Optimal Provision of Multiple Excludable Public Goods*

Hanming Fang†     Peter Norman‡

First Version: October 2002
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 characterize the optimal mechanism, and show that it involves bundling. 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.

Keywords: Public Goods Provision; Bundling; Exclusion

JEL Classification Number: H41.

*This paper generalizes and supercedes Fang and Norman [12], 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, hanming.fang@duke.edu.
‡Department of Economics, UNC-Chapel Hill, 107 Gardner Hall, Chapel Hill, NC 27599-3305, normanp@email.unc.edu.
1 Introduction

This paper studies the optimal provision mechanism for a natural monopolist selling 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, which is akin to a hazard rate condition, is satisfied. If the regularity condition is violated, which happens if 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 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, which in turn depends on whether or not a regularity condition, akin to a hazard rate condition, on the distribution of valuations is satisfied. Under the regularity condition, the optimal mechanism also prescribes a very simple rule for uses access to the public goods if 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 only be given access to 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 then 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. We find it interesting that this relates 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\footnote{See, e.g., McAfee, McMillan and Whinston [20] and Fang and Norman [13].}. Our result here shows that this logic falls apart when valuations are too positively correlated.

It is also worth commenting that the analytical tractability for the multidimensional mechanism design from 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 [15] and Norman [23]). 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.

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. 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 and Appendix B contains the proofs for more technical proofs of results in Section 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 setting up the mechanism design problem (Section 2.3). We then gradually show, sometimes with additional restrictions on the environment, that it is without loss of generality to consider a smaller and more tractable class of simple, anonymous and symmetric mechanisms (Sections 2.4 and 2.5). The main results of this section are Propositions 1 and 2, which are used in later Sections to reduce the dimensionality of the design problem.

2.1 The Environment

There are $M$ excludable public goods, labeled by $j \in \mathcal{J} = \{1, ..., M\}$ and $n$ consumers, indexed by $i \in \mathcal{I} = \{1, ..., n\}$. Each public good is indivisible, and the cost of providing good $j$, denoted $C^j(n)$, is independent of which of the other goods are provided. Since $n$ is the size of the economy and not the number of users, all goods are fully non-rival. The rationale for indexing cost by $n$ is to be able to analyze large economies without making the public goods a free lunch in the limit. We therefore allow for the existence of $\hat{c}^j > 0$ such that $\lim_{n \to \infty} C^j(n)/n = \hat{c}^j > 0$. 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 good $j$, so that the type is given by a vector $\theta_i = (\theta_i^1, ..., \theta_i^M) \in \Theta \subset R^M$. Agent $i$ has preferences represented by the utility function,

$$\sum_{j \in \mathcal{J}} I^j_i \theta_i^j - t_i, \quad (1)$$

where $I^j_i$ 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.

The type $\theta_i$ 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 $\theta_i$. Valuations across goods may be correlated for the individual agent. For brevity of notation, we let $\theta \equiv (\theta_1, \ldots, \theta_n) \in \Theta^n$, which will be referred to as a type profile. In the usual fashion, we let $\theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n)$.

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

$$A = \left\{0, 1\right\}^M \times \left\{0, 1\right\}^{m \times n} \times R^n.$$  

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 $\Theta^n$ to $A$. We represent a randomized mechanism in analogy with the representation of mixed strategies in Aumann [5]. That is, let $\Xi \equiv [0, 1]$, and think of $\vartheta \in \Xi$ as the outcome of a fictitious lottery, where, without loss of generality, $\vartheta$ is uniformly distributed and independent of $\theta$. A random direct mechanism is then a measurable mapping $G : \Theta^n \times \Xi \rightarrow A$. A conceptual advantage of this representation is that it allows for a useful decomposition. That is, we may write $G$ as a $(2M + 1)$-tuple, $G = (\{\zeta^j\}_{j \in J}, \{\omega^j\}_{j \in J}, \tau)$ where,

- **Provision Rule:** $\zeta^j : \Theta^n \times \Xi \rightarrow \{0, 1\}$
- **Inclusion Rule:** $\omega^j : \Theta^n \times \Xi \rightarrow \{0, 1\}^n$
- **Cost-sharing Rule:** $\tau : \Theta^n \rightarrow R^n$.  

We refer to $\zeta^j$ as the provision rule for good $j$, and interpret $E_{\Xi} \zeta^j (\theta, \vartheta)$ as the probability of provision given announcements $\theta$. The rule $\omega^j = (\omega^j_1, \ldots, \omega^j_n)$ is the inclusion rule for good $j$, and $E_{\Xi} \omega^j_i (\theta, \vartheta)$ is interpreted as the probability that agent $i$ gets access to good $j$ when announcements are $\theta$, conditional on good $j$ being provided. Finally, $\tau = (\tau_1, \ldots, \tau_n)$ is the cost-sharing rule, where $\tau_i(\theta)$ is the transfer from agent $i$ to the mechanism designer given announced valuations $\theta$. 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 $(\theta_{-i}, \vartheta)$. A mechanism is incentive compatible if truth-telling is a Bayesian Nash equilibrium in the revelation game induced by $G$,

$$E_{-i}\left[ \sum_{j \in J} \zeta^j(\theta, \vartheta) \omega^j_i(\theta, \vartheta) \theta^j_i - \tau_i(\theta) \right] \geq E_{-i}\left[ \sum_{j \in J} \zeta^j(\hat{\theta}_i, \theta_{-i}, \vartheta) \omega^j_i(\hat{\theta}_i, \theta_{-i}, \vartheta) \theta^j_i - \tau_i(\hat{\theta}_i, \theta_{-i}) \right],$$

$\forall i \in I, \theta \in \Theta^n, \hat{\theta}_i \in \Theta$. (IC)

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

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

$$E_{-i}\left[ \sum_{j \in J} \zeta^j(\theta, \vartheta) \omega^j_i(\theta, \vartheta) \theta^j_i - \tau_i(\theta) \right] \geq 0, \quad \forall i \in I, \theta_i \in \Theta.$$  (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)⁴. Thus a mechanism is thus constrained efficient if it maximizes

$$\sum_{j \in J} \mathbb{E}\zeta^j(\theta, \vartheta) \left[ \sum_{i \in I} \omega^j_i(\theta, \vartheta) \theta^j_i - C^j(n) \right],$$  (4)

over all incentive feasible mechanisms.

It is ex post efficient to provide good $j$ if and only if $\sum_{i \in I} \theta^j_i \geq C^j(n)$, 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 the discussions in Mailath and Postlewaite [17].) Our setup is thus a second best problem.

---

² As shown in Borgers and Norman [9] 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.

³ The qualifier ex ante is crucial for the equivalence. See Ledyard and Palfrey [18] for a characterization of interim efficiency.

⁴ 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 [14].
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 \(g = (\{\rho^j\}_{j \in J}, \{\eta^j\}_{j \in J}, t)\) such that for each good \(j \in J\),

- **Provision Rule**: \(\rho^j : \Theta^n \to [0, 1]\)
- **Inclusion Rule**: \(\eta^j : \Theta \to [0, 1]\)
- **Cost-sharing Rule**: \(t : \Theta \to \mathbb{R}\),

where \(\rho^j\) is the provision rule for good \(j\), \(\eta^j\) 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 \(\theta\), 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 \(\theta_{-i}\); and fourth, all agents are treated symmetrically in terms of the transfer and inclusion rules.

Agent anonymity is built into the notion of a simple mechanism in the inclusion and transfer 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 \(j \in J\), \(\rho^j(\theta) = \rho^j(\theta')\) for every \((\theta, \theta') \in \Theta^n \times \Theta^n\) such that \(\theta'\) can be obtained from \(\theta\) 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 \(G\) of the form (3), there exists an anonymous simple 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 \(\theta_{-i}\), 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.

\[5\text{The actual proof is a bit more complex than simply randomizing with equal probabilities over the } n! \text{ permutations. The reason is that inclusion and provision probabilities are potentially correlated.}\]
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 \( \theta_i = (\theta_1^i, \ldots, \theta_M^i) \) is an exchangeable random variable, that is \( F(\theta_i) = F(\theta'_i) \) whenever \( \theta'_i \) is a permutation of \( \theta_i \); and that there exists \( C(n) \) such that \( C^j(n) = C(n) \) for all \( j \).

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

**Definition 3** Mechanism \( g \) is symmetric if for every \( \theta \) and every permutation \( P : J \rightarrow J \):

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

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 \( \theta^P = (\theta_1^P, \theta_2^P) = ((l, h), (h, l)) \). 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. The result is:

**Proposition 2** Suppose that \( \theta_i \) is an exchangeable random variable and that there exists \( C(n) \) such that \( C^j(n) = C(n) \) for all \( j \in J \). 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 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. Incentive feasibility of the new mechanism follows from incentive feasibility of the original mechanism. Proposition 2 generalizes this procedure by permuting the goods ($M!$ possibilities) and creating a symmetric mechanism by averaging over these.

3 The 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 $i \in I$ and $j \in J$. The probability of any type $\theta_i \in \Theta \equiv \{l, h\}^M$ is denoted $\beta_i (\theta_i)$ and, using standard conventions and independence $\beta (\theta) = \times_{i=1}^n \beta_i (\theta_i)$ and $\beta_{-i} (\theta_{-i}) = \times_{j \neq i} \beta_{-i} (\theta_{-i})$ denote the probabilities of type profile $\theta$ and $\theta_{-i}$ respectively.

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 $\theta_i$ is an exchangeable random variable, we therefore have that the probability that an agent has a high valuation for exactly $m$ goods is given by

$$\Pr \left[ \# \{ j \in J | \theta_i^j = h \} = m \right] \equiv \beta_m = \frac{M!}{m!(M-m)!} \beta_i (\theta_i),$$

(6)

where $\theta_i$ is any type with $m$ high valuations. This formulation allows the willingness to pay between goods to be correlated, but the simplest case is when valuations for different goods are i.i.d., in which case, for any $\theta_i$ with $m$ high valuations, $\beta_i (\theta_i) = \alpha^m (1 - \alpha)^{M-m}$ and $\beta_m = \frac{M!}{m!(M-m)!} \alpha^m (1 - \alpha)^{M-m}$, where $\alpha$ 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,

---

Footnote: 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.
\[
\max_{\{\theta, \eta, t\} \in \Theta^n} \beta(\theta) \sum_{j=1}^{M} \rho^j(\theta) \left[ \sum_{i=1}^{n} \eta^j_i(\theta_i) \theta^j_i - cn \right]
\]  
\[0 \leq \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta) \eta^j_i(\theta_i) \theta^j_i - t_i(\theta_i) - \left[ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i}, \theta'_i) \eta^j_i(\theta'_i) \theta^j_i - t_i(\theta'_i) \right]
\]
for every \(\theta_i, \theta'_i \in \Theta \times \Theta\)
\[0 \leq \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i}, 1) \eta^j_i(1) l - t_i(1)
\]
\[0 \leq \sum_{i=1}^{n} \sum_{\theta_i \in \Theta} \beta_i(\theta_i) t_i(\theta_i) - \sum_{\theta \in \Theta^n} \beta(\theta) \sum_{j=1}^{M} \rho^j(\theta) cn
\]
\[
\rho^j(\theta) \in [0, 1], \quad \eta^j_i(\theta_i) \in [0, 1],
\]
where \(\Theta_{-i}\) 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 \(l = (l, \ldots, l)\) \(9\), and the resource constraint \(10\). 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.

