Optimal Provision of Multiple Excludable Public Goods*

Hanming Fang[†] Peter Norman[‡]

This Version: May 2010

Abstract

This paper studies the optimal provision mechanism for multiple excludable public goods. For a class of problems with symmetric goods and binary valuations, we show that the optimal mechanism involves bundling if a regularity condition, akin to a hazard rate condition, on the distribution of valuations is satisfied. Relative to separate provision mechanisms, the optimal bundling mechanism may increase the asymptotic provision probability of socially efficient public goods from zero to one, and decreases the extent of use exclusions. If the regularity condition is violated, the optimal solution replicates the separate provision outcome for the two-good case.

Keywords: Public Goods Provision; Bundling; Exclusion
JEL Classification Number: H41.

*This paper substantially generalizes and supersedes Fang and Norman (2003), a manuscript that was circulated under the title "An Efficiency Rationale for Bundling of Public Goods," and Fang and Norman (2006a). Both considered only the case of two 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. We are most grateful to two anonymous referees and Co-editor Robert Porter for comments that significantly improved the paper. All remaining errors are ours.

†Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104-6297. Email: hanming.fang@econ.upenn.edu

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

1 Introduction

This paper studies the constrained efficient mechanism for multiple excludable public goods. By "excludable public goods," we mean goods that allow the provider to exclude consumers from access, but, nevertheless, are fully non-rival in consumption. There are many real-world examples of excludable public goods: cable TV, electronic journals, computer software and digital music files are all almost perfect examples. Importantly, most of these goods are provided in bundles. For example, cable TV customers purchase most programming as components of a few big bundles; access to electronic libraries are usually provided through site licences that allow access to every issue of every journal in the library; and digital music files, computer software and other digital files are commonly sold in a bundled format.

Most of the economics literature focuses on the implications of bundling on the revenue of *producers*.[1] To the best of our knowledge, a normative benchmark addressing the pros and cons of bundling for *consumers* and *social welfare,* while still explicitly incorporating the non-rival nature of these goods, does not yet exist. Our paper aims to fill this gap and is a step toward a better understanding of the welfare consequences of bundling of non-rival goods.[2]

We consider a simple model with M excludable public goods, satisfying standard (but strong) separability assumptions on both the supply and demand side. On the demand side, a consumer is described by a vector of single-good valuations, and her willingness to pay for a bundle is the sum of the valuations of the goods included in the bundle; on the supply side, the cost of providing any good is independent of which other goods are also provided. Under these separability assumptions, the first-best benchmark is to provide good j whenever the average valuation of the good exceeds the per capita cost of provision, and to exclude no consumer from usage. There is thus no role for either bundling or use exclusion if there is perfect information.

If, instead, preferences are private information, as we assume in this paper, then consumers must be given appropriate incentives to truthfully reveal their willingness to pay. Together with balanced budget and participation constraints, it is impossible to implement the non-bundling full information first-best solution. Then, as we show in this paper, bundling is often useful because it facilitates revenue extraction from the consumers and thus relaxes a binding budget constraint.

We consider an environment with n agents and M excludable public goods where valuations for each good takes only two values, and the costs and valuation distributions are symmetric for the M goods. We obtain a full characterization of the constrained efficient mechanism. The solution is rather striking. When there are a large number of agents, i.e., as $n \to +\infty$, the optimal mechanism either provides all goods with probability close to one, or provides all goods with probability close to zero. Which of these two scenarios applies depends on whether a *monopolist profit maximizer* that provides the goods for sure could break even. If a regularity condition on the valuation distribution – which can be interpreted much like a hazard rate condition – is satisfied, then the optimal mechanism also prescribes a simple rule for user access to the public goods that are provided. The rule can be described as follows. All agents will fall into one of three groups depending on their numbers of high-valuation goods. A first group consists

[1] See, e.g., Adams and Yellen (1976), McAfee, McMillan and Whinston (1989), Armstrong (1996, 1999), Bakos and Brynjolfsson (1999), Fang and Norman (2006b) and Vincent and Manelli (2006).

[2] For the cable TV industry, see Crawford (2008), Cullen and Crawford (2007), Crawford and Yorokoglu (2009) and Yorokoglu (2009) for some recent studies on the welfare effects of *à la carte* pricing on the consumers. These studies assume that the provision of the programs will not be affected by the *à la carte* pricing regulation. See Section 5 for more discussions on the implications of our results on this issue.

of all agents whose numbers of high-valuation goods strictly exceed a threshold. These agents are given access to the grand bundle consisting of all goods. A second group consists of all agents whose numbers of high-valuation goods are strictly lower than the threshold. For these agents access is granted only to the goods for which they have high valuations. Finally, there is a third group, consisting of those with exactly the threshold-level number of high-valuation goods. These agents receive random access to their low-valuation goods and full access to their high-valuation goods.[3]

If the regularity condition discussed above is violated, it is in general difficult to identify the binding constraints. However, the two-good case is both tractable and instructive. For the two-good case, violations of the regularity condition can then be interpreted as when the valuations of the two goods are "too positively correlated." This invalidates the approach we use for solving the regular case. When the valuations of the two goods are "too positively correlated,", there are too few "mixed types" (those with one high and one low valuation) to justify giving them preferential treatment in terms of accessing their low-valuation good, leading to a violation of the monotonicity that is required for the standard approach. Surprisingly, we show that the optimal solution for this case is in fact identical to the solution when both goods must be separately provided.

The tractability of the multidimensional mechanism design problem studied in this paper comes from the fact that types are naturally ordered in terms of the number of high-valuation goods. This allows us to exploit some important similarities with unidimensional problems that help us determine which constraints are likely to bind. In addition, for unidimensional problems it is known that maximizing social surplus subject to budget and participation constraints leads to a Lagrangian characterization that can be interpreted as a compromise between profit and welfare maximization (see Hellwig 2003, and Norman 2004). In Hellwig (2003) and Norman (2004), incentive compatibility, balanced budget and participation constraints can be combined into a single constraint that the monopolistic provider's profit must be non-negative. This cannot be done in the multidimensional model considered in this paper. However, Lagrange multipliers of "adjacent constraints" are linked in a way that allows for an analogous characterization.

The remainder of the paper is structured as follows. Section 2 describes the model and the class of simple, anonymous, symmetric mechanisms we will consider without loss of generality. Section 3 characterizes the optimal mechanism for the regular case. In Section 4 we use some special cases to better interpret the characterization in Section 3. In particular, we characterize the optimal mechanism when the regularity condition is violated in the two-good case. Section 5 contains a brief discussion of the relevance of our analysis with respect to anti-trust issues. Appendix A contains the proofs of some of the key results in Section 3.[4]

2 The Model

The Environment. There are M *excludable* and indivisible public goods, labeled by $j \in \mathcal{J} = \{1, ..., M\}$ and n agents, indexed by $i \in \mathcal{I} = \{1, ..., n\}$. The cost of providing good j, denoted $C^j(n)$, is independent of which of the other goods are provided. We assume that $C^j(n) = cn$ for all $j \in \mathcal{J}$ where $c > 0$. Notice that $C^j(n)$ depends on n, the number of *agents* in the economy, and not on the number of *users*. This

[3] Note, however, the third group of agents will be rather small when there are many goods.

[4] An online appendix available at the authors' websites contains all the omitted proofs.

assumption captures the fully non-rival nature of the public goods.[5]

Agent i is described by a valuation for each good $j \in \mathcal{J}$. Her type is given by a vector $\theta_i = \left(\theta_i^1, ..., \theta_i^M\right) \in \Theta$. We assume that the valuation for each good j is now either high or low, so that $\theta_i^j \in \{l, h\}$ for each $i \in \mathcal{I}$ and $j \in \mathcal{J}$. Thus, $\Theta \equiv \{l, h\}^M$. Agents' preferences are represented by the utility function

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

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

An agent's type θ_i is her private information. The probability of any type $\theta_i \in \Theta \equiv \{l, h\}^M$ is denoted $\beta(\theta_i)$. Unlike much of the bundling literature, we allow the valuations for different goods for a given agent to be correlated across goods. However, types are independent across agents, which allows us to denote the probabilities of *type profile* $\theta \equiv (\theta_1, ..., \theta_n) \in \Theta^n$ and $\theta_{-i} \equiv (\theta_1, .., \theta_{i-1}, \theta_{i+1}, ...\theta_n) \in \mathbf{\Theta}_{-i} \equiv \Theta^{n-1}$ by $\boldsymbol{\beta}(\theta) = \Pi_{i=1}^n \beta(\theta_i)$ and $\boldsymbol{\beta}_{-i}(\theta_{-i}) = \Pi_{i' \neq i} \beta(\theta_{i'})$ respectively. For simplicity, we assume that the valuations of the M goods are *symmetrically* distributed in the sense that $\theta_i = \left(\theta_i^1, \cdots, \theta_i^M\right)$ is an *exchangeable* random vector, that is, we assume that $\beta(\theta_i) = \beta\left(\theta_i'\right)$ if θ_i' is a permutation of θ_i.

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

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

By the revelation principle, we only consider direct mechanisms for which truth-telling is a Bayesian Nash equilibrium. A pure direct mechanism maps Θ^n onto A. In general, we should consider *direct randomized mechanisms*, which can be represented analogously to the representation of mixed strategies in Aumann (1964).[6] However, it is shown in Fang and Norman (2006a, Propositions 1 and 2) that, in our environment, it is *without loss of generality* to consider a class of *simple, anonymous and symmetric* mechanisms.[7] Specifically, we say that a mechanism is *simple* if it can be expressed as a $(2M+1)$-tuple $g = (\rho, \eta, t) \equiv (\{\rho^j\}_{j \in \mathcal{J}}, \{\eta^j\}_{j \in \mathcal{J}}, t)$ such that for each good $j \in \mathcal{J}$, $\rho^j : \Theta^n \to [0,1]$ is the provision rule for good j; $\eta^j : \Theta \to [0,1]$ is the inclusion rule for good j; and $t : \Theta \to \mathrm{R}$ is the transfer rule, same for all agents. A simple mechanism is also *anonymous* if for every $j \in \mathcal{J}$, $\rho^j(\theta) = \rho^j\left(\theta'\right)$ for every

[5]The fact that the per capita cost of the public good is constant independent of n enables us to analyze large economies without making the public goods a "free lunch" in the limit.

[6]Specifically, let $\Xi \equiv [0,1]$ and think of $\vartheta \in \Xi$ as the outcome of a fictitious lottery where, without loss of generality, ϑ is uniformly distributed and independent of θ. A random direct mechanism is then a measurable mapping $\mathcal{G} : \Theta^n \times \Xi \to A$, which can be decomposed as $(\{\zeta^j(\theta, \vartheta)\}_{j \in \mathcal{J}}, \{\omega^j(\theta, \vartheta)\}_{j \in \mathcal{J}}, \tau)$ where $\zeta^j : \Theta^n \times \Xi \to \{0,1\}$ is the provision probability for good j when the type profile announcement is θ and the lottery outcome realization is ϑ; analogously, $\omega^j \equiv \left(\omega_1^j, ..., \omega_n^j\right) : \Theta^n \times \Xi \to \{0,1\}^n$ is the *inclusion rule* for good j; and $\tau \equiv (\tau_1, ..., \tau_n) : \Theta^n \to \mathrm{R}^n$ is the *cost-sharing rule,* where $\tau_i(\theta)$ is the transfer from agent i to the mechanism designer given announced valuation profile θ. Note that in principle, transfers could also be random, but the pure cost-sharing rule is without loss of generality due to risk neutrality.

[7]The simplification results from the symmetries we assumed in our environment, namely, cost functions are identical and the valuations are exchangeable across goods.

$(\theta, \theta') \in \Theta^n \times \Theta^n$ such that θ' can be obtained from θ by permuting the indices of the agents. Finally, we say that a simple anonymous mechanism is also *symmetric* if for every θ and every permutation $P : \mathcal{J} \to \mathcal{J}$ with inverse P^{-1}, we have $\rho^{P^{-1}(j)}\left(\theta^P\right) = \rho^j(\theta)$ and $\eta^{P^{-1}(j)}\left(\theta_i^P\right) = \eta^j(\theta_i)$ for every $j \in \mathcal{J}$, and $t\left(\theta_i^P\right) = t(\theta_i)$, where $\theta_i^P = \left(\theta_i^{P^{-1}(1)}, \theta_i^{P^{-1}(2)} ..., \theta_i^{P^{-1}(M)}\right)$ denotes the permutation of agent i's type by changing the role of the goods in accordance to P, and $\theta^P \equiv \left(\theta_1^P, ..., \theta_n^P\right)$ denotes the valuation profile obtained when the role of the goods is changed in accordance to P for every $i \in \mathcal{I}$.

The class of simple, anonymous and symmetric mechanisms have three properties. First, the inclusion and transfer rules are the same for all agents, and the inclusion and transfer rules for any agent i are independent of θ_{-i}. Second, all agents are treated symmetrically in the transfer, inclusion and provision rules. Third, conditional on θ, the provision probability $\rho^j(\theta)$ is independent of all other provision probabilities as well as all inclusion probabilities.

3 Characterizing the Optimal Mechanism

We now describe the mechanism design problem. As we mentioned in the previous section, it is without loss of generality to only consider simple, anonymous and symmetric mechanisms. Of course, an incentive feasible mechanism has to be incentive compatible, individually rational and budget balanced. A mechanism $g = (\rho, \eta, t) \equiv (\{\rho^j\}_{j\in\mathcal{J}}, \{\eta^j\}_{j\in\mathcal{J}}, t)$ is *incentive compatible* if truth-telling is a Bayesian Nash equilibrium in the revelation game induced by g, i.e., for all $i \in \mathcal{I}$, $(\theta_i, \theta_i') \in \Theta^2$,

$$\sum_{\theta_{-i}\in\boldsymbol{\Theta}_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i^j - t_i(\theta_i) - \left[\sum_{\theta_{-i}\in\boldsymbol{\Theta}_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i}, \theta_i')\,\eta_i^j(\theta_i')\,\theta_i^j - t_i(\theta_i')\right] \geq 0. \quad (3)$$

A mechanism g satisfies *individual rationality* at the interim stage if the expected surplus from truth-telling for the "lowest" type, i.e., type-$\mathbf{l} \equiv (l, ..., l)$, is non-negative:[8]

$$\sum_{\theta_{-i}\in\boldsymbol{\Theta}_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i}, \mathbf{l})\,\eta_i^j(\mathbf{l})\,l - t_i(\mathbf{l}) \geq 0,\ \forall i \in \mathcal{I}. \quad (4)$$

A mechanism g is *ex ante* budget-balanced if:[9]

$$\sum_{i=1}^{n} \sum_{\theta_i\in\Theta} \beta(\theta_i)\,t_i(\theta_i) - \sum_{\theta\in\Theta^n} \boldsymbol{\beta}(\theta) \sum_{j=1}^{M} \rho^j(\theta)\,cn \geq 0. \quad (5)$$

Because of the assumed transferrable utility in (1), the constrained *ex ante* Pareto efficient allocations are characterized by a fictitious social planner's problem to maximize social surplus:

$$\max_{\{\rho,\eta,t\}} \sum_{\theta\in\Theta^n} \boldsymbol{\beta}(\theta) \sum_{j=1}^{M} \rho^j(\theta) \left[\sum_{i=1}^{n} \eta_i^j(\theta_i)\,\theta_i^j - cn\right] \quad (6)$$

s.t. (3), (4), (5),

and $\rho^j(\theta) \in [0,1],\ \forall \theta \in \Theta^n$ and $j \in \mathcal{J}$ (7)

$\eta_i^j(\theta_i) \in [0,1],\ \forall \theta_i \in \Theta$ and $i \in \mathcal{I},\ j \in \mathcal{J}$. (8)

[8]It is routine to show that individual rationality for type-$\mathbf{l}$ and the incentive compatibility constraints (3) ensure that the individual rationality constraints for all other types are satisfied.

[9]As shown in Borgers and Norman (2008) it is without loss of generality to consider a balanced budget constraint in *ex ante* form.

where constraints (7) and (8) ensure that inclusion and provision rules always generate valid probabilities.

To simplify our discussions below, we will refer to an incentive constraint in (3) as a *downward* incentive constraint if $\theta_i^j \geq \theta_i'^j$ for all $j \in \mathcal{J}$, as an *upward* incentive constraint if $\theta_i^j \leq \theta_i'^j$ for all $j \in \mathcal{J}$, as a *diagonal* incentive constraint if there exists $j, k \in \mathcal{J}$ so that $\theta_i^j > \theta_i'^j$ and $\theta_i^k < \theta_i'^k$, and as an *adjacent* incentive constraint if θ_i and θ_i' differ only in a single coordinate.

3.1 The Relaxed Optimization Problem

Based on intuition from unidimensional mechanism design problems (e.g. Matthews and Moore 1987) and similar to Armstrong and Rochet (1999), we now formulate a *relaxed problem* where all incentive constraints in (3) except the *downward adjacent* incentive constraints are removed.[10] Denote by $\theta_i|l_k$ the type that is obtained from θ_i if the k-th coordinate is changed from h to l, the relaxed problem can be written as:

$$\max_{\{\rho,\eta,t\}} \sum_{\theta\in\Theta^n} \boldsymbol{\beta}(\theta) \sum_{j=1}^{M} \rho^j(\theta) \left[\sum_{i=1}^{n} \eta_i^j(\theta_i)\,\theta_i^j - cn\right] \tag{9}$$

$$\text{s.t.} \quad \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta)\,\eta_i^j(\theta_i)\,\theta_i^j - t_i(\theta_i)$$

$$- \left[\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i},\theta_i|l_k)\,\eta_i^j(\theta_i|l_k)\,\theta_i^j - t_i(\theta_i|l_k)\right] \geq 0 \tag{10}$$

$$\forall \theta_i \in \Theta, \text{ and } \forall k \text{ such that } \theta_i^k = h, \text{ and } \forall i \in \mathcal{I}$$

and (4), (5), (7) and (8).

The existence of solutions to problem (9) can be established by first compactifying the constraint set and then applying Weierstrass maximum theorem.

In what follows we will first present a sequence of intermediate results that will be used to characterize the solution to the relaxed problem (9); then in Section 3.5 we will present conditions on the primitives to ensure that the solution to the relaxed problem also solves the full problem (6).

3.2 Relating the Multipliers

Since the environment we consider is symmetric across goods in both cost function and valuation distributions, it can be shown that the optimal provision and inclusion rules can be restricted to be symmetric across goods without loss of generality (see Proposition 2 of Fang and Norman 2006a). In addition, strong duality in linear programming ensures that the value of the multiplier associated with any of the downward adjacent incentive constraint (10) depends only on *the number of goods* for which the consumer has a high valuation. To state the result formally, we introduce some notation. For any $\theta_i \in \Theta$, write $m(\theta_i) \in \{0, ..., M\}$ as the number of goods for which $\theta_i^j = h$, i.e.,

$$m(\theta_i) = \#\left\{j \in \mathcal{J} : \theta_i^j = h\right\}. \tag{11}$$

[10] Also see Chapter 6 of Bolton and Dewatripont (2005) for a detailed analysis of multidimensional mechanism problems for a revenue-maximizing monopolist selling private goods.

Thus, for every $i \in \mathcal{I}$ and every $\theta_i \in \Theta$ there are $m(\theta_i)$ downward adjacent incentives in (10). Given any $m \in \{0, ..., M\}$ there are $\frac{M!}{m!(M-m)!}$ types $\theta_i \in \Theta$ such that $\theta_i^j = h$ for exactly m goods. Since θ_i is an exchangeable random vector all these types are equally likely, the probability that an agent has high valuations for exactly m goods, denoted by β_m, is given by

$$\beta_m = \frac{M!}{m!\,(M-m)!}\beta(\theta_i), \tag{12}$$

where θ_i is any type with m high valuations.[11] The following lemma follows from strong duality in linear programming:

Lemma 1 *For every $m \in \{1, ..., M\}$, it is without loss of generality to assume that there exists some $\lambda(m) \geq 0$ such that $\lambda(m)$ is the multiplier associated with every constraint (10) such that $m(\theta_i) = m$.*

From now on, we denote by $\lambda(m)$ the multiplier for all downward adjacent incentive constraints for types with $m \geq 1$ high valuations. Also, we let $\lambda(0)$ denote the multiplier associated with the participation constraint (4) for type $\mathbf{l} = (l, ..., l)$, and Λ denote the multiplier to the resource constraint (5). Our next result show that the multipliers $\lambda(m), m \in \{0, ..., M\}$ and Λ are systematically linked:

Lemma 2 *For every $m \in \{0, ..., M\}$, the value of $\lambda(m)$ is related to Λ in accordance with*

$$\lambda(m) = \begin{cases} \Lambda \text{ if } m = 0, \\ \frac{(m-1)!(M-m)!}{M!}\Lambda\sum_{j=m}^{M}\beta_j = \frac{(m-1)!(M-m)!}{M!}\Lambda \Pr[m(\theta_i) \geq m] \text{ if } m = 1, ..., M \end{cases} \tag{13}$$

where β_j is defined in (12).

Lemma 2 is a key step in solving (9). Its role is similar to the characterization of incentive feasibility in terms of a single integral constraint in unidimensional mechanism design problems (i.e., the approach in Myerson 1981 and others). In multidimensional problems, it is impossible to collapse all the constraints into a single constraint. Instead, Lemma 2 allows us to *indirectly* relate all optimality conditions to a single constraint.

The easiest way to understand Lemma 2 is to consider the first order condition of problem (9) with respect to $t_i(\theta_i)$ where θ_i is a type with $m(\theta_i) = m \in \{1, ..., M-1\}$ high-valuation goods. The constraints in problem (9) where $t_i(\theta_i)$ appears can be delineated as follows. First, $t_i(\theta_i)$ appears in the incentive constraints (10). There are m ways to change a single h-coordinate into an l, so there are m different adjacent downward deviations from θ_i. In addition, there are $M-m$ types θ_i' with $m(\theta_i') = m+1$ such that an adjacent downward deviation from θ_i' can "turn into" type θ_i. Second, $t_i(\theta_i)$ appears in the balanced budget constraint (5). With these observations, the optimality conditions to the program (9) with respect to $t_i(\theta_i)$ yields:

$$-\lambda(m)\,m + \lambda(m+1)(M-m) + \Lambda\beta_i(\theta_i) = 0. \tag{14}$$

When $m(\theta_i) = 0$, which is the case if and only if $\theta_i = \mathbf{l} = (l, ..., l)$, there is no possible downward adjacent deviation from type-$\mathbf{l}$; instead, $t_i(\mathbf{l})$ appears in the participation constraint (4). Thus the optimality condition with respect to $t_i(\mathbf{l})$ is:

$$-\lambda(0) + \lambda(1)\,M + \Lambda\beta_i(\mathbf{l}) = 0. \tag{15}$$

[11] For example, in the simplest case where valuations for different goods are i.i.d. with α being the probability that $\theta_i^j = h$, we have that for any θ_i such that $m(\theta_i) = m$, $\beta(\theta_i) = \alpha^m(1-\alpha)^{M-m}$ and $\beta_m = \frac{M!}{m!(M-m)!}\alpha^m(1-\alpha)^{M-m}$.

Similarly, when $m(\theta_i) = M$, which is the case if and only $\theta_i = \mathbf{h} = (h, ..., h)$, there is no possible downward adjacent deviation to type-$\mathbf{h}$, thus the optimality condition with respect to $t_i(\mathbf{h})$ is:

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

Because $\beta_M = \beta_i(\mathbf{h})$, it is immediate from (16) that the characterization in (13) holds for $m = M$. The rest of the proof is an induction argument using (14) and (15) recursively. The identity (12) is finally used to translate terms involving $\beta(\theta_i)$ into β_m.

A rough intuition for Lemma 2 is as follows. If an extra unit of revenue can be extracted from types with m high valuations without upsetting any constraint, then an extra unit can be extracted from all higher types as well. This roughly follows from the fact that transfers enter the utility function (1) linearly, hence the payoff difference for an agent with m high valuations between truth-telling and announcing $m-1$ high valuations depends only on the difference in the transfers. Multipliers are therefore proportional to the probability that the number of valuations exceeds m.

3.3 Optimal Inclusion Rules

Our next two lemmas establish some properties of the optimal inclusion rules when a public good is provided. Lemma 3 shows that individuals will always be able to access the goods for which they have high valuations. In contrast, Lemma 4 shows that access to goods for which the consumer has a low valuation is determined by the implications of allowing access on a non-trivial trade-off between social welfare and revenue extraction, where the weights on revenue extraction is a scaling of Λ, the multiplier Λ associated with the resource constraint (5).

Lemma 3 ***(Optimal Inclusion Rule for High-Valuation Goods)*** *Suppose that $\theta_i \in \Theta$ such that $\theta_i^j = h$ and that $\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta_i, \theta_{-i}) > 0$. Then, $\eta_i^j(\theta_i) = 1$ in any optimal solution to (9).*[12]

Lemma 3 may appear to be the well-known "no distortion at the top" result from the unidimensional mechanism design problem. However, this is not an accurate interpretation as we are in a multi-dimensional setting where *a priori* it might be optimal to restrict access to high-valuation goods for types with too few high valuations. Instead, Lemma 3 states that an agent should be given access to her high-valuation goods *irrespective of* her total number of high-valuation goods. The result is indeed best understood in terms of the relationship between the multipliers in (14). Providing an extra high-valuation good to a type with m high valuations relaxes the downward adjacent incentive constraints (10) for every type with exactly m high valuations. There are m such incentive constraints, so the effect associated with higher utility from truth-telling for these types is $\lambda(m)mh$. However, giving access to a high-valuation good j makes it more tempting for a type with $m+1$ high valuations to announce only m high valuations. There are $M-m$ types that could change a single coordinate from high to low and pretend to have m high valuation goods. The negative effect on the utility of these types from mis-reporting as a type with m high valuations is $-\lambda(m+1)(M-m)h$. It then follows from (14) that the first positive effect from making truth-telling more appealing always dominates the second negative effect; hence all consumers always get access to their high-valuation goods.

[12]The condition $\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta_i, \theta_{-i}) > 0$ is needed in Lemmas 3 and 4 because the inclusion rules have no effect on either the objective function or the constraints when the conditional probability of provision is zero.

The characterization of the optimal inclusion rule for low-valuation goods is more complicated. Some exclusions from access to low-valuation goods are essential in order to extract sufficient revenue to cover the costs of provision; but at the same time, exclusions diminish social welfare. Optimal inclusions to low-valuation goods are therefore determined from a non-trivial trade-off between revenue extraction and social welfare maximization. It turns out that this trade-off is neatly captured by the functions $G_m(\Phi), m = 0, ..., M-1$, defined by

$$G_m(\Phi) = (1-\Phi)(M-m)l\beta_m + \Phi\left[\beta_m(M-m)l - (h-l)\sum_{j=m+1}^{M}\beta_j\right], \tag{17}$$

where $G_m(\Phi)$ is relevant for each type θ_i with $m(\theta_i) = m \in \{0, ..., M-1\}$. We will first state the formal result, and then discuss why $G_m(\Phi)$ summarizes the compromise between welfare and profit maximization.

Lemma 4 ***(Optimal Inclusion Rule for Low-Valuation Goods)*** *Suppose that $\theta_i \in \Theta$ with $m(\theta_i) = m \in \{0, ..., M-1\}$ where $\theta_i^j = l$ for some $j \in \mathcal{J}$, and that $\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta_i, \theta_{-i}) > 0$. Then,*

$$\eta_i^j(\theta_i) = \eta(m) \equiv \begin{cases} 0 & \text{if } G_m(\Phi) < 0 \\ z \in [0,1] & \text{if } G_m(\Phi) = 0 \\ 1 & \text{if } G_m(\Phi) > 0, \end{cases} \tag{18}$$

in any optimal solution to (9) where $\Phi = \frac{\Lambda}{1+\Lambda}$.

To interpret (17) and gain intuition for Lemma 4, imagine that all goods are provided with certainty and consider the following two potential inclusion rules: (a) All agents with at least m high valuations receive access to all the goods, while those with $m-1$ or fewer high valuations only receive access to their high valuation goods; (b) All agents with at least $m+1$ high valuations receive access to all the goods, while those with m or fewer high valuations only receive access to their high valuation goods. It is easy to see that $(M-m)l\beta_m$, the first term in (17), is the gain in *social welfare* if the planner changes the inclusion rule from (b) to (a). Less obvious is that the term in square bracket in (17) is the difference in *revenue* between the two inclusion rules for the provider of the public goods. To see this, note that under inclusion rule (a), the provider can charge $mh + (M-m)l$ to agents who consume all goods, and charge h per good to the remaining agents who only obtain access to their high valuation goods. The expected revenue under inclusion rule (a) and the above pricing policy, denoted by $R(m)$, is then:[13]

$$R(m) = [mh + (M-m)l]\sum_{j=m}^{M}\beta_j + h\sum_{j=1}^{m-1}\beta_j j. \tag{19}$$

The expected revenue from inclusion rule (b) and analogous pricing policy is $R(m+1)$. It follows from some algebra that:

$$R(m) - R(m+1) = \beta_m(M-m)l - (h-l)\sum_{j=m+1}^{M}\beta_j. \tag{20}$$

Thus, the term in square bracket in (17) represents the effect on the provider's *revenue,* which may be positive or negative, if the planner changes the inclusion rule from (b) to (a). Consequently, $G_m(\Phi)$ is

[13] A profit maximizing monopolist would simply pick m to maximize $R(m)$ provided that the pubic good is provided.

simply a *weighted average of the effect on social welfare and revenue* from giving, instead of excluding, the agents with m high valuations access to all their low-valuation goods where Φ, the weight on revenue, is a normalization of the multiplier on the resource constraint (5).[14]

The optimal inclusion rule for low valuation goods as characterized in Lemma 4 thus states that if the trade-off is in favor of social welfare, i.e. when $G_m(\Phi) > 0$, then the agents with m high valuations should be provided with full access to all their low-valuation goods; if $G_m(\Phi) < 0$, then they should be excluded from their low-valuation goods.

3.4 Optimal Provision Rules

We will now characterize the optimal provision rule. To simplify exposition, it is useful to introduce the following notation. For a given type profile $\theta = (\theta_1, ..., \theta_n)$, we denote by $H^j(\theta, m)$ the number of agents who have a *high* valuation for good j and a total of m high valuations; similarly, we denote by $L^j(\theta, m)$ the number of agents with a *low* valuation for good j and a total of m high valuations.[15] The optimal provision rule can then be expressed in terms of $H^j(\theta, m)$, $L^j(\theta, m)$, and $G_m(\Phi)$ as follows:

Lemma 5 ***(Optimal Provision Rules)*** *The provision rule for good j in the optimal solution to (9) satisfies:*

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

Similar to the optimal inclusion rule (18), the optimal provision rule (21) can also be interpreted as a compromise between welfare and revenue maximization. To see this, recall that Lemma 3 established that all agents always get access to all high-valuation goods that are provided. Thus, the term $\sum_{m=1}^{M} H^j(\theta, m)\, h$ is the social surplus from all consumers with a high valuation for good j. To understand the second term in (21), recall that $G_m(\Phi)$ is the effect from giving an agent with m high valuations access to *all* their low-valuation goods. There are $L^j(\theta, m)$ agents with a low valuation for j and a total of m high valuations, but $L^j(\theta, m)\, G_m(\Phi)$ must be scaled by $\frac{1}{\beta_m (M-m)}$ in order to express the effect in the same units as $\sum_{m=1}^{M} H^j(\theta, m)\, h$. The factor $\frac{1}{M-m}$ is straightforward as $G_m(\Phi)$ measures the effect of providing access to *all* $M - m$ low-valuation goods to agents with m high valuations, so scaling by $\frac{1}{M-m}$ yields the relevant *per good* effect. The need to scale up by $\frac{1}{\beta_m}$ is a consequence of $G_m(\Phi)$ being expressed as an unconditional value, while (21) is expressed as a value conditional on a *fixed* type profile θ. Hence, the first two terms in (21) may be thought of as the effect from providing good j on a combination of social welfare and revenue, whereas the third term obviously is the associated cost of provision.

3.5 Linking the Relaxed Problem (9) and the Full Problem (6)

We are now in a position to link the solution to the relaxed problem (9) with that of the full problem (6). Many steps of this analysis are similar to Matthews and Moore (1987), but the multidimensional nature of our environment leads to some important differences (see also Armstrong and Rochet 1996 for a similar analysis).

[14] This property has been shown in unidimensional settings (see Hellwig 2003 and Norman 2004); but to the best of our knowledge, no analogous result in a multidimensional setting has been shown in the literature.

[15] Note that $H^j(\theta, m)$ and $L^j(\theta, m)$ are simple accounting summaries.

Analogous to the standard approach in solving unidimensional problems, a key condition for the solution to the relaxed problem (9) to also solve the full problem (6) is that the mechanism that solves the relaxed problem is *monotonic* in the following sense:

Definition 1 *A mechanism (ρ, η, t) is monotonic if $\eta_i^j(\theta_i) \le \eta_i^j(\theta_i')$ and $\rho^j(\theta_{-i}, \theta_i) \le \rho^j(\theta_{-i}, \theta_i')$ whenever $m(\theta_i) \le m(\theta_i')$ and $\theta_i^j \le \theta_i'^j$.*

One can show that if the solution to the relaxed problem (9), (ρ, η, t), is monotonic and all the downward adjacent constraints *bind*, then (ρ, η, t) also satisfies all the other incentive constraints in the full problem (6). Moreover, one can show that the downward adjacent incentive constraints in (9) will indeed bind if (ρ, η, t) is monotonic and is *not ex post* efficient.[16] Thus, we have:

Proposition 1 *Let (ρ, η, t) be an optimal solution to (9). If (ρ, η, t) is monotonic and is **not** ex post efficient, then (ρ, η, t) is also an optimal solution to the full problem (6).*

Now we provide a useful sufficient condition under which the solution to the relaxed problem (9) is indeed monotonic in the sense of Definition 1:

Proposition 2 *Let (ρ, η, t) be a solution to the relaxed problem (9), then (ρ, η, t) is monotonic if $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ is strictly decreasing in m on $\{0, ..., M-1\}$.*

The sufficient condition identified in Proposition 2 is almost, but not quite, a hazard rate condition, as the term $\frac{1}{M-m}$ makes the condition different from a simple hazard rