\beta_i(\theta_i)} h + L^j(\theta,m) \frac{m}{\beta_i(\theta_i)} \eta(m)\, l \right] \\ &- \sum_{m=0}^{M-1} \lambda(m+1) \left\{ H^j(\theta,m) \frac{(M-m)}{\beta_i(\theta_i)} h + L^j(\theta,m) \frac{1}{\beta_i(\theta_i)} \eta(m) \left[(M-m)\, l + (h-l) \right] \right\} \\ &- \Lambda cn + \frac{\gamma^j(\theta) - \phi^j(\theta)}{\beta(\theta)} \;=\; 0. \end{aligned}$$

Collecting terms, we get

$$\begin{aligned} 1 + &\sum_{m=0}^{M} H^j(\theta,m) \left\{ \frac{h}{\beta_i(\theta_i)} \left[\lambda(m)\, m - \lambda(m+1)(M-m) \right] \right\} \\ &+ \sum_{m=0}^{M} L^j(\theta,m) \frac{\eta(m)}{\beta_i(\theta_i)} \left\{ \lambda(m)\, ml - \lambda(m+1) \left[(M-m)\, l + (h-l) \right] \right\} \\ &- \Lambda cn + \frac{\gamma^j(\theta) - \phi^j(\theta)}{\beta(\theta)} \;=\; 0. \end{aligned}$$

Using the difference equation for the multipliers in (15), we can simplify this further to:

$$1+\sum_{m=0}^{M} H^{j}(\theta,m)\,h\Lambda+\sum_{m=0}^{M} L^{j}(\theta,m)\left\{\Lambda\eta(m)\,l-\frac{\eta(m)}{\beta_i(\theta_i)}\lambda(m+1)(h-l)\right\}-\Lambda cn+\frac{\gamma^{j}(\theta)-\phi^{j}(\theta)}{\beta(\theta)} \;=\; 0.$$

Using (A9), we can eliminate $\lambda(m+1)$ and get

$$1+\sum_{m=0}^{M} H^{j}(\theta,m)\,h\Lambda+\sum_{m=0}^{M} L^{j}(\theta,m)\left\{\Lambda\eta(m)\,l-\frac{\eta(m)(h-l)}{\beta_m(M-m)}\Lambda\sum_{j=m+1}^{M}\beta_j\right\}-\Lambda cn+\frac{\gamma^{j}(\theta)-\phi^{j}(\theta)}{\beta(\theta)} \;=\; 0.$$

Furthermore,

$$\Lambda\eta(m)\,l-\frac{\eta(m)(h-l)}{\beta_m(M-m)}\Lambda\sum_{j=m+1}^{M}\beta_j=\frac{\Lambda\eta(m)}{\beta_m(M-m)}\left[\beta_m(M-m)\,l-(h-l)\sum_{j=m+1}^{M}\beta_j\right]=\frac{\Lambda}{\beta_m(M-m)}\max\{0,G_m(1)\},$$

where $G_m(\cdot)$ is defined in (18). Substituting this back into (A8) gives

$$1+\Lambda\sum_{m=0}^{M}\left[H^{j}(\theta,m)\,h+L^{j}(\theta,m)\frac{\max\{0,G_m(1)\}}{\beta_m(M-m)}-cn\right]+\frac{\gamma^{j}(\theta)-\phi^{j}(\theta)}{\beta(\theta)}=0.$$

which, by combining with the complementary slackness conditions, gives the result. ■

Before getting into the details of the proof, it is also useful to observe that (A20) can be rewritten as

$$\rho^{j}(\theta)=\begin{cases}0 & \text{if } 1+\Lambda\left[\sum_{i=1}^{n} Z^{j}(\theta_i)-cn\right]<0\\ z\in[0,1] & \text{if } 1+\Lambda\left[\sum_{i=1}^{n} Z^{j}(\theta_i)-cn\right]=0\\ 1 & \text{if } 1+\Lambda\left[\sum_{i=1}^{n} Z^{j}(\theta_i)-cn\right]>0,\end{cases}$$

where

$$Z^{j}\left(\theta^{i}\right)=\begin{cases}h & \text{if } \theta_i^{j}=h\\ \frac{\max\{0,G_m(1)\}}{\beta_m(M-m)} & \text{if } \theta_i^{j}=l \text{ and } \theta_i^{k}=h \text{ for exactly } m \text{ goods } k\in\mathcal{J}.\end{cases}$$

The point with this formulation is that $\left\{Z^{j}\left(\theta^{i}\right)\right\}_{i=1}^{n}$ is a sequence of i.i.d. random variables. Observe that $Z^{j}\left(\theta^{i}\right)$ is bounded below by 0 and above by h, so the variance is bounded. This allows us to use the central limit theorem to establish a "generalized paradox of voting", which simply states the intuitively obvious fact that the influence of a single agent approaches zero as the number of agents goes out of bounds. To state this, we will now need to be careful about the fact that the solution depends on the number of agents in the economy, and a mechanism with n agents will therefore now be denoted by (ρ_n,η_n,t_n) and the multiplier on the resource constraint will be denoted by Λ_n.

(Proof of Lemma 15, continued): Consider a mechanism that solves (A18) first. As $Z^{j}\left(\theta^{i}\right)\in[0,h]$ for every $\theta^{i}\in\Theta$, it follows that for any pair $\left(\theta_i',\theta_i''\right)$ we have

$$\begin{aligned}\mathrm{E}\left[\rho_n^{j}(\theta)\,|\theta_i'\right]-\mathrm{E}\left[\rho_n^{j}(\theta)\,|\theta_i''\right] \;&\leq\; \Pr\left[1+\Lambda_n\left(h+\sum_{k\neq i} Z^{j}(\theta_k)\right)-cn\geq 0\right]-\Pr\left[1+\Lambda_n\left(0+\sum_{k\neq i} Z^{j}(\theta_k)\right)-cn>0\right]\\ &=\; \Pr\left[\frac{1-cn}{\Lambda_n}-h\leq\sum_{k\neq i} Z^{j}(\theta_k)\leq\frac{1-cn}{\Lambda_n}\right].\end{aligned}$$

Let $\sigma = \text{VAR}\left(Z^j\left(\theta^i\right)\right)$. We can then rewrite the probability statement as

$$\Pr\left[k_n \leq \frac{\sum_{k\neq i} Z^j\left(\theta_k\right) - \text{E}Z^j\left(\theta_i\right)}{\sigma\sqrt{n-1}} \leq k_n + \frac{h}{\sigma\sqrt{n-1}}\right],$$

where $k_n \equiv \frac{1-cn}{\Lambda_n\sigma\sqrt{n-1}} - \frac{h}{\sigma\sqrt{n-1}} - \frac{\text{E}Z^j(\theta_i)}{\sigma\sqrt{n-1}}$. As σ is finite, $\text{E}\left\{Z^j\left(\theta_k\right) - \text{E}Z^j\left(\theta_i\right)\right\} = 0$, and $\left\{Z^j\left(\theta^i\right)\right\}_{i=1}^n$ is an i.i.d. sequence, we know that the central limit theorem is applicable, thus $\frac{\sum_{k\neq i} Z^j(\theta_k) - \text{E}Z^j(\theta_i)}{\sigma\sqrt{n-1}}$ is asymptotically distributed as a standard normal distribution. Moreover, $\frac{h}{\sigma\sqrt{n-1}} \to 0$, which together with the convergence in distribution implies that for every real number k and any $\varepsilon > 0$, there exists $N < \infty$ such that

$$\Pr\left[k \leq \frac{\sum_{k\neq i} Z^j\left(\theta_k\right) - \text{E}Z^j\left(\theta_i\right)}{\sigma\sqrt{n-1}} \leq k + \frac{h}{\sigma\sqrt{n-1}}\right] \leq \frac{1}{\sqrt{2\pi}}\int_k^{k+\varepsilon} \exp\left(\frac{y^2}{2}\right) dy + \varepsilon$$

for $n \geq N$. But, the standard normal is symmetric and single-peaked, so for $n \geq N$,

$$\begin{aligned}\Pr\left[k_n \leq \frac{\sum_{k\neq i} Z^j\left(\theta_k\right) - \text{E}Z^j\left(\theta_i\right)}{\sigma\sqrt{n-1}} \leq k_n + \frac{h}{\sigma\sqrt{n-1}}\right] \quad &\leq \quad \frac{1}{\sqrt{2\pi}}\int_{k_n}^{k_n+\varepsilon} \exp\left(\frac{y^2}{2}\right) dy + \varepsilon \\ &\leq \quad \frac{1}{\sqrt{2\pi}}\int_{-\frac{\varepsilon}{2}}^{\frac{\varepsilon}{2}} \exp\left(\frac{y^2}{2}\right) dy + \varepsilon \to 0\end{aligned}$$

as $\varepsilon \to 0$. The result follows.

For the case in which $\{\rho_n, \eta_n, t_n\}_{n=1}^{\infty}$ is a sequence from the solution to (11), we replace $Z^j\left(\theta_i\right)$ above with

$$Z_n^j\left(\theta_i\right) = \begin{cases} h & \text{if } \theta_i^j = h \\ \frac{\max\{0, G_m(\Phi_n)\}}{\beta_m(M-m)} & \text{if } \theta_i^j = l \text{ and if } \theta_i^k = h \text{ and } m\left(\theta_i\right) = m. \end{cases}$$

By choice of subsequences such that $G_m\left(\Phi_n\right) \to G_m^*$, we can approximate $\frac{\max\{0, G_m(\Phi_n)\}}{\beta_m(M-m)}$ with $\frac{\max\{0, G_m^*\}}{\beta_m(M-m)}$. The rest of the argument follows the one above step by step. ■

Proof of Proposition 3.

Proof. **[PART 1]** For contradiction, let there be a subsequence of feasible mechanisms with $\lim_{n\to\infty} \text{E}\left[\rho_n^j\left(\theta\right)\right] = \rho^*$ and $\lim_{n\to\infty} \eta_n\left(\widetilde{m}\right) = \eta^*$, where $\widetilde{m}$ is the threshold number of high goods in the sense of Lemma 14, which without loss is taken to be constant along the sequence by choice of converging subsequence. We use the following notations below:

- $\text{E}\left[\rho_n^j\left(\theta\right)|m, \theta_i^j = h\right]$ denotes the conditional provision probability for good j perceived by agent i given that $\theta_i^j = h$ and $\theta_i^k = h$ for exactly m goods $k \in J$ (including j). Notice that this probability is the same for all goods by Proposition 2.

- $\text{E}\left[\rho_n^j\left(\theta\right) \mid m, \theta_i^j = l\right]$ denotes the conditional provision probability for good j perceived by agent i given that $\theta_i^j = l$ and $\theta_i^k = h$ for exactly m goods $k \in J$. Again, this probability is the same for all goods by Proposition 2.

Note that Propositions 1 and 2 jointly imply that for every $m \in \{0, .., M\}$, there exists some $t_n(m)$ such that the transfer $t_n(m)$ is paid by every agent i with m high valuations. The utility (from truth-telling) is,

$$\begin{array}{rl} mh\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right] - t_n(m) & \text{if } m < \widetilde{m} \\ \widetilde{m}h\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i^j = h\right] + (M - \widetilde{m})\eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i^j = l\right] l - t_n(\widetilde{m}) & \text{if } m = \widetilde{m} \\ mh\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right] + (M - m)\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = l\right] l - t_n(m) & \text{if } m > \widetilde{m}. \end{array}$$

Hence,

$$t_n(m) \le mh\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right] \text{ for } \theta_i \text{ with } m < \widetilde{m}, \tag{A21}$$

since otherwise these agents would be better off not to participate.[18]. Similarly,

$$t_n(\widetilde{m}) \le \widetilde{m}h\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i^j = h\right] + (M - \widetilde{m})\eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i^j = l\right] l \text{ for } \theta_i \text{ with } m = \widetilde{m}, \tag{A22}$$

again immediately from the participation constraint (or the downwards adjacent incentive constraint combined with (A21)). Using (A22), we find

$$\begin{array}{rrl} & & (\widetilde{m}+1)h\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = h\right] + (M - (\widetilde{m}+1))\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = l\right] l - t_n(\widetilde{m}+1) \\ & \ge & \widetilde{m}h\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = h\right] + \eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = l\right]([M - (\widetilde{m}+1)]\,l + h) - t_n(\widetilde{m}) \\ /(\text{A22})/ & \ge & \eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = l\right]([M - (\widetilde{m}+1)]\,l + h) - (M - \widetilde{m})\eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i^j = l\right] l \\ & = & \eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = l\right](h - l)\,. \end{array}$$

Hence,

$$\begin{array}{rcl} t_n(\widetilde{m}+1) & \le & (\widetilde{m}+1)h\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = h\right] + (M - \widetilde{m} - 1)\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = l\right] l \\ & & -\eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = l\right](h - l)\,. \end{array} \tag{A23}$$

Finally, for every for every $m > \widetilde{m} + 1$, (A23) implies that

$$\begin{array}{rl} & mh\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right] + (M - m)\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = l\right] l - t_n(m) \\ \ge & (\widetilde{m}+1)h\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = h\right] + \mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = l\right][(M - m)\,l + (m - \widetilde{m} - 1)\,h] - t_n(\widetilde{m}+1) \\ \ge & \mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = l\right][(m - \widetilde{m} - 1)(h - l)] + \eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = l\right](h - l)\,, \end{array}$$

hence,

$$\begin{array}{rcl} t_n(m) & \le & mh\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right] + (M - m)\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = l\right] l \\ & & -\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}+1, \theta_i^j = l\right][(m - \widetilde{m} - 1)(h - l)] - \eta_n(\widetilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = l\right](h - l)\,. \end{array} \tag{A24}$$

From from Lemma 15, we have

$$\lim_{n\to\infty}\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right] = \lim_{n\to\infty}\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = l\right] = \lim_{n\to\infty}\mathrm{E}\left[\rho_n^j(\theta)\,|\widetilde{m}, \theta_i'^j = l\right] = \lim_{n\to\infty}\mathrm{E}\left[\rho_n^j(\theta)\right] = \rho^*. \tag{A25}$$

[18]Strictly speaking, we only impose the participation constraint on type **l**. However, the downwards adjacent incentive constraint together with the participation constraint for type **l** imply that all participation constraints hold.

Combining (A25) with (A21), (A22), (A23) and (A24), it follows that, for every $\varepsilon > 0$ there is some N such that, whenever $n \geq N$,

1. $t_n(m) \leq \rho^* mh + \varepsilon$ if $m < \widetilde{m}$
2. $t_n(\widetilde{m}) \leq \rho^* [\widetilde{m}h + (M - \widetilde{m})\eta^* l] + \varepsilon$
3. $t_n(\widetilde{m}+1) \leq \rho^* [(\widetilde{m}+1)h + (M - \widetilde{m} - 1)l - \eta^*(h-l)] + \varepsilon$
4. $t_n(m) \leq \rho^* [(\widetilde{m}+1)h + (M - \widetilde{m} - 1)l - \eta^*(h-l)] + \varepsilon$ if $m > \widetilde{m}+1$.

Summing over m, we find that the expected per capita transfer revenue satisfies:

$$\begin{aligned}
\sum_{m=0}^{M} \beta_m t_n(m) &\leq \rho^* \left[\left(\sum_{m=\widetilde{m}+1}^{M} \beta_m \right) \{(\widetilde{m}+1)h + (M-\widetilde{m}-1)l\} + h \sum_{m=0}^{\widetilde{m}} \beta_m m \right] \\
&\quad + \rho^* \eta^* \left[\beta_{\widetilde{m}} (M-\widetilde{m}) l - \left(\sum_{m=\widetilde{m}+1}^{M} \beta_m \right) (h-l) \right] + \varepsilon \\
&= \rho^* (1-\eta^*) \left[\left(\sum_{m=\widetilde{m}+1}^{M} \beta_m \right) \{(\widetilde{m}+1)h + (M-\widetilde{m}-1)l\} + h \sum_{m=0}^{\widetilde{m}} \beta_m m \right] \\
&\quad + \rho^* \eta^* \left[\beta_{\widetilde{m}} (M-\widetilde{m}) l + \left(\sum_{m=\widetilde{m}+1}^{M} \beta_m \right) (\widetilde{m}h + (M-\widetilde{m})l) + h \sum_{m=0}^{\widetilde{m}} \beta_m m \right] + \varepsilon \\
&= \rho^* (1-\eta^*) \left[\left(\sum_{m=\widetilde{m}+1}^{M} \beta_m \right) \{(\widetilde{m}+1)h + (M-\widetilde{m}-1)l\} + h \sum_{m=0}^{\widetilde{m}} \beta_m m \right] \\
&\quad + \rho^* \eta^* \left[\left(\sum_{m=\widetilde{m}}^{M} \beta_m \right) (\widetilde{m}h + (M-\widetilde{m})l) + h \sum_{m=0}^{\widetilde{m}-1} \beta_m m \right] + \varepsilon \\
&= \rho^* \{(1-\eta^*) R(\widetilde{m}+1) + \eta^* R(\widetilde{m})\} + \varepsilon.
\end{aligned} \tag{A26}$$

Moreover, for any ε, we can find N such that $M \sum_{\theta \in \Theta^n} \beta(\theta) \rho_n^j(\theta) c \geq \rho^* Mc - \varepsilon$, which in conjunction with (A26) means that

$$\begin{aligned}
\sum_{i=1}^{n} \sum_{\theta_i \in \Theta} \beta_i(\theta_i) t_{in}(\theta_i) - \sum_{\theta \in \Theta^n} \beta(\theta) \sum_{j=1}^{M} \rho_n^j(\theta) cn &= n \sum_m \beta_m t_n(m) - M \sum_{\theta \in \Theta^n} \beta(\theta) \rho_n^j(\theta) c \\
&\leq \rho^* \underbrace{\{(1-\eta^*) R(\widetilde{m}+1) + \eta^* R(\widetilde{m}) - Mc\}}_{<0 \text{ by hypothesis}} + 2\varepsilon.
\end{aligned} \tag{A27}$$

We conclude that corresponding to any $\rho^* > 0$, there exists $\varepsilon > 0$ such that the right hand side of (A27) is negative. Hence, the budget constraint (14) is violated for large n along any sequence with non-vanishing provision probability. It thus follows that $\lim_{n \to \infty} \mathrm{E}\left[\rho_n^j(\theta)\right] = 0$ for any convergent subsequence.

[PARTS 2 and 3] Consider the sequence of mechanisms $\left(\widehat{\rho}_n, \widehat{\eta}_n, \widehat{t}_n\right)$ where $\widehat{\rho}_n^j(\theta) = 1$ for every n and θ, where all agents get to consume all high valuation goods and the inclusion rule for low valuation goods is

$$\widehat{\eta}_n(m) = \begin{cases} 0 & \text{if } m < m^* \\ \frac{R(m^*+1)-cM}{R(m^*+1)-R(m^*)} & \text{if } m = m^* \\ 1 & \text{if } m > m^*, \end{cases}$$

and where the transfer tule is

$$\widehat{t}_n(m) = \begin{cases} mh & \text{if } m < m^* \\ m^*h + (M - m^*)\,\eta l & \text{if } m = m^* \\ (m^*+1)\,h + (M - m^* - 1)l - \eta^*(h - l) & \text{if } m > m^*, \end{cases}$$

for

$$\eta^* = \frac{R(m^*+1) - cM}{R(m^*+1) - R(m^*)}. \tag{A28}$$

It is easy to check that truth-telling is incentive compatible and individually rational. The per capita expected transfer under the above mechanism is

$$\begin{aligned} \sum_{m=0}^{M} \beta_m \widehat{t}_n(m) &= h \sum_{m=0}^{m^*-1} \beta_m m + \beta_{m^*}\left[m^*h + (M - m^*)\,\eta^* l\right] \qquad \text{(A29)} \\ &\quad + \left(\sum_{m=m^*+1}^{M} \beta_m\right)\left[(m^*+1)\,h + (M - m^* - 1)l - \eta^*(h - l)\right] \\ &= (1-\eta^*)\left\{h \sum_{m=0}^{m^*} \beta_m m + \left(\sum_{m=m^*+1}^{M} \beta_m\right)\left[(m^*+1)\,h + (M - m^* - 1)l\right]\right\} \\ &\quad + \eta^*\left[h \sum_{m=0}^{m^*-1} \beta_m m + \left(\sum_{m=m^*}^{M} \beta_m\right)\left[m^*h + (M - m^*)l\right]\right] \\ &= (1-\eta^*)\,R(m^*+1) + \eta^* R(m^*) \\ /(\text{A28})/ \quad &= \left(\frac{cM - R(m^*)}{R(m^*+1) - R(m^*)}\right) R(m^*+1) + \frac{R(m^*+1) - cM}{R(m^*+1) - R(m^*)} R(m^*) = cM. \end{aligned}$$

The feasibility constraint thus holds exactly. Thus we conclude that $\left(\widehat{\rho}_n, \widehat{\eta}_n, \widehat{t}_n\right)$ is incentive feasible for every n. The associated social surplus with the above mechanism is

$$S_n^*(\theta) = \sum_{j=1}^{M}\left\{h \sum_{m=1}^{M} H^j(\theta, m) + l\left[\eta^* L^j(\theta, m^*) + \sum_{m=m^*+1}^{M} L^j(\theta, m)\right]\right\} - cM, \tag{A30}$$

where the notations $H^j(\theta, m)$ and $L^j(\theta, m)$ were explained earlier. By Lemma 14, we know that the surplus maximizing mechanism is also characterized by some threshold m^{**} and some inclusion probability given $m = m^{**}$ denoted by η^*, so that the associated social surplus is given by

$$S_n^{**}(\theta) = \sum_{j=1}^{M} \rho^j(\theta)\left[h \sum_{m=1}^{M} H^j(\theta, m) + l\left\{\eta^{**} L^j(\theta, m^{**}) + \sum_{m=m^{**}+1}^{M} L^j(\theta, m)\right\} - c\right]. \tag{A31}$$

Notice that $\frac{H^j(\theta,m)}{n}$ converges in probability to $\frac{m}{M}\beta_m$ and $\frac{L^j(\theta,m)}{n}$ converges in probability to $\frac{M-m}{M}\beta_m$. These immediately imply that

$$\lim_{n\to\infty} \Pr\left[\left|\frac{S_n^*(\theta)}{n} - h \sum_{m=0}^{M} m\beta_m - l\left[\eta^*(M - m^*)\,\beta_{m^*} + \sum_{m=m^*+1}^{M} (M - m)\,\beta_m\right] + cM\right| > \varepsilon\right] = 0 \tag{A32}$$

for any $\varepsilon > 0$. Next, note that

$$\sum_{m=1}^{M} H^j(\theta, m)\, h + \sum_{m=0}^{M} L^j(\theta, m) \frac{1}{\beta_m (M-m)} \max\{0, G_m(\Phi)\} - cn \tag{A33}$$

converges in probability to

$$\begin{aligned}
&\sum_{m=1}^{M} \frac{m}{M}\beta_m h + \sum_{m=0}^{M} \frac{M-m}{M}\beta_m \frac{1}{\beta_m (M-m)} \max\{0, G_m(\Phi)\} - cn \qquad \text{(A34)}\\
=\;& \sum_{m=0}^{M} \left\{ \frac{m}{M}\beta_m h + \max\{0, G_m(\Phi)\} \right\} - cn\\
=\;& \sum_{m=1}^{M} \frac{m}{M}\beta_m h + \sum_{m=m^{**}+1}^{M} (1-\Phi)(M-m)\, l\beta_m + \Phi\left[\beta_m (M-m)\, l - (h-l)\sum_{j=m+1}^{M}\beta_j\right] - cn.
\end{aligned}$$

Recalling the form of the surplus maximizing provision rule in (21), we see that good j is provided is and only if the expression in (A33) is positive. Now:

1. Suppose that $m^{**} > m^{*}$. Then, $R(m^{**}) - cM < 0$ and an argument identical with part 1 establishes that $\lim_{n\to\infty} \mathrm{E}\left[\rho_n^j(\theta)\right] \to 0$.
2. Suppose that $m^{**} < m^{*}$ and $R(m^{**}) - cM < 0$. Then an argument identical with part 1 establishes that $\lim_{n\to\infty} \mathrm{E}\left[\rho_n^j(\theta)\right] \to 0$.
3. Suppose that $m^{**} < m^{*}$ and $R(m^{**}) - cM \geq 0$. Then, the expression in (A34) is strictly positive and $\lim_{n\to\infty} \mathrm{E}\left[\rho_n^j(\theta)\right] \to 1$. The social surplus must therefore converge towards the same expression as in (A32), but with m^{*} replaced by m^{**}. But, if $m^{**} < m^{*}$ surplus is smaller because there are more exclusions.
4. Suppose that $m^{**} = m^{*}$. If $\widehat{\eta}_n(m^{*}) \to \eta^{**} < \eta^{*}$ an argument identical with part 1 establishes that $\lim_{n\to\infty} \mathrm{E}\left[\rho_n^j(\theta)\right] \to 0$. If $\widehat{\eta}_n(m^{*}) \to \eta^{**} > \eta^{*}$ we have that $\lim_{n\to\infty} \mathrm{E}\left[\rho_n^j(\theta)\right] \to 1$ and, again, social surplus converges towards the same expression as in (A32), but with η^{*} replaced by η^{**}. Again, surplus is smaller if $\eta^{**} > \eta^{*}$ since exclusions are higher. ■

<u>*Proof of Proposition 4.*</u>

Proof. To prove Proposition 4, we add the (non-adjacent) constraint that the highest type should not have an incentive to pretend to be the lowest type:

$$0 \leq \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i}, \mathbf{h})\, h - t_i(\mathbf{h}) - \left[\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i}, \mathbf{l})\, \eta_i^j(\mathbf{l})\, h - t_i(\mathbf{l})\right] \tag{A35}$$

to program (11). Let ψ denote the multiplier on constraint (A35). The optimality conditions with respect to t are then

$$\begin{aligned}
-2\lambda(2) - \psi + \Lambda\beta_2 &= 0\\
\lambda(2) - \lambda(1) + \Lambda\frac{1}{2}\beta_1 &= 0\\
2\lambda(1) + \psi - \lambda(0) + \Lambda\beta_0 &= 0,
\end{aligned}$$

which are obtained just like (A2) by considering types with 2, 1 and 0 high types respectively in the special case with two high valuations and the additional constraint (A35). It follows immediately that:

Lemma A1 *Consider $m = 2$. Then, there are three possibilities in the solution to the program where (A35) has been added to (11):*

1. *The only binding constraints are the adjacent constraints in (11). In this case the solution is unchanged relative to program (11) and $\lambda(2) = \frac{1}{2}\Lambda\beta_2$, $\lambda(1) = \frac{1}{2}\Lambda(\beta_1 + \beta_2)$ and $\lambda(0) = \Lambda$*
2. *The only binding constraints are (A35) and the downwards adjacent constraints for types with $m = 1$. In this case $\psi = \Lambda\beta_2, \lambda(1) = \Lambda\frac{1}{2}\beta_1$ and $\lambda(0) = \Lambda$*
3. *All constraints bind, in which case $2\lambda(2) - \psi = \Lambda\beta_2$, $\lambda(1) \in \left[\Lambda\frac{1}{2}\beta_1, \Lambda\frac{1}{2}(\beta_1 + \beta_2)\right]$ and $\lambda(0) = \Lambda$.*

We now consider these possibilities in turn:

Claim A3 *Suppose that $\frac{\beta_1}{2} < \frac{\beta_0\beta_2}{1-\beta_0}$. Then (A35) binds.*

Proof of Claim A3: If (A35) does not bind, Lemma A1 implies that the solution is unchanged from (11). Simplifying the inclusion rule (19) for the special case with $M = 2$, we have that

$$\eta(1) \equiv \begin{cases} 0 & \text{if } G_1(\Phi) < 0 \\ z \in [0,1] & \text{if } G_1(\Phi) = 0 \\ 1 & \text{if } G_1(\Phi) > 0, \end{cases}$$

$$\eta(0) \equiv \begin{cases} 0 & \text{if } G_0(\Phi) < 0 \\ z \in [0,1] & \text{if } G_0(\Phi) = 0 \\ 1 & \text{if } G_0(\Phi) > 0, \end{cases}$$

where

$$\begin{aligned} G_1(\Phi) &= \beta_1 l - \Phi(h-l)\beta_2 \\ G_0(\Phi) &= \beta_0 2l - \Phi(h-l)(\beta_1 + \beta_2) \end{aligned}$$

and the provision rule (21) simplifies to

$$\rho^j(\theta) = \begin{cases} 0 & \text{if } \left[H^j(\theta,1) + H^j(\theta,2)\right]h + \frac{L^j(\theta,0)}{2\beta_0}\max\{0, G_0(\Phi)\} + \frac{L^j(\theta,1)}{\beta_1}\max\{0, G_1(\Phi)\} - cn < 0 \\ z \in [0,1] & \text{if } \left[H^j(\theta,1) + H^j(\theta,2)\right]h + \frac{L^j(\theta,0)}{2\beta_0}\max\{0, G_0(\Phi)\} + \frac{L^j(\theta,1)}{\beta_1}\max\{0, G_1(\Phi)\} - cn = 0 \\ 1 & \text{if } \left[H^j(\theta,1) + H^j(\theta,2)\right]h + \frac{L^j(\theta,0)}{2\beta_0}\max\{0, G_0(\Phi)\} + \frac{L^j(\theta,1)}{\beta_1}\max\{0, G_1(\Phi)\} - cn > 0. \end{cases}$$

Notice that

$$\begin{aligned} \frac{G_1(\Phi)}{\beta_1} &= l - \frac{\Phi(h-l)\beta_2}{\beta_1} < l - \frac{\Phi(h-l)\beta_2}{2\frac{\beta_0\beta_2}{1-\beta_0}} \\ &= l - \frac{\Phi(h-l)(\beta_1+\beta_2)}{2\beta_0} = \frac{G_0(\Phi)}{2\beta_0}, \end{aligned}$$

implying that $\eta(1) \leq \eta(0)$ and $\rho^j(\theta_{-i}, \theta_i) \leq \rho^j(\theta_{-i}, ll)$ if $\theta_i^j = l$ and $\theta_i^k = h$ for $k \neq j$. Let $t(0), t(1)$ and $t(2)$ denote transfers from types with zero, one and two high valuations respectively. Because the downwards adjacent incentive constraint for type hh and the types with one high valuation are binding, it follows that

$$\begin{aligned}
&\underbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, hh) + \rho^2(\theta_{-1}, hh)\right] h - t(2)}_{\text{utility from truth-telling for } hh} \\
\text{/binding DAIC for } hh\text{/} \quad = & \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, hl)\, h + \rho^2(\theta_{-1}, hl)\,\eta(1)\, h\right] - t(1) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, hl)\, h + \rho^2(\theta_{-1}, hl)\,\eta(1)\, l\right] - t(1) \\
& + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\,\rho^2(\theta_{-1}, hl)\,\eta(1)(h-l) \\
\text{/binding DAIC for } hl\text{/} \quad = & \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, ll)\,\eta(0)\, h + \rho^2(\theta_{-1}, ll)\,\eta(0)\, l\right] - t(0) \\
& + \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\,\rho^2(\theta_{-1}, hl)\,\eta(1)(h-l) \\
= & \underbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, ll)\,\eta(0)\, h + \rho^2(\theta_{-1}, ll)\,\eta(0)\, h\right] - t(0)}_{\text{utility if } hh \text{ announces } ll} \\
& \underbrace{+ \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^2(\theta_{-1}, hl)\,\eta(1) - \rho^2(\theta_{-1}, ll)\,\eta(0)\right](h-l),}_{\leq 0 \text{ as } \eta(1)\leq\eta(0) \text{ and } \rho^j(\theta_{-i},hl)\leq\rho^j(\theta_{-i},ll) \text{ for every } \theta_{-i}}
\end{aligned}$$

implying that (A35) binds or is violated, a contradiction. ■

Claim A4 *Suppose that* $\frac{\beta_1}{2} < \frac{\beta_0\beta_2}{1-\beta_0}$. *Then the downwards adjacent constraints for type hh binds.*

Proof of Claim A4: For contradiction, assume that

$$\sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, hh) + \rho^2(\theta_{-1}, hh)\right] h - t(2) > \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, hl)\, h + \rho^2(\theta_{-1}, hl)\,\eta(1)\, h\right] - t(1). \tag{A36}$$

With slack in the downwards adjacent constraints for hh and a binding (A35) the optimality conditions with respect to $\eta_i^2(hl)$ my be written as

$$\sum_{\theta_{-i}\in\Theta_{-i}} \beta(\theta_{-i}, hl)\,\rho^2(\theta_{-i}, hl)\, l + \lambda(1) \sum_{\theta_{-i}\in\Theta_{-i}} \beta_{-i}(\theta_{-i})\,\rho^2(\theta_{-i}, hl)\, l + \gamma_i^2(hl) - \phi_i^2(hl) = 0.$$

This implies that $\eta_i^2(hl) = \eta_i^1(lh) = \eta(1) = 1$ because a small increase in the low valuation usage probability for the types with one high valuation has no effect on incentives for hh if the downwards adjacent constraints for type hh are nonbinding. After a simplification (that removes the awkwardness of the notation mentioned in footnote 17), the optimality conditions with respect to $\rho^1(\theta)$ are given by the complementary slackness conditions together with

$$\begin{aligned}
&\left[h\left\{H^1(\theta,1)+H^1(\theta,2)\right\}+l\left\{L^1(\theta,1)+\eta(0)L^1(\theta,0)\right\}-cn\right] \\
&+\lambda(1)\frac{2}{\beta_1}\left[H^1(\theta,1)h+L^1(\theta,1)\eta(0)l\right]+\psi\frac{1}{\beta_2}H^1(\theta,2)h \\
&-\psi\frac{1}{\beta_0}\beta_{-i}(\theta_{-i})\eta(0)H^1(\theta,2)-\lambda(1)\frac{\eta(0)}{\beta_0}\left[H^1(\theta,1)h+L^1(\theta,1)l\right] \\
&+\lambda(0)\frac{1}{\beta_0}L^1(\theta,0)l-\Lambda cn+\frac{\gamma^1(\theta)-\phi^1(\theta)}{\beta(\theta)} = 0,
\end{aligned} \tag{A37}$$

where $\gamma^1(\theta)$ is for the constraint the multiplier $\rho^1(\theta)\geq 0$ and $\phi^1(\theta)$ is the multiplier for the constraint $1-\rho^1(\theta)\geq 0$. It is easy to see from (A37) that $\rho^1(\theta_{-i},lh)\geq\rho^1(\theta_{-i},ll)$ for every θ_{-i} as increasing $L^1(\theta,1)$ by a unit and decreasing $L^1(\theta,1)$ by a unit. Together with the fact that $\eta(1)\geq\eta(0)$, it follows that

$$\sum_{\theta_{-i}\in\Theta_{-i}}\beta(\theta_{-i})\rho^1(\theta_{-i},lh)\eta(1)\geq\sum_{\theta_{-i}\in\Theta_{-i}}\beta(\theta_{-i})\rho^1(\theta_{-i},ll)\eta(0),$$

which means that the perceived probability to consume a low valuation good is weakly higher for a consumer with a high valuation for the other good. Hence,

$$\begin{aligned}
&\underbrace{\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},hh)+\rho^2(\theta_{-1},hh)\right]h-t(2)}_{\text{utility from truth-telling for } hh} \\
\text{/(A35)binds/}\quad &=\underbrace{\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},ll)\eta(0)h+\rho^2(\theta_{-1},ll)\eta(0)h\right]-t(0)}_{\text{utility if } hh \text{ announces } ll} \\
&=\underbrace{\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},ll)\eta(0)l+\rho^2(\theta_{-1},ll)\eta(0)h\right]-t(0)}_{\text{utility if } lh \text{ announces } ll} \\
&+(h-l)\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^1(\theta_{-i},ll)\eta(0) \\
\text{/binding DAIC for } lh\text{/}\quad &=\underbrace{\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},lh)\eta(1)l+\rho^2(\theta_{-1},lh)h\right]-t(1)}_{\text{utility from truth-telling for } lh} \\
&+(h-l)\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^1(\theta_{-i},ll)\eta(0) \\
&=\underbrace{\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},lh)\eta(1)h+\rho^2(\theta_{-1},lh)h\right]-t(1)}_{\text{utility if } hh \text{ announces } lh} \\
&\underbrace{+(h-l)\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^1(\theta_{-i},ll)\eta(0)-(h-l)\sum_{\theta_{-i}\in\Theta_{-i}}\beta_{-i}(\theta_{-i})\rho^1(\theta_{-i},lh)\eta(1)}_{\leq 0},
\end{aligned}$$

which contradicts (A36). ■

(Proof of Proposition 4, continued): We conclude that all constraints must bind. This immediately implies that $\eta(1)=\eta(0)$. Assuming that $l<c<\alpha h$ it is easy to check that $0<\eta(1)=\eta(0)<1$

for sufficiently large n, which can only hold if $G_1(\Phi) = G_0(\Phi) = 0$. That means that the provision rule for good j becomes

$$\rho^j(\theta) = \begin{cases} 0 & \text{if } \left[H^j(\theta,1) + H^j(\theta,2)\right]h - cn < 0 \\ z \in [0,1] & \text{if } \left[H^j(\theta,1) + H^j(\theta,2)\right]h - cn = 0 \\ 1 & \text{if } \left[H^j(\theta,1) + H^j(\theta,2)\right]h - cn > 0, \end{cases}$$

which coincides with the provision rule if good j is provided as a stand alone good. Since the inclusion rule is also the same as that when only good is provided, the result follows. ■