To simplify our discussions in the analysis that follows we use the following terminology:

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

1. **downwards incentive constraint** if \(\theta^j_i \geq \theta'^j_i\) for all \(j \in J\)

2. **upwards incentive constraint** if \(\theta^j_i \leq \theta'^j_i\) for all \(j \in J\)

3. **diagonal incentive constraint** if there exists \(j, k \in J\) so that \(\theta^j_i > \theta'^j_i\) and \(\theta^k_i < \theta'^k_i\)

Incentive constraints that rule out deviations where a single coordinate is misrepresented are henceforth called *adjacent*. If the \(j\)th coordinate in \(\theta_i\) is changed from \(h\) to \(l\) we write \(\theta_i|h_j\) and if the \(k\)th coordinate in \(\theta_i\) is changed from \(l\) to \(h\) we write \(\theta_i|h_k\). Finally, the number of goods for which \(\theta^j_i = h\) is denoted by \(m(\theta_i) \in \{0, \ldots, M\}\).

### 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. In analogue to standard approach to unidimensional problems, we will
show 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 \((\rho, \eta, t)\) is said to be monotonic if \(\eta^i_j(\theta_i) \leq \eta^i_j(\theta'_i)\) and \(\rho^j(\theta_{-i}, \theta_i) \leq \rho^j(\theta_{-i}, \theta'_i)\) whenever \(m(\theta_i) \leq m(\theta'_i)\) and \(\theta^j_i \leq \theta^j_i\).

We will provide sufficient conditions in terms of primitives for monotonicity in the analysis that follow.

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

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

\[
0 \leq \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^M \rho^j(\theta) \eta^i_j(\theta_i) \theta^j_i - t_i(\theta_i)
\]

\[
- \left[ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^M \rho^j(\theta_{-i}, \theta_i | l_k) \eta^i_j(\theta_i | l_k) \theta^j_i - t_i(\theta_i | l_k) \right]
\]

for every \(k\) such that \(\theta^k_i = h\) \([m(\theta_i)\) constraints for every \(i\) and \(\theta_i] \]

\[
0 \leq \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^M \rho^j(\theta_{-i}, 1) \eta^i_j(1) l - t_i(1)
\]

\[
0 \leq \sum_{i=1}^n \sum_{\theta_i \in \Theta} \beta_i(\theta_i) t_i(\theta_i) - \sum_{\theta \in \Theta^a} \beta(\theta) \sum_{j=1}^M \rho^j(\theta) cn
\]

\[
\rho^j(\theta) \in [0, 1], \eta^i_j(\theta_i) \in [0, 1].
\]

Next, we will state a sequence of intermediate results that will be used to characterize the solution to (11) and, eventually, the full problem in (7). Proofs of these results are in the appendix.

### 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 argue 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 $m = 1, \ldots, M$ 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 $i$ and let $\Lambda$ denote the value of the multiplier for the resource constraint (14).

Let $\theta_i$ be a type for which agent $i$ has a high valuation for $m \in \{0, \ldots, M - 1\}$ goods. There are $m$ different ways to replace a coordinate in $\theta_i$ that corresponds to a high valuation. Also, there are $M - m$ types of $\theta'_i$ that can be “turned into” type $\theta_i$ 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

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

whereas the optimality condition for $\theta_i = h = (h, \ldots, h)$ is

$$\lambda(M) + \Lambda \beta_i (h) = 0$$  \hfill (16)

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

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

$$m \lambda(m) = \frac{m!(M-m)!}{M!} \Lambda \sum_{j=m}^{M} \beta_j$$  \hfill (17)

$$= \frac{m!(M-m)!}{M!} \Lambda \Pr [m(\theta_i) \geq m]$$

where $\beta_j$ 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).

At first blush, Lemma 4 may appear like a “no distortion at the top result”. However, this is not the case as it says that a consumer should get her high valuation good irrespective of the number of 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 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 \( m \) high valuations. There are \( M - m \) ways to change one coordinate into a lie with \( m \) high valuations, so the increased 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, so that all consumers always get access to the high valuation goods.

In order to simplify the discussion of the inclusion rules for coordinates \( j \) with \( \theta_i^j = l \) it is useful to define

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

(18)

It is then straightforward to derive the following characterization from (11):

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

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

(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

\footnote{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.}
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

\[
R(m) = \Pr [\bar{m}(\theta_i) \geq m] \{mh + (M - m)l\} + h \sum_{j=1}^{m-1} \beta_j j
\]

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

\[
R(m) - R(m + 1) = \sum_{j=m}^{M} \beta_j \{mh + (M - m)l\} + h \sum_{j=1}^{m-1} \beta_j j
\]

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. This property is well known in many settings (see Hellwig [15] and Norman [23]), where incentive feasibility can be characterized by a single constraint using methods from Myerson [21]. While it is intuitively plausible that the relation between profit and welfare maximization should be the same in our multidimensional setting, we don’t know of any general results that is applicable.

### 3.4 Provision Rules

We are now in a position to start the characterization of 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. 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. 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

$$
\rho^j (\theta) = \begin{cases} 
0 & \text{if } \sum_{m=1}^{M} H^j (\theta, m) + \sum_{m=0}^{M} L^j (\theta, m) \frac{1}{\beta_m (M-m)} \max \{0, G_m (\Phi)\} - cn < 0 \\
\begin{array}{ll}
\frac{z}{2} & \text{if } \sum_{m=1}^{M} H^j (\theta, m) + \sum_{m=0}^{M} L^j (\theta, m) \frac{1}{\beta_m (M-m)} \max \{0, G_m (\Phi)\} - cn = 0 \\
1 & \text{if } \sum_{m=1}^{M} H^j (\theta, m) + \sum_{m=0}^{M} L^j (\theta, m) \frac{1}{\beta_m (M-m)} \max \{0, G_m (\Phi)\} - cn > 0 
\end{array}
\end{cases}
$$

(21)

Just like the optimal inclusion rule, (21) may be understood as a compromise between welfare maximization and profit maximization. That is, 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 $\frac{1}{\beta_m}$. The reason for $\frac{1}{\beta_m}$ is that the easiest way to think about inclusion rules involves a trade-off between the possibility that an agent has exactly $m$ high valuations (probability $\beta_m$) and the probability that the number of high valuations exceed $m$ (probability $\sum_{j=m+1}^{M} \beta_j$). In contrast, the natural units for the welfare created by providing good $j$ are for a fixed $\theta$, 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 factor $\frac{1}{\beta_m}$ 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 (11) to solve (7)

We will now discuss when the solution to the relaxed problem (11) solves also the full problem (7). Many steps in this analysis are similar to Matthews and Moore [19], but, the multidimensional nature of our environment leads to some differences.

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

Lemma 7 Suppose that $(\rho, \eta, t)$ is monotonic and that all downwards and upwards incentive constraints hold. Then, all incentive constraints in (7) are satisfied.

Lemma 8 Suppose that $(\rho, \eta, t)$ is monotonic and that all adjacent downwards incentive constraints hold. Then, all downwards constraints in (7) are satisfied.

Lemma 9 Suppose that $(\rho, \eta, t)$ 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 \((\rho, \eta, t)\) 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 for the relaxed problem (11). As a first step we establish that monotonicity guarantees that the constraints bind.

**Lemma 11** Let \((\rho, \eta, t)\) be an optimal solution to (11). Suppose that (1) \((\rho, \eta, t)\) is not ex post optimal; and (2) \((\rho, \eta, t)\) is monotonic. Then, every downwards adjacent incentive constraint and the participation constraint for type \(i\) 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 \(\theta^i\) where there is slack. To deal with this possibility we instead extract some additional cash from type \(\theta^i\). An inductive argument shows that this additional cash can be used to improve the allocation unless the solution is 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 \((\rho, \eta, t)\) be an optimal solution to (11). Suppose that (1) \((\rho, \eta, t)\) is not ex post optimal; and (2) \((\rho, \eta, t)\) is monotonic. Then, \((\rho, \eta, t)\) is also an optimal solution to the full problem (7).

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

**Lemma 13** Let \((\rho, \eta, t)\) be a solution to (11). Then,

1. \((\eta, \rho)\) is monotonic in the sense that \(\eta^i (\theta^i_1) \leq \eta^i (\theta^i_0)\) and \(\rho^j (\theta_{-i}, \theta^i) \leq \rho^j (\theta_{-i}, \theta^j_0)\) whenever \(m(\theta^i) < m(\theta^j_1)\) and \(\theta^j_1 \leq \theta^i_0\) if \(\frac{1}{M-m} \sum_{j=m+1}^{M} \beta^j_j\) is decreasing in \(m\) on \(\{0, ..., M-1\}\).

2. A sufficient condition for \(\frac{1}{M-m} \sum_{j=m+1}^{M} \beta^j_j\) to be strictly decreasing on \(\{0, ..., M-1\}\) is that the valuation for good \(j\) is independent to any other good \(j'\).
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 \( \beta_m \) a consumer 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 \( \frac{1}{M-m} \sum_{j=m+1}^{M} \beta_j \) the number of high valuations exceed \( m \), and for such a consumer the revenue is reduced by \((h - l)\). In general, \( \frac{1}{M-m} \sum_{j=m+1}^{M} \beta_j \) 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.

### 3.6 The Main Result

Consider the problem to maximize the average probability of provision under the same constraints as the relaxed problem (11).

\[
\max_{\{\rho, \eta, t\}} \sum_{\theta \in \Theta} \beta(\theta) \sum_{j=1}^{M} \rho^j(\theta) \quad \text{s.t. } (12), (13), (14). \tag{22}
\]

Not surprisingly, the characterization of the solution to (22) is qualitatively similar to the constrained welfare problem (see Lemma A2 in the appendix for details). In particular, regardless of whether one considers the constrained welfare problem in (11) or the problem to maximize the probability of provision in (22) we have that the monotonicity condition on \( \frac{1}{M-m} \sum_{j=m+1}^{M} \beta_j \) acts like an “increasing virtual surplus assumption” in that we get a discrete analogue of a threshold characterization:

**Proposition 3** Let \( \rho, \eta, t \) be a solution to either (11) or (22) and suppose that \( \frac{1}{M-m} \sum_{j=m+1}^{M} \beta_j \) is strictly decreasing on \( \{0, \ldots, M - 1\} \). Then, there exists some \( \tilde{m} \) such that:

1. \( \eta^j_i(\theta_i) = \eta(m) = 0 \) for every \( \theta_i \) with \( \theta^l_i = l \) and \( \theta^h_i = h \) for \( m < \tilde{m} \) goods \( k \) in \( J \setminus \{j\} \).
2. \( \eta^j_i(\theta_i) = \eta(m) = 1 \) for every \( \theta_i \) with \( \theta^l_i = l \) and \( \theta^h_i = h \) for \( m > \tilde{m} \) goods \( k \) in \( J \setminus \{j\} \).

At this point we will start to look at sequences of solutions to (11) and (22), where the number of agents \( n \) goes out of bounds. As a result, 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. This is useful because a version of the paradox of voting applies, which simplifies the description of the optimal mechanism tremendously as, for large \( n \), it is almost as if the provision rule is constant in type announcements. Formally, we use the central limit theorem to establish:

**Lemma 14** Let \((\rho_n, \eta_n, \theta_n)\) be a solution to either the constrained welfare problem (11) or to the problem (22) that establishes an upper bound on the probability of provision. Then, \(E \left[ \rho_n^i (\theta) | \theta_i^j \right] - E \left[ \rho_n^i (\theta) | \theta_i^{0j} \right] \to 0\) as \( n \to \infty \) for any \( i \) and any pair \( \theta_i^j, \theta_i^{0j} \in \Theta \).

Lemma 14 immediately implies that \(E \left[ \rho_n^i (\theta) | \theta_i^j \right] - E \left[ \rho_n^i (\theta) \right] \to 0\). Hence, all conditional probabilities appearing in the incentive constraints may be approximated by the ex ante probability of provision, which greatly simplifies the analysis.

**Proposition 4** Suppose that \( \frac{1}{M-m} \sum_{j=m+1}^{M} \beta_j \) is strictly decreasing on \( \{0, ..., M - 1\} \) and let \( R(m) \) be defined as in (20). Then,

1. \( \lim_{n \to \infty} E \rho_n^i (\theta) = 0 \) for every \( j \) if \( \max_m R(m) - cM < 0 \) in any sequence of feasible solutions to the full problem (7);
2. \( \lim_{n \to \infty} E \rho_n^i (\theta) = 1 \) 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 \( \theta_i \) with \( m(\theta_i) \geq m^* \), \( \eta_n^i (\theta_i) = 1 \) for every good \( j \);
   (b) for every type \( \theta_i \) with \( m(\theta_i) = m^* - 1 \), \( \eta_n^i (\theta_i) = 1 \) for every good \( j \) such that \( \theta_i^j = h \), and
   \[ \eta_n^i (\theta_i) \to \frac{R(m^*) - cM}{R(m^*) - R(m^* - 1)} \]
   for every good \( j \) such \( \theta_i^j = l \);
   (c) for every type \( \theta_i \) with \( m(\theta_i) \leq m^* - 2 \), \( \eta_n^i (\theta_i) = 1 \) for every good \( j \) such that \( \theta_i^j = h \) and \( \eta_n^i (\theta_i) = 0 \) for every good \( j \) such that \( \theta_i^j = l \).

The proof of Proposition 4 requires quite a bit of technical work. Still, the key ideas are rather simple. Once we have established that the solution must have a threshold characterization, and that the influence of any individual agent’s announcement vanishes as the number of agents go out of bounds, 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 \( \theta \). 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

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. Since there are no complementarities in preferences it should be obvious that the analysis of a providing a single excludable public good is the same as the analysis of the case where \( M \) goods are provided by independent agencies is the same.\footnote{If goods are complements or substitutes this is no longer true.}

With only a single good, the condition on \( \frac{1}{M-m} \sum_{j=m+1}^{M} \beta_j \) is satisfied by default, so Proposition \ref{prop:one_good} is always applicable. Furthermore,

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

Hence, Proposition \ref{prop:one_good} says that \( \lim_{n \to \infty} E \rho_n (\theta) = 0 \) if \( \max \{ l, \alpha h \} - c < 0 \) and that \( \lim_{n \to \infty} E \rho_n (\theta) = 1 \) if the inequality is reversed.

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 \( c \). With this case ruled out we have that the probability for access for type \( l \) converges to

\[
\frac{R(1) - c}{R(1) - R(0)} = \frac{\alpha h - c}{\alpha h - l} \in (0, 1)
\]

as \( n \) goes out of bounds. As this implies that \( G_1 (\Phi) = 0 \) we have that the provision rule in (21) simplifies to

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

which curiously is exactly the same provision rule as the one that would be optimal for a profit maximizing monopolist. In other words, the welfare loss associated with a for profit monopolist relative to the constrained social optimum is due to over-exclusion only. Asymptotically, this is true for \( M \geq 2 \) too, but the provision rule (24) is part both of the profit and surplus maximizing solutions for any \( n \). With more than one good, there is typically under-provision by a for profit provider for finite \( n \).

Finally, note that the case when valuations are perfectly correlated is also equivalent to the case with a single good.
4.2 Two Goods

In the case with two goods we have that \( \frac{M}{m} \sum_{j=m+1}^{M} \beta_j \) is strictly decreasing in \( m \) on \( \{0, 1\} \) if and only if
\[
\frac{\beta_2}{\beta_1} < \frac{1}{2} \left( \frac{\beta_1 + \beta_2}{\beta_0} \right) = \frac{1}{2} \left( \frac{1 - \beta_0}{\beta_0} \right) \iff \frac{\beta_1}{2} > \frac{\beta_0 \beta_2}{1 - \beta_0}.
\]
(25)

Furthermore
\[
R(m) = \begin{cases} 
2l & \text{for } m = 0 \\
(\beta_1 + \beta_2) (h + l) & \text{for } m = 1 \\
2\alpha = 2 \left( \beta_1 + \beta_2 \right) & \text{for } m = 2.
\end{cases}
\]

Ruling out the trivial case when \( l \geq c \) we see from Proposition 3 that there are three possibilities:

1. \( \lim_{n \to \infty} E \rho_i^n (\theta) = 0 \) for \( j = 1, 2 \);

2. \( (\beta_1 + \beta_2) (h + l) > 2c \), in which case \( \lim_{n \to \infty} E \rho_i^n (\theta) = 1 \) for \( j = 1, 2 \) and all consumers with at least 1 high valuation good get access to both goods and the agents 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. \( 2\alpha > 2c > (\beta_1 + \beta_2) (h + l) \), in which case \( \lim_{n \to \infty} E \rho_i^n (\theta) = 1 \) for \( j = 1, 2 \) and consumers with 2 high valuations get access to bot goods, while consumers with 2 low valuations don’t get any access at all, whereas consumers with one high valuation get access to the low valuation good with probability
\[
\frac{R(2) - 2c}{R(2) - R(1)} = \frac{2\alpha h - 2c}{2\alpha h - (\beta_1 + \beta_2) (h + l)} \in (0, 1).
\]
(26)

There is a non-empty parameter region where \( (\beta_1 + \beta_2) (h + l) > 2c > 2ah \) in which we get asymptotic non-provision if goods are provided separately, but where the optimal bundling mechanism provides bot goods for sure.

In addition, the optimal bundling mechanism leads to strict efficiency gains relative the non-bundling regime because the probability of inclusion for low-valuation agents is increased, even in cases when the goods can be provided without bundling. To see this, suppose that \( ah > c \) so that both public goods will be asymptotically provided with probability one with or without bundling. From (23), we know that under the best separate provision mechanism, the probability for access to a low valuation agent is \( (ah - c) / (ah - l) \). In contrast, (26) 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
\[
\frac{\frac{\beta_1}{\beta_0 + \beta_1}}{2\alpha h - 2c} = \frac{2\alpha h - 2c}{2\alpha h - (\beta_1 + \beta_2) (h + l)}.
\]
(27)

(27)
Some algebra shows that (27) is larger than \( \frac{\alpha h - c}{\alpha h - l} \) whenever \( \frac{\beta_1}{2} > \frac{\beta_2}{1 - \beta_0} \) which is precisely the condition under which Proposition 4 is applicable. Fewer consumers are thus excluded in the optimal bundling mechanism. A similar calculation applies to the case 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 \( \frac{1}{M-m} \sum_{j=m+1}^{M} \frac{1}{\beta_j} \) fails. As shown in (25) this reduces to the condition that \( \frac{\beta_1}{2} \leq \frac{\beta_2}{1 - \beta_0} \), 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 5** Assume that \( \frac{\beta_1}{2} \leq \frac{\beta_2}{1 - \beta_0} \) 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 \( \theta_i \) with \( \theta_i^j = l \).

The interesting aspect of Proposition 5 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. Ruling out the trivial case with \( l > c \), the question is thus whether \( (\beta_1 + \beta_2)(h + l) \) exceeds \( 2\alpha h \) or not. Without further parameter restrictions the inequality can go either way, but, positive asymptotic provision requires that the revenue exceeds \( 2c \) and when \( \alpha h < c \) and \( l < c \) are both satisfied it follows that

\[
(\beta_1 + \beta_2)(h + l) < (\beta_1 + \beta_2)(h + c) < (\beta_1 + \beta_2) \left( \frac{1}{\alpha} + 1 \right) c
\]

\[
= \left[ 2 + \frac{(1 - \beta_0) \left( \frac{\beta_1 + \beta_2}{2} - \beta_2 \right)}{\beta_1 + \beta_2} \right] c,
\]

\[
< \left[ 2 + \frac{(1 - \beta_0) \left( \frac{\beta_2 - \beta_0}{1 - \beta_0} + \beta_2 \right)}{\beta_1 + \beta_2} \right] c = 2c,
\]

20
when $\frac{1}{2}\beta_1 \leq \frac{\beta_0\beta_2}{1 - \beta_0^2}$. 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 $E\hat{\rho}_n(\theta) \to 0$ is the same as for the previous analysis.

For the second part of Proposition 5, consider the case when either separate provision or bundling provide sufficient revenue for provision. In the solution characterized by Proposition 5 the asymptotic ex ante probability of being excluded from consumption from 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 $\frac{\beta_1}{2} > \frac{\beta_0\beta_2}{1 - \beta_0^2}$ is used the ex ante probability from being excluded 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 (27) shows that ex ante probability of getting access and therefore also the social surplus is actually smaller using the bundling mechanism in this case.

5 Discussion

5.1 Our Model as a Positive Theory of Bundling

Bundled discounts are common in many markets, both highly competitive and more concentrated. In fact, it is sometimes argued that almost any commodity sold is best viewed as a bundle of characteristics. That is, a car is considered a “bundle” consisting of a chassis, an engine, cup holders, a stereo system, air conditioning etc. There is obviously nothing wrong with this point of view, but our analysis takes a more narrow perspective in that we only consider commodities that are viable “stand-alone goods” in the sense that a final consumer may have a substantial willingness to pay for the commodity even if no other goods in the bundle are consumed. Most components in the bundle we refer to as a “car” seem to fail this test.

Once we make the restriction to goods that are or could be sold separately, it seems that many commonly bundled commodities indeed are ones with non-rival characteristics, or, more generally, goods with a substantial fixed cost component. Software, music, electronic libraries, and TV programming are obvious examples. Given the increased importance of these kind of goods we believe that is important to develop a theory of bundling that explicitly incorporates the non-rival characteristics.

In the interest of brevity, we focused in this paper on the (harder) problem to maximize social welfare, but the profit maximizing selling strategy comes out as a by-product of this analysis. We haven’t stressed it, but it is easy to see that the only change in the asymptotic characterization of
the profit maximizing mechanism is that the threshold number of high valuation goods a consumer needs in order to get access to the low valuation goods will increase. Or, put differently, the price of the grand bundle will be higher than in Proposition 4 as we now seek as high a profit as possible rather than to break even. Everything else in Proposition 4 is unchanged.

As a positive theory of bundling, the key difference between our model and the standard model with constant unit costs is that the decision when at all to provide a commodity becomes non-trivial, which in turn makes it necessary to go beyond the standard setup with a single consumer. This makes the characterization of the optimal selling mechanism somewhat complicated in the case of a given finite number of consumers. However, taking limits of the *exact solutions* we obtain an intuitive characterization of the optimal mechanism in a large finite economy.

5.2 Creating a Regulatory Benchmark for Bundling with Large Fixed Costs

Bundling may under some circumstances be in violation to current US anti-trust legislation, and we note that the legality and desirability of bundling of software (a non-rival good) have been a critical point of debate in several actual cases in recent years. While economists have been involved in this discussion, we note that economic theory lacks a clear-cut normative benchmark for this case. Almost all formal models are covering the case with constant unit costs, which has the somewhat awkward implication that market power seems hard to justify unless there are some barriers to entry.

We hope that the analysis in this paper will be a small step towards the development of a useful normative benchmark for bundling in natural monopoly situations. Obviously, our model is highly stylized. However, it is flexible enough to generate a non-trivial trade-off. On the one side, it is a logical possibility 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, so a requirement to “unbundle” is strictly worse for the consumers in this case, regardless of what other (budget balancing) remedies the decision to unbundle would be combined with. On the other side, the fact that bundling makes it easier for the monopolist to extract consumer surplus (which is also what drives the potential positive effect of bundling) 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 seems to be counter factual in many cases. Moreover, arguments based on alleged complementarities are often used to justify bundling. It is rather clear that in itself, complementarities are not sufficient to provide a rationale for bundling. If many consumers like consume a particular computer program jointly with their computer, they may choose to do

---

10 This line of reasoning was an important part of the motivation in the decision by the Office of Fair Trading [24] in the UK on alleged anti-competitive mixed bundling by the British Sky Broadcasting Limited.
so regardless of whether the computer program is bundled with the computer or not. 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 very easy to allow the consumer to opt out of the standard software bundle (a choice which is usually not available). Hence, it seems that the argument that “bundling is good for the customer because the customers like the goods together” is in general somewhat suspect. However, it may be that complementarities could 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


A  Appendix A: Proofs of Results in Section 3

Proof of Lemma 7.

Proof. The only variables that are not automatically in a compact set are the transfers. However, 
\( t_i(1) \leq ML < Mh \) 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 without loss of generality may bound \( t_i(\theta_i) \) from above by \( M^2 h \). 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 without loss may bound \( t_i(\theta_i) \) from below by \( -M^2 h \). Hence, existence of a solution to (11) follows from Weierstrass maximum theorem. \( \blacksquare \)

Proof of Lemma 2.

Proof. Let \( m(\theta_i) = m(\theta_i | l_k) \) denote the multiplier associated with one of the \( m - 1 \)
downwards adjacent constraints for type \( \theta_i \) and let \( \lambda_i(\theta_i | l_k) \) denote the multiplier associated with
one of the \( m - 1 \) downwards adjacent constraints for type \( \theta_i \). 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 \( \lambda_i(\theta_i, \theta_i') = \lambda_i(\bar{\theta}_i, \bar{\theta}_i') \).

Proof of Lemma 3.

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

\[
0 \leq \sum_{\theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho_j(\theta) \eta^j(\theta_i) \theta^j_i - t_i(\theta_i) - \sum_{\theta_{-i}} \beta_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho_j(\theta_{-i}, \theta_i | l_k) \eta^j(\theta_i | l) \theta^j_i - t_i(\theta_i | l_k)
\]

(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, and also in the resource
constraint (14). The optimality condition with respect to \( t_i(\theta_i) \) can thus be written as

\[
-m\lambda(m) + (M - m) \lambda(m + 1) + \Lambda \beta_i(\theta_i) = -m\lambda(m) + (M - m) \lambda(m + 1) + \Lambda m! (M - m)! \beta_m = 0,
\]

(A2)

where we use (6) to relate the probability of \( \theta_i \) with the probability of \( m(\theta_i) \).

For \( m = M \) condition (A2) reads \( M\lambda(M) = \Lambda \beta_M \), implying that Lemma 3 is also true for
\( m = M \).

We now assume that \( \lambda(m) \) is given by the expression in (17) for some \( m \leq M \). The optimality
condition (A2) with respect to $t_i (\theta'_i)$ where $m (\theta'_i) = m - 1$ then reads:

$$0 = -(m - 1) \lambda (m - 1) + (M - m + 1) \lambda (m) + \frac{(m - 1)! (M - m + 1)!}{M!} \beta_{m - 1}$$

$$= -(m - 1) \lambda (m - 1) + \frac{(M - m + 1)!}{m} m \lambda (m) + \frac{(m - 1)! (M - m + 1)!}{M!} \beta_{m - 1}$$

$$= -(m - 1) \lambda (m - 1) + \frac{(M - m + 1)!}{m} (M - m)! \Lambda \sum_{j=m}^{M} \beta_j + \frac{(m - 1)! (M - m + 1)!}{M!} \beta_{m - 1}$$

$$= -(m - 1) \lambda (m - 1) + \frac{(m - 1)! (M - m + 1)!}{M!} \Lambda \sum_{j=m-1}^{M} \beta_j,$$

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

Proof of Lemma 4.

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

$$\sum_{\theta_{-i} \in \Theta_{-i}} \beta (\theta) \rho^j_i (\theta) h + \lambda (m) m \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \rho^j_i (\theta) h$$

$$- \lambda (m + 1) (M - m) \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \rho^j_i (\theta) h + \gamma^j_i (\theta_i) - \phi^j_i (\theta_i) = 0$$

(A3)

$$\gamma^j_i (\theta_i) \eta^j_i (\theta_i) = 0 \text{ and } \phi^j_i (\theta_i) \left(1 - \eta^j_i (\theta_i)\right) = 0,$$

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

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

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

$$\beta_i (\theta_i) h + \lambda (m) mh - \lambda (m + 1) (M - m) h + \gamma^j_i (\theta_i) - \phi^j_i (\theta_i) = 0,$$

(A5)

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

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

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 then that

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

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:

\[
\beta_i(\theta_i) l + \lambda(m) ml - \lambda(m+1) [(M-m-1)l + h] + \gamma_i^j(\theta_i) - \phi_i^j(\theta_i) = 0.
\]

Using the relations between the multipliers we find that

\[
\beta_i(\theta_i) l + \lambda(m) ml - \lambda(m+1) [(M-m-1)l + h] = \gamma_i^j(\theta_i) - \phi_i^j(\theta_i)
\]

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

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

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
$$\beta(\theta) \left[ \sum_{i=1}^{n} \eta_i^j (\theta_i) \theta_i^j - cn \right] + \sum_{m=0}^{M} \lambda(m) m \left[ H_j^j (\theta, m) \beta_{-i} (\theta_{-i}) h + L_j^j (\theta, m) \beta_{-i} (\theta_{-i}) \eta(m) l \right]$$

$$- \sum_{m=0}^{M-1} \lambda(m+1) \left[ H_j^j (\theta, m) \beta_{-i} (\theta_{-i}) (M-m) h + L_j^j (\theta, m) \beta_{-i} (\theta_{-i}) \eta(m) \{(M-m) l + (h-l)\} \right]$$

$$- \lambda \beta(\theta) cn + \gamma_j^j (\theta) - \phi_j^j (\theta) = 0,$$

together with complementary slackness conditions. Recall that $H_j^j (\theta, m) (L_j^j (\theta, m))$ is the number of agents that has a high (low) valuation for good $j$ and $m$ high valuations in total, where we abuse previously used notation by using $\eta(m)$ for the probability that an agent gets access to his low valuation goods when having $m$ high valuations, and where $\gamma_j^j (\theta)$ and $\phi_j^j (\theta)$ are the multipliers for the boundary constraints. After using (A4) and the fact that

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

we can write the condition as

$$\sum_{m=0}^{M} H_j^j (\theta, m) h + \sum_{m=0}^{M} L_j^j (\theta, m) \eta(m) l + \sum_{m=0}^{M} \lambda(m) \left[ H_j^j (\theta, m) \frac{m}{\beta_i (\theta_i)} h + L_j^j (\theta, m) \frac{m}{\beta_i (\theta_i)} \eta(m) l \right]$$

$$- \sum_{m=0}^{M-1} \lambda(m+1) \left[ H_j^j (\theta, m) \frac{(M-m)}{\beta_i (\theta_i)} h + L_j^j (\theta, m) \frac{1}{\beta_i (\theta_i)} \eta(m) \{(M-m) l + (h-l)\} \right]$$

$$- (1+\Lambda) cn + \frac{\gamma_j^j (\theta) - \phi_j^j (\theta)}{\beta(\theta)} = 0.$$

Collecting terms we get

$$\sum_{m=0}^{M} H_j^j (\theta, m) \left\{ h + \frac{h}{\beta_i (\theta_i)} [\lambda(m) m - \lambda(m+1) (M-m)] \right\}$$

$$\sum_{m=0}^{M} L_j^j (\theta, m) \left\{ \eta(m) l + \frac{\eta(m)}{\beta_i (\theta_i)} [\lambda(m) ml - \lambda(m+1) \{(M-m) l + (h-l)\}] \right\}$$

$$- (1+\Lambda) cn + \frac{\gamma_j^j (\theta) - \phi_j^j (\theta)}{\beta(\theta)} = 0.$$

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

$$\sum_{m=0}^{M} H_j^j (\theta, m) \left\{ h + \Lambda h \right\} + \sum_{m=0}^{M} L_j^j (\theta, m) \left\{ \eta(m) l + \eta(m) \left[ \Lambda l - \frac{\lambda(m+1)}{\beta_i (\theta_i)} (h-l) \right] \right\}$$

$$(A8)$$

$$- (1+\Lambda) cn + \frac{\gamma_j^j (\theta) - \phi_j^j (\theta)}{\beta(\theta)} = 0.$$

\[11\] The notation is bad in that $\beta_{-i} (\theta_{-i})$ would be more appropriately denoted by $\beta_{-i} (\theta_{i})$ where $\theta_i$ would describe a type that would enter in the particular term (and therefore change with each term). Concretely, $H_j^j (\theta, m) \beta_{-i} (\theta_{-i})$ would correspond to $\beta_{-i} (\theta_{i})$ for $\theta_i$ with $\theta_i = h$ 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

$$\frac{\lambda (m + 1)}{\beta_i (\theta_i)} / \beta_i (\theta_i) = \frac{1}{\beta_i (\theta_i)} \frac{m! (M - m - 1)!}{M!} \Lambda \sum_{j=m+1}^{M} \beta_j \quad (A9)$$

Hence

$$\eta (m) l + \eta (m) \left[ \Lambda l - \frac{\lambda (m + 1)}{\beta_i (\theta_i)} (h - l) \right] = \eta (m) l + \eta (m) \left[ \Lambda l - \frac{1}{\beta_m (M - m)} \Lambda \sum_{j=m+1}^{M} \beta_j (h - l) \right]$$

$$= (1 + \Lambda) \eta (m) \left\{ l - \frac{1}{\Lambda} \sum_{j=m+1}^{M} \beta_j (h - l) \right\}$$

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

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

$$/ (A9) / = \frac{(1 + \Lambda)}{\beta_m (M - m)} \max \{ 0, G_m (\Phi) \} .$$

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.

\textbf{Proof of Lemma 7}

Proof. The solution to (11) violates an incentive constraint in (7) if there exists $\theta_i, \theta'_i$ such that

$$\sum_{\theta_i \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \sum_{j=1}^{M} \rho^j (\theta) \eta^j_i (\theta_i) \theta^j_i - t_i (\theta_i) < \sum_{\theta_i \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \sum_{j=1}^{M} \rho^j (\theta_{-i}, \theta'_i) \eta^j_i (\theta'_i) \theta^j_i - t_i (\theta'_i) . \quad (A10)$$

Since $\theta_i$ is exchangeable and costs are identical for all goods, we can apply Proposition 2 (in conjunction with Lemma 3) to conclude that it is without loss of generality to assume that there exists $\{ \eta (m), P^h (\theta_{-i}, m), P^l (\theta_{-i}, m), t (m) \}_{m=0}^{M}$, such that:

1. $\eta^j_i (\theta_i) = \eta (m)$ for every $(\theta_i, j)$ such that $\theta^j_i = l$ for good $j$ and $\theta^k_i = h$ for exactly $m$ goods $k \in J$. 

29
2. \( \rho^j (\theta_{-i}, \theta^j_i) = P^h (\theta_{-i}, m) \) for every \((\theta_i, j)\) such that \(\theta^j_i = h\) for good \(j\) and \(\theta^k_i = h\) for exactly \(m\) goods \(k \in \mathcal{J}\) (including good \(j\))

3. \( \rho^j (\theta_{-i}, \theta^j_i) = P^l (\theta_{-i}, m) \) for every \((\theta_i, j)\) such that \(\theta^j_i = l\) for good \(j\) and \(\theta^k_i = h\) for exactly \(m\) goods \(k \in \mathcal{J}\).

4. \( t_i (\theta_i) = t (m) \) for every \(\theta_i\) such that \(\theta^k_i = h\) for exactly \(m\) goods \(k \in \mathcal{J}\).

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

\[
\sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left[ P^h (\theta_{-i}, m) mh + P^l (\theta_{-i}, m) \eta (m) (M - m) l \right] - t (m)
\]

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

\[
\sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left\{ P^h (\theta_{-i}, m') \left[ (r + 1) h + (m' - r - 1) l \right] + P^l (\theta_{-i}, m') \eta (m') \left[ (s + 1) l + (M - m' - s - 1) h \right] \right\} - t (m')
\]

where the inequality follows from the assumed monotonicity [i.e. \( P^h (\theta_{-i}, m') \geq P^l (\theta_{-i}, m') \)] and \(\eta (m') \leq 1\). This is a violation of the upward incentive constraints, a contradiction to the postulate of the Lemma that all upward 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

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

\[
\geq \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left[ P^h (\theta_{-i}, k) Kh + P^l (\theta_{-i}, K) \eta (K) (M - m) l + (m - K) h \right] - t (K)
\]

\[
= U (m, K)
\]
is satisfied by hypothesis. By assumption the downwards adjacent incentive constraint for type $K$ holds, implying that

$$U(K, K) = \sum_{\theta_i \in \Theta_i} \beta_{-i}(\theta_i) \left[ p^h(\theta_{-i}, K) Kh + p^l(\theta_{-i}, K) \eta(K)(M - K) l \right] - t(K) \tag{A13}$$

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

$$\sum_{\theta_i \in \Theta_i} \beta_{-i}(\theta_i) \left[ p^h(\theta_{-i}, K - 1) (K - 1) h + p^l(\theta_{-i}, K) \eta(K)(M - K) l + (m - K + 1) h \right] - t(K - 1)$$

$$\sum_{\theta_i \in \Theta_i} \beta_{-i}(\theta_i) \left[ p^h(\theta_{-i}, K - 1) (K - 1) h + p^l(\theta_{-i}, K) \eta(K)(M - K) l + h \right] - t(K - 1)$$

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

**Proof of Lemma 9**

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

$$U(m, m) = \sum_{\theta_i \in \Theta_i} \beta_{-i}(\theta_i) \left\{ p^h(\theta_{-i}, m) mh + p^l(\theta_{-i}, m) \eta(m)(M - m) l \right\} - t(m)$$

$$\geq \sum_{\theta_i \in \Theta_i} \beta_{-i}(\theta_i) \left\{ p^h(\theta_{-i}, k) [mh + (k - m) l] + p^l(\theta_{-i}, k) \eta(k)(M - k) l \right\} - t(k)$$

$$= U(m, k)$$

for all $k \in \{m + 1, \ldots, K\}$. If $K = M$ all upwards constraint hold by assumption. If $K < M$, the
We then note that

\[ U (K, K) = \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^h (\theta_{-i}, K) Kh + P^l (\theta_{-i}, K) \eta (K) (M - K) l - t (K) \]

(A14)

\[ \geq \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^h (\theta_{-i}, K + 1) [Kh + l] + P^l (\theta_{-i}, K + 1) \eta (K + 1) (M - K - 1) l - t (K + 1) \]

\[ = U (K, K + 1). \]

We then note that

\[ U (m, m) - U (m, K + 1) \geq U (m, K) - U (m, K + 1) \]

\[ = \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left\{ P^h (\theta_{-i}, K) [mh + (K - m) l] + P^l (\theta_{-i}, K) \eta (K) (M - K) l \right\} - t (K) \]

\[ - \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^h (\theta_{-i}, K + 1) [mh + (K + 1 - m) l] + P^l (\theta_{-i}, K + 1) \eta (K + 1) (M - K - 1) l - t (K + 1) \]

\[ = \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left\{ P^h (\theta_{-i}, K) Kh + P^l (\theta_{-i}, K) \eta (K) (M - K) l \right\} - t (K) \]

\[ + \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left\{ P^h (\theta_{-i}, K) (K - m) (l - h) \right\} \]

\[ - \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^h (\theta_{-i}, K + 1) [Kh + l] + P^l (\theta_{-i}, K + 1) \eta (K + 1) (M - K - 1) l - t (K + 1) \]

\[ + \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^h (\theta_{-i}, K + 1) (K - m) (h - l) \]

\[ = U (K, K) - U (K, K + 1) + \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left[ P^h (\theta_{-i}, K + 1) - P^h (\theta_{-i}, K) \right] (K - m) (h - l), \]

\[ \geq 0 \text{ by A14} \]

\[ \geq 0 \text{ by monotonicity of } P^h \]

implying that type \( m \) has no incentive to mis-report as type \( K + 1 \). 

\[ \square \]

**Proof of Lemma 10.**

Proof. Consider types \( \theta_i \) and \( \theta'_i \) 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:

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

\[ U (m, m + 1) = \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left\{ P^h (\theta_{-i}, m + 1) [mh + l] + P^l (\theta_{-i}, m + 1) \eta (m + 1) (M - m - 1) l \right\} - t (m + 1). \]
Proof. From Lemma 6, we know that
\[
U (m,m) - U (m,m+1)
\]
\[
= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^h (\theta_{-i}, m) mh + P^l (\theta_{-i}, m) \eta (m) (M-m) l - t (m)
\]
\[
- \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) [P^h (\theta_{-i}, m+1) [mh+l] + \eta (m+1) (M-m-1) l] - t (m+1)
\]
\[
= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^h (\theta_{-i}, m) mh + P^l (\theta_{-i}, m) \eta (m) \{ (M-m-1) l + h \} - t (m)
\]
\[
= U (m+1, m)
\]
\[
+ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) P^l (\theta_{-i}, m) \eta (m) (l-h)
\]
\[
- \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) [P^h (\theta_{-i}, m+1) (m+1) h + P^l (\theta_{-i}, m+1) \eta (m+1) (M-m-1) l] - t (m+1)
\]
\[
= U (m+1, m+1)
\]
\[
+ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) [P^h (\theta_{-i}, m+1) [h-l]]
\]
\[
= U (m+1, m) - U (m+1, m+1) + \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \{P^h (\theta_{-i}, m+1) - P^l (\theta_{-i}, m) \eta (m)\} [h-l]
\]
\[
= 0 \text{ by Lemma } \ref{LemmaA1}\]
\[
\geq 0 \text{ as } P^h (\theta_{-i}, m+1) - P^l (\theta_{-i}, m) \eta (m) \geq 0
\]
\[
\geq 0 \text{.}
\]

Proof of Lemma \ref{LemmaA1}.

Proof. We first prove a lemma:

Lemma A1 In any solution to (11), \( \rho^i (\theta) = 0 \) if \( \sum_{i=1}^n \eta^j_i (\theta_i) \theta^j_i - cn \leq 0 \).

Proof. 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
\[
\sum_{i=1}^n \eta^j_i (\theta_i) \theta^j_i - cn = \sum_{m=1}^M H^j (\theta, m) h + \sum_{m=0}^M L^j (\theta, m) l - cn,
\]
and
\[
\frac{1}{\beta_m (M-m)} G_m (\Phi) = l - \Phi \left[ (h-l) \sum_{j=m+1}^M \beta_j \right] < l \text{.}
\]
(Proof of Lemma \ref{LemmaA1} continued:) As we assume way the case when \( (\rho, \eta, t) \) is ex post optimal, it follows that \( (n-1) h > nc \), as otherwise the ex post optimal mechanism is either to implement if and only if \( \theta^j_i = h \) for all \( i \) (if \( (n-1) h < nc < nh \)) or to never provide any good (if \( nh \leq nc \)). 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 \(\theta_i\) such that

\[
\sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \sum_{j=1}^M \rho^j (\theta) \eta^j_i (\theta_i) \theta_i^j - t_i (\theta_i) > \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \sum_{j=1}^M \rho^j (\theta_{-i}, \theta_i \setminus l_k) \eta^j_i (\theta_i \setminus l_k) \theta_i^j - t_i (\theta_i \setminus l_k)
\]

If \(\eta^j_i (\theta_i \setminus l_k) < 1\) we can increase \(\eta^j_i (\theta_i \setminus l_k)\) slightly without violating the constraint. Since \((n - 1)h > nc\) it follows that \(\sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \rho^j (\theta_{-i}, \theta_i \setminus l_k) > 0\) for every \(j\). Hence, the value of the objective function increases, which contradicts the assumption that we considered an optimum in the first place. Suppose instead that \(\eta^j_i (\theta_i \setminus l_k) = 1\). Then, we can find some \(\varepsilon > 0\) and raise the fee to \(\bar{t}_i (\theta_i) = t_i (\theta_i) + \varepsilon\) for \(\theta_i\) and every \(\theta'_i\) with \(m (\theta'_i) = m (\theta_i)\). This raises some extra revenue, implying that the fee can be lowered slightly for some other types without upsetting the resource constraint. We can therefore proceed inductively as follows. First consider types \(\theta'_i\) with \(m (\theta'_i) = m (\theta_i) - 2\). If \(\eta^j_i (\theta'_i) < 1\) we can increase the surplus by increasing \(\eta^j_i (\theta'_i)\) for \(\theta'_i\) 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 \(\eta^j_i (\theta'_i) = 1\) and \(m (\theta_i) - 2 \geq 1\) repeat the argument for types \(\theta'_i\) with \(m (\theta'_i) = m (\theta_i) - 3\). By Lemma A1 it follows that the associated increased access is always desirable from the point of view of society, so, by induction, we conclude that either there exists a feasible allocation with a higher surplus or that \(\eta^j_i (\theta_i) = 1\) for all \(\theta_i\) and \(j\), meaning that there are no active use 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 \(\theta\) 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 that it is immediate from (21) that the probability of provision given \((h, \theta_{-i})\) is always weakly higher than under \((\theta_i, \theta_{-i})\) for any \(\theta_{-i} \in \Theta_{-i}\). Next, it is easy to use (19) and (21) to check that a sufficient condition for monotonicity is that \(\frac{1}{\beta_m (M - m)} \max \{0, G_m (\Phi)\}\) is increasing in \(m\) one \(\{0, \ldots, M - 1\}\). But,

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

(A15)
which proves the first part of the claim.

(2) If valuations are independent across goods so that the probability that an agent has a high valuation for any good \( j \) is \( \alpha \), then the probability that an agent has \( m \) high valuations is given by

\[
\beta_m = \frac{M!}{m!(M-m)!} \alpha^m (1-\alpha)^{M-m},
\]

which 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

\[
\frac{1}{\beta_{m+1}(M-m-1)} \sum_{j=m+2}^{M} \beta_j = \frac{1}{\beta_{m+1}(M-m-1)} \sum_{j=m+1}^{M-1} \beta_{j+1} = \frac{(m+1)(1-\alpha)}{(M-m)\alpha(M-m-1)} \beta_m \sum_{j=m+1}^{M-1} \frac{(M-j)\alpha}{(j+1)(1-\alpha)} \beta_j
\]

\[
= \frac{1}{(M-m)\beta_m} \sum_{j=m+1}^{M-1} \frac{(m+1)(M-j)}{(j+1)(M-m-1)} \beta_j
\]

\[/j \geq m + 1 \text{ and } \beta_M > 0/ < \frac{1}{(M-m)\beta_m} \sum_{j=m+1}^{M} \beta_j.
\]

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

Proof of Proposition 3.

Proof. For the problem (11) 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 \( G_m (\Phi) \leq 0 \), which is immediate from (A15). For the problem (22) we have that the only role of the inclusion rule is to maximize revenue, and the optimal inclusion rule is then

\[
\eta_i^j (\theta_i) = \eta (m) \equiv \begin{cases} 
0 & \text{if } G_m (1) < 0 \\
\ z \in [0,1] & \text{if } G_m (1) = 0 \\
1 & \text{if } G_m (1) > 0. 
\end{cases} \tag{A18}
\]

Proof of Lemma 14.

Proof. We begin by establishing a Lagrangian characterization of the solution to (22) similar to that of the welfare problem (11):

Lemma A2 Let \((\rho, \eta, t)\) be an optimal solution to (22) 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 (A18);
2. there exists some $\Lambda \geq 0$ such that the provision rule for good $j$ satisfies

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

Proof. The derivation of the inclusion rules follow the analysis of the constrained welfare problem step by step and is omitted. For the provision rule, we have that the optimality conditions for $\rho^j (\theta)$ associated with the problem [22] may be written as

$$
\beta (\theta) + \sum_{m=0}^{M} \lambda (m) m \left[ H^j (\theta, m) \beta_{-i} (\theta_{-i}) h + L^j (\theta, m) \eta (m) l \right]
$$

$$
- \sum_{m=0}^{M-1} \lambda (m + 1) \left[ H^j (\theta, m) \beta_{-i} (\theta_{-i}) (M - m) h + L^j (\theta, m) \beta_{-i} (\theta_{-i}) \eta (m) \{(M - m) l + (h - l)\} \right]
$$

$$
- \Lambda \beta (\theta) cn + \gamma^j (\theta) - \phi^j (\theta) = 0,
$$

Together with complementary slackness, where we recall that $H^j (\theta, m) \ (L^j (\theta, m))$ is the number of agents that has a high (low) valuation for good $j$ and $m$ high valuations in total, where we abuse previously used notation by using $\eta (m)$ for the probability that an agent gets access to his low valuation goods when having $m$ high valuations, and where $\gamma^j (\theta)$ and $\phi^j (\theta)$ are the multipliers for the boundary constraints. By noting that $\frac{\beta_{-i} (\theta_{-i})}{\beta (\theta)} = \frac{1}{\beta_i (\theta_i)}$, we can write the condition as

$$
1 + \sum_{m=0}^{M} \lambda (m) \left[ H^j (\theta, m) \frac{m}{\beta_i (\theta_i)} h + L^j (\theta, m) \frac{m}{\beta_i (\theta_i)} \eta (m) l \right]
$$

$$
- \sum_{m=0}^{M-1} \lambda (m + 1) \left[ H^j (\theta, m) \frac{(M - m)}{\beta_i (\theta_i)} h + L^j (\theta, m) \frac{1}{\beta_i (\theta_i)} \eta (m) \{(M - m) l + (h - l)\} \right]
$$

$$
- \Lambda cn + \frac{\gamma^j (\theta) - \phi^j (\theta)}{\beta (\theta)} = 0.
$$

Collecting terms we get

$$
1 + \sum_{m=0}^{M} H^j (\theta, m) \left\{ \frac{h}{\beta_i (\theta_i)} \left[ \lambda (m) m - \lambda (m + 1) (M - m) \right] \right\}
$$

$$
+ \sum_{m=0}^{M} L^j (\theta, m) \left\{ \frac{\eta (m)}{\beta_i (\theta_i)} \left[ \lambda (m) m l - \lambda (m + 1) \{(M - m) l + (h - l)\} \right] \right\}
$$

$$
- \Lambda cn + \frac{\gamma^j (\theta) - \phi^j (\theta)}{\beta (\theta)} = 0.
$$
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 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

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

Furthermore

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

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 (A19) 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

\[
Z^j(\theta^i) = \begin{cases} 
h \frac{\max\{0, G_m(1)\}}{\beta_m(M - m)} & \text{if } \theta^j_i = h \\
\theta^j_i & \text{if } \theta^j_i = l \quad \text{and} \quad \theta^j_i = h \quad \text{for exactly } m \text{ goods } k \in J
\end{cases}
\]

The point with this formulation is that \( \{ Z^j(\theta^i) \}_{i=1}^{n} \) is a sequence of i.i.d. random variables. It is easy to see that \( Z^j(\theta^i) \) 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 $\langle \rho_n, \eta_n \rangle$ and the multiplier on the resource constraint will be denoted by $\Lambda_n$.

(Proof of Lemma 14 continued): Proof. Consider a mechanism that solves (22) first. As $Z^j(\theta^i) \in [0, h]$ for every $\theta^i \in \Theta$ it follows that for any pair $(\theta', \theta''_i)$ we have that

$$E [\rho^j_n(\theta) | \theta'] - E [\rho^j_n(\theta) | \theta''] \leq \Pr \left[ 1 + \Lambda_n \left( h + \sum_{k \neq i} Z^j(\theta_k) \right) - cn \geq 0 \right] - \Pr \left[ 1 + \Lambda_n \left( 0 + \sum_{k \neq i} Z^j(\theta_k) \right) - cn > 0 \right]$$

$$= \Pr \left[ \frac{1 - cn}{\Lambda_n} - h \leq \sum_{k \neq i} Z^j(\theta_k) \leq \frac{1 - cn}{\Lambda_n} \right]$$

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

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

where $k_n \equiv \frac{1 - cn}{\Lambda_n \sigma \sqrt{n - 1}} - \frac{h}{\sigma \sqrt{n - 1}} - \frac{E Z^j(\theta_i)}{\sigma \sqrt{n - 1}}$. As $\sigma$ is finite, $E \{ Z^j(\theta_k) - EZ^j(\theta_i) \} = 0$ and $\{ Z^j(\theta^i) \}_{i=1}^n$ is an i.i.d. sequence, we have that the central limit theorem is applicable, so that $\frac{\sum_{k \neq i} Z^j(\theta_k) - EZ^j(\theta_i)}{\sigma \sqrt{n - 1}}$ is asymptotically distributed in accordance with 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

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

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

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

$$\leq \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon}^{\varepsilon} \exp \left( \frac{y^2}{2} \right) dy + \varepsilon \to 0$$

as $\varepsilon \to 0$. The result follows. For the case with $\{ \rho_n \}_{n=1}^\infty$ being a sequence from the solution to (11) we replace $Z^j(\theta_i)$ with

$$Z^j_n(\theta_i) = \begin{cases} h & \text{if } \theta^j_i = h \\ \max \{ 0, G_m(\Phi_n) \} & \text{if } \theta^j_i = l \text{ and if } \theta^k_i = h \text{ for exactly } m \text{ goods } k \neq j \end{cases}$$

By choice of subsequences such that $G_m(\Phi_n) \to G^*_m$ we can approximate $\max \{ 0, G_m(\Phi_n) \}$ with $\max \{ 0, G^*_m \}$ for $m \to \infty$. The rest of the argument follows the one above step by step.
Proof of Proposition 4.

Proof. [PART 1] For contradiction let there be a subsequence of feasible mechanisms with lim_{n→∞} E [ρ^j_n (θ) | m, θ^j_i = h] = ρ^* and lim_{n→∞} η_n (m) = η^*, where m is the threshold number of high goods in the sense of Proposition 3 which without loss is taken to be constant along the sequence by choice of converging subsequence. We use the following notations below:

- \( E [ρ^j_n (θ) | m, θ^j_i = h] \) denotes the conditional provision probability for good \( j \) perceived by agent \( i \) given that \( θ^j_i = h \) and \( θ^k_i = h \) for exactly \( m \) goods \( k \in J \) (including \( j \)). Notice that this probability is the same for all goods by Proposition 2.

- \( E [ρ^j_n (θ) | m, θ^j_i = l] \) denotes the conditional provision probability for good \( j \) perceived by agent \( i \) given that \( θ^j_i = l \) and \( θ^k_i = h \) 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 \( m \in \{0, \ldots, M\} \) 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,

\[
\begin{align*}
 mh \left[ ρ^j_n (θ) | m, θ^j_i = h \right] - t_n (m) & \quad \text{if } m < \bar{m} \\
 \bar{m} h \left[ ρ^j_n (θ) | m, θ^j_i = h \right] + (M - \bar{m}) \eta_n (\bar{m}) E \left[ ρ^j_n (θ) | m, θ^j_i = l \right] l - t_n (\bar{m}) & \quad \text{if } m = \bar{m} \\
 mh \left[ ρ^j_n (θ) | m, θ^j_i = h \right] + (M - m) E \left[ ρ^j_n (θ) | m, θ^j_i = l \right] l - t_n (m) & \quad \text{if } m > \bar{m}
\end{align*}
\]

Hence,

\[ t_n (m) \leq mh \left[ ρ^j_n (θ) | m, θ^j_i = h \right] \quad \text{for } θ_i \text{ with } m < \bar{m}, \quad (A20) \]

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

\[ t_n (\bar{m}) \leq \bar{m} h \left[ ρ^j_n (θ) | m, θ^j_i = h \right] + (M - \bar{m}) \eta_n (\bar{m}) E \left[ ρ^j_n (θ) | m, θ^j_i = l \right] l \quad \text{for } θ_i \text{ with } m = \bar{m}, \quad (A21) \]

again immediately from the participation constraint (or the downwards adjacent incentive constraint combined with (A20)). Using (A21) we find

\[
\begin{align*}
 (\bar{m} + 1) h E [ρ^j_n (θ) | \bar{m} + 1, θ^j_i = h] + (M - (\bar{m} + 1)) E [ρ^j_n (θ) | \bar{m} + 1, θ^j_i = l] l - t_n (\bar{m} + 1) & \\
 \geq \bar{m} h E [ρ^j_n (θ) | \bar{m}, θ^j_i = h] + \eta_n (\bar{m}) E [ρ^j_n (θ) | \bar{m}, θ^j_i = l] l ([M - (\bar{m} + 1)] l + h) - t_n (\bar{m}) \quad / (A21) \quad / & \\
 \geq \eta_n (\bar{m}) E [ρ^j_n (θ) | \bar{m}, θ^j_i = l] ([M - (\bar{m} + 1)] l + h) - (M - \bar{m}) \eta_n (\bar{m}) E [ρ^j_n (θ) | \bar{m}, θ^j_i = l] l & \\
 = \eta_n (\bar{m}) E [ρ^j_n (θ) | \bar{m}, θ^j_i = l] (h - l).
\end{align*}
\]

\(^{12}\)Strictly speaking, we only impose the participation constraint on type I. However, the downwards adjacent incentive constraint together with the participation constraint for type I imply that all participation constraints hold.
Hence, 

\[ t_n (\bar{m} + 1) \leq (\bar{m} + 1) h E \left[ \rho_n^j (\theta) | \bar{m} + 1, \theta_i^j = 1 \right] + (M - \bar{m} - 1) E \left[ \rho_n^j (\theta) | \bar{m} + 1, \theta_i^j = 1 \right] (A22) \]

\[ - \eta_n (\bar{m}) E \left[ \rho_n^j (\theta) | \bar{m}, \theta_i^j = 1 \right] (h - l). \]

Finally, for every for every \( m > \bar{m} + 1 \) we have that (A22) implies that

\[ m h E \left[ \rho_n^j (\theta) | m, \theta_i^j = 1 \right] + (M - m) E \left[ \rho_n^j (\theta) | m, \theta_i^j = 1 \right] l - t_n (m) \]

\[ \geq (\bar{m} + 1) h E \left[ \rho_n^j (\theta) | \bar{m} + 1, \theta_i^j = 1 \right] E \left[ \rho_n^j (\theta) | \bar{m} + 1, \theta_i^j = 1 \right] ((M - m) l + (m - \bar{m} - 1) h) - t_n (\bar{m} + 1) \]

\[ \geq E \left[ \rho_n^j (\theta) | \bar{m} + 1, \theta_i^j = 1 \right] ((m - \bar{m} - 1) (h - l) + \eta_n (\bar{m}) E \left[ \rho_n^j (\theta) | \bar{m}, \theta_i^j = 1 \right] (h - l), \]

and, hence

\[ t_n (m) \leq m h E \left[ \rho_n^j (\theta) | m, \theta_i^j = 1 \right] + (M - m) E \left[ \rho_n^j (\theta) | m, \theta_i^j = 1 \right] l \]

\[ - E \left[ \rho_n^j (\theta) | \bar{m} + 1, \theta_i^j = 1 \right] ((m - \bar{m} - 1) (h - l)) - \eta_n (\bar{m}) E \left[ \rho_n^j (\theta) | \bar{m}, \theta_i^j = 1 \right] (h - l). \]  

(A23)

From from Lemma 14 we have that

\[ \lim_{n \to \infty} E \left[ \rho_n^j (\theta) | m, \theta_i^j = 1 \right] = \lim_{n \to \infty} E \left[ \rho_n^j (\theta) | m, \theta_i^j = 1 \right] = \lim_{n \to \infty} E \left[ \rho_n^j (\theta) | \bar{m}, \theta_i^j = 1 \right] = \lim_{n \to \infty} E \left[ \rho_n^j (\theta) \right] = \rho^*. \]

(A24)

Combining (A24) with (A20), (A21), (A22) and (A23) it follows that, for every \( \varepsilon > 0 \) there is some \( N \) such that, whenever \( n \geq N \),

1. \( t_n (m) \leq \rho^* m h + \varepsilon \) if \( m < \bar{m} \)

2. \( t_n (\bar{m}) \leq \rho^* [\bar{m} h + (M - \bar{m}) \eta^* l] + \varepsilon \)

3. \( t_n (\bar{m} + 1) \leq \rho^* [\bar{m} + 1 + (M - \bar{m} - 1) l - \eta^* (h - l)] + \varepsilon \)

4. \( t_n (m) \leq \rho^* [(m + 1) h + (M - \bar{m} - 1) l - \eta^* (h - l)] + \varepsilon \) if \( m > \bar{m} + 1 \).
Summing over \( m \) we find that the expected per capita transfer revenue satisfies:

\[
\sum_{m=0}^{M} \beta_m t_n (m) \leq \rho^* \left[ \left( \sum_{m=m+1}^{M} \beta_m \right) \{(\tilde{m}+1)h + (M - \tilde{m} - 1)l \} + h \sum_{m=0}^{\tilde{m}} \beta_m m \right] (A25)
\]

\[
+ \rho^* \eta \left[ \beta_{\tilde{m}} (M - \tilde{m}) - \left( \sum_{m=m+1}^{M} \beta_m \right) (h - l) \right] + \varepsilon
\]

\[
= \rho^*(1 - \eta^*) \left[ \left( \sum_{m=m+1}^{M} \beta_m \right) \{(\tilde{m}+1)h + (M - \tilde{m} - 1)l \} + h \sum_{m=0}^{\tilde{m}} \beta_m m \right]
\]

\[
+ \rho^* \eta \left[ \beta_{\tilde{m}} (M - \tilde{m}) + \left( \sum_{m=m+1}^{M} \beta_m \right) (\tilde{m}h + (M - \tilde{m})l) + h \sum_{m=0}^{\tilde{m}} \beta_m m \right] + \varepsilon
\]

\[
= \rho^*(1 - \eta^*) \left[ \left( \sum_{m=m+1}^{M} \beta_m \right) \{(\tilde{m}+1)h + (M - \tilde{m} - 1)l \} + h \sum_{m=0}^{\tilde{m}-1} \beta_m m \right] + \varepsilon
\]

\[
= \rho^* \{(1 - \eta^*) R (\tilde{m} + 1) + \eta^* R (\tilde{m}) \} + \varepsilon.
\]

Moreover, for any \( \varepsilon \), we can find \( N \) such that \( M \sum_{\theta \in \Theta} \beta (\theta) \rho_n^\dagger (\theta) c \geq \rho^* M c - \varepsilon \), which in conjunction with (A25) means that

\[
\sum_{i=1}^{\Theta} \sum_{\theta_i \in \Theta} \beta_i (\theta_i) t_{in} (\theta_i) - \sum_{\theta \in \Theta^n} \beta (\theta) \sum_{j=1}^{M} \rho_n^\dagger (\theta) c n = \sum_{m} \beta_m t_n (m) - M \sum_{\theta \in \Theta^n} \beta (\theta) \rho_n^\dagger (\theta) c \leq \rho^* \{(1 - \eta^*) R (\tilde{m} + 1) + \eta^* R (\tilde{m}) - M c \} + 2\varepsilon.
\]

\( \leq 0 \) by hypothesis

We conclude that corresponding to any \( \rho^* > 0 \) there exists \( \varepsilon > 0 \) such that the right hand side of (A26) is negative. Hence, the budget constraint (14) is violated for large \( n \) along any sequence with non-vanishing provision probability, and it follows that \( \lim_{n \to \infty} E \left[ \rho_n^\dagger (\theta) \right] = 0 \) for any convergent subsequence.

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

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

and where the transfer rule is

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

41
for
\[ \eta^* = \frac{R(m^* + 1) - cM}{R(m^* + 1) - R(m^*)}. \] (A27)

It is easy to check that truth-telling is incentive compatible and individually rational. The per capita expected transfer is
\[ \sum_{m=0}^{M} \beta_m \tilde{t}_n(m) = h \sum_{m=0}^{m^*-1} \beta_m m + \beta_{m^*} [m^* h + (M - m^*) \eta^* l] \]
(A28)
\[ + \left( \sum_{m=m^*+1}^{M} \beta_m \right) [(m^* + 1) h + (M - m^*) l - \eta^* (h - l)] \]
\[ = (1 - \eta^*) \left\{ h \sum_{m=0}^{m^*-1} \beta_m m + \left( \sum_{m=m^*+1}^{M} \beta_m \right) [(m^* + 1) h + (M - m^*) l] \right\} \]
\[ + \eta^* \left[ h \sum_{m=0}^{m^*-1} \beta_m m + \left( \sum_{m=m^*}^{M} \beta_m \right) [m^* h + (M - m^*) l] \right] \]
\[ = (1 - \eta^*) R(m^* + 1) + \eta^* R(m^*) \]
\[ \frac{cM - R(m^*)}{R(m^* + 1) - R(m^*)} R(m^* + 1) + \frac{R(m^* + 1) - cM}{R(m^* + 1) - R(m^*)} R(m^*) = cM. \]

The feasibility constraint thus holds exactly and we conclude that \((\tilde{\rho}_n, \tilde{\eta}_n, \tilde{t}_n)\) is incentive feasible for every \(n\). The associated social surplus is
\[ S^*_n(\theta) = \sum_{j=1}^{M} \left\{ h \sum_{m=1}^{M} H^j(\theta, m) + l \left[ \eta^* L^j(\theta, m^*) + \sum_{m=m^*+1}^{M} L^j(\theta, m) \right] \right\} - cM, \] (A29)

where the notations \(H^j(\theta, m)\) and \(L^j(\theta, m)\) were introduced and explained in the paragraph before Proposition 6. By Proposition 3 we know that the surplus maximizing mechanism is also characterized by some threshold \(m^{**}\) and some inclusion probability given \(m = m^{**}\) denoted by \(\eta^*\), so that the associated social surplus is given by
\[ S^*_n(\theta) = \sum_{j=1}^{M} \rho_j^j(\theta) \left[ h \sum_{m=1}^{M} H^j(\theta, m) + l \left\{ \eta^{**} L^j(\theta, m^{**}) + \sum_{m=m^{**}+1}^{M} L^j(\theta, m) \right\} - c \right] \] (A30)

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 \(\frac{M - m}{M} \beta_m\), which immediately implies that
\[ \lim_{n \to \infty} \Pr \left[ \left| \frac{S^*_n(\theta)}{n} - h \sum_{m=0}^{M} m \beta_m - l \left[ \eta^* (M - m^*) \beta_{m^*} + \sum_{m=m^*+1}^{M} (M - m) \beta_m \right] + cM \right| > \varepsilon \right] = 0 \] (A31)

for any \(\varepsilon > 0\). Next, note that
\[ \sum_{m=1}^{M} H^j(\theta, m) h + \sum_{m=0}^{M} L^j(\theta, m) \frac{1}{\beta_m (M - m)} \max \{0, G_m(\Phi)\} - cn \] (A32)
converges in probability to
\[
\sum_{m=1}^{M} \frac{m}{M} \beta_m h + \sum_{m=0}^{M} \frac{M-m}{M} \beta_m \frac{1}{\beta_m (M-m)} \max \{0, G_m(\Phi)\} - cn
\]

(A33)

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 (A32) is positive. Now:

1. Suppose that \(m^{**} > m^*\). Then, \(R(m^{**}) - cM < 0\) and an argument identical with part 1 establishes that \(\lim_{n \to \infty} E[\rho^i_n(\theta)] \to 0\).

2. Suppose that \(m^{**} < m^*\) and \(R(m^{**}) - cM < 0\). Then an argument identical with part 1 establishes that \(\lim_{n \to \infty} E[\rho^i_n(\theta)] \to 0\).

3. Suppose that \(m^{**} < m^*\) and \(R(m^{**}) - cM \geq 0\). Then, the expression in (A33) is strictly positive and \(\lim_{n \to \infty} E[\rho^i_n(\theta)] \to 1\). The social surplus must therefore converge towards the same expression as in (A31), but with \(m^*\) replaced by \(m^{**}\). But, if \(m^{**} < m^*\) surplus is smaller because there are more exclusions.

4. Suppose that \(m^{**} = m^*\). If \(\tilde{h}_n(m^*) \to \eta^{**} < \eta^*\) an argument identical with part 1 establishes that \(\lim_{n \to \infty} E[\rho^i_n(\theta)] \to 0\). If \(\tilde{h}_n(m^*) \to \eta^{**} > \eta^*\) we have that \(\lim_{n \to \infty} E[\rho^i_n(\theta)] \to 1\) and, again, social surplus converges towards the same expression as in (A31), but with \(\eta^*\) replaced by \(\eta^{**}\). Again, surplus is smaller if \(\eta^{**} > \eta^*\) since exclusions are higher.

Proof of Proposition 5.

Proof. To prove Proposition 5 we add the (non-adjacent) constraint that the highest type should not have an incentive to pretend to be the lowest type.

\[
0 \leq \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \sum_{j=1}^{M} \rho^i (\theta_{-i}, h) h - t_i(h) \]

(A34)

\[
- \left[ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \sum_{j=1}^{M} \rho^i (\theta_{-i}, l) \eta^i (l) h - t_i (l) \right]
\]

to program (11). Let \(\psi\) denote the multiplier on constraint (A34). The optimality conditions with respect to \(t\) are then

\[
-2\lambda \psi + \Lambda \beta_2 = 0
\]

(2)

\[
\lambda (2) - \lambda (1) + \Lambda \frac{1}{2} \beta_1 = 0
\]

(1) + (0) + \Lambda \beta_0 = 0,
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 \([A34]\). It follows immediately that:

**Lemma A3** Consider \(m = 2\). Then, there are three possibilities in the solution the the program where \([A34]\) has been added onto \((11)\):

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 \([A34]\) 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 \(2 \lambda(2) - \psi = \Lambda \beta_2, \lambda(1) \in [\Lambda \frac{1}{2} \beta_1, \Lambda \frac{1}{2} (\beta_1 + \beta_2)]\) and \(\lambda(0) = \Lambda\).

We now consider these possibilities in turn:

**Lemma A4** Suppose that \(\frac{\beta_1}{2} < \frac{\beta_0 \beta_2}{1 - \beta_6}\). Then \([A34]\) binds.

**Proof.** If \([A34]\) does not bind, Lemma A3 implies that the solution is unchanged from \((11)\). Simplifying the inclusion rule \((19)\) for the special case with \(M = 2\), we have that

\[
\eta(1) \equiv \begin{cases} 
0 & \text{if } G_1(\Phi) < 0 \\
\begin{array}{ll}
0 & \text{if } G_0(\Phi) < 0 \\
1 & \text{if } G_0(\Phi) > 0 
\end{array} & \text{if } G_1(\Phi) = 0 \\
1 & \text{if } G_0(\Phi) > 0
\end{cases},
\]

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

where

\[
G_1(\Phi) = \beta_1 l - \Phi (h - l) \beta_2
\]

\[
G_0(\Phi) = \beta_0 2 l - \Phi (h - l) (\beta_1 + \beta_2)
\]

and the provision rule \((21)\) simplifies to

\[
\rho^j(\theta) = \begin{cases} 
0 & \text{if } [H^j(\theta, 1) + H^j(\theta, 2)] h + \frac{L^j(\theta, 0)}{2 \beta_0} \max \{0, G_0(\Phi)\} + \frac{L^j(\theta, 1)}{\beta_1} \max \{0, G_1(\Phi)\} - cn < 0 \\
\begin{array}{ll}
0 & \text{if } [H^j(\theta, 1) + H^j(\theta, 2)] h + \frac{L^j(\theta, 0)}{2 \beta_0} \max \{0, G_0(\Phi)\} + \frac{L^j(\theta, 1)}{\beta_1} \max \{0, G_1(\Phi)\} - cn = 0 \\
1 & \text{if } [H^j(\theta, 1) + H^j(\theta, 2)] h + \frac{L^j(\theta, 0)}{2 \beta_0} \max \{0, G_0(\Phi)\} + \frac{L^j(\theta, 1)}{\beta_1} \max \{0, G_1(\Phi)\} - cn > 0
\end{array} & \text{if } [H^j(\theta, 1) + H^j(\theta, 2)] h + \frac{L^j(\theta, 0)}{2 \beta_0} \max \{0, G_0(\Phi)\} + \frac{L^j(\theta, 1)}{\beta_1} \max \{0, G_1(\Phi)\} - cn < 0
\end{cases}
\]
Notice that
\[
\frac{G_1(\Phi)}{\beta_1} = l - \frac{\Phi(h-l)\beta_2}{\beta_1} < l - \frac{\Phi(h-l)\beta_2}{2\beta_0} = \frac{G_0(\Phi)}{2\beta_0},
\]

implying that \( \eta(1) \leq \eta(0) \) and \( \rho^i(\theta_{-i},\theta_i) \leq \rho^j(\theta_{-i},\theta_j) \) 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 no, 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

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

utility from truth-telling for \( hh \)

/binding DAIC for \( hh \) / 
\[
= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[ \rho^1(\theta_{-i},hl) h + \rho^2(\theta_{-i},hl) \eta(1)h \right] - t(1)
\]

= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[ \rho^1(\theta_{-i},hl) h + \rho^2(\theta_{-i},hl) \eta(1)h \right] - t(1)

+ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \rho^2(\theta_{-i},hl) \eta(1)(h-l)

/binding DAIC for \( hl \) / 
\[
= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[ \rho^1(\theta_{-i},ll) \eta(0) h + \rho^2(\theta_{-i},ll) \eta(0)h \right] - t(0)
\]

= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[ \rho^1(\theta_{-i},ll) \eta(0) h + \rho^2(\theta_{-i},ll) \eta(0)h \right] - t(0)

+ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \rho^2(\theta_{-i},hl) \eta(1)(h-l)

utility if \( hh \) announces \( ll \)

+ \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[ \rho^2(\theta_{-i},hl) \eta(1) - \rho^2(\theta_{-i},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}

\]

implying that (A34) binds or is violated, a contradiction.

**Lemma A5** Suppose that \( \frac{\beta_1}{2} < \frac{\beta_0\beta_2}{1-\beta_0} \). Then the downwards adjacent constraints for type \( hh \) binds.

**Proof.** For contradiction, assume that

\[
\sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i}(\theta_{-i}) \left[ \rho^1(\theta_{-i},hh) + \rho^2(\theta_{-i},hh) \right] h - t(2)
\]

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

(A35)
With slack in the downwards adjacent constraints for \( hh \) and a binding \( (A34) \) the optimality conditions with respect to \( \eta^2_i (hl) \) may be written as

\[
\sum_{\theta_{-i} \in \Theta_{-i}} \beta (\theta_{-i}, hl) \rho^2 (\theta_{-i}, hl) l + \lambda (1) \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \rho^2 (\theta_{-i}, hl) l + \gamma^2_i (hl) - \phi^2_i (hl) = 0,
\]

implying that \( \eta^2_i (hl) = \eta^1_i (lh) = \eta (1) = 1 \) for the simple reason that 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 11) the optimality conditions with respect to \( \rho^1 (\theta) \) are given by the complementary slackness conditions together with

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

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 (A36) that \( \rho^1 (\theta_{-i}, lh) \geq \rho^1 (\theta_{-i}, ll) \) for every \( \theta_{-i} \) 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,

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

utility from truth-telling for \(hh\)

/\[A34\] binds/

\[
= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left[ \rho^1 (\theta_{-i}, ll) \eta (0) h + \rho^2 (\theta_{-i}, ll) \eta (0) h \right] - t (0)
\]

utility if \(hh\) announces \(ll\)

\[
= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left[ \rho^1 (\theta_{-i}, ll) \eta (0) l + \rho^2 (\theta_{-i}, ll) \eta (0) h \right] - t (0)
\]

utility if \(lh\) 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/\]

\[
= \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \left[ \rho^1 (\theta_{-i}, lh) \eta (1) h + \rho^2 (\theta_{-i}, lh) h \right] - t (1)
\]

utility from truth-telling for \(lh\)

\[
+ (h - l) \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \rho^1 (\theta_{-i}, ll) \eta (0)
\]

utility if \(hh\) announces \(lh\)

\[
+ (h - l) \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \rho^1 (\theta_{-i}, ll) \eta (0) - (h - l) \sum_{\theta_{-i} \in \Theta_{-i}} \beta_{-i} (\theta_{-i}) \rho^1 (\theta_{-i}, lh) \eta (1),
\]

\(\leq 0\)

which contradicts (A35).

\(\text{(Proof of Proposition 5, continued):}\) We conclude that all constraints must bind. This immediately implies that \(\eta (1) = \eta (0)\). Assuming that \(l < c < ah\) 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

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

which coincides with the provision rule if good \(j\) is provided as a stand alone good. Since the inclusion rule is also as when only good \(i\) is provided, the result follows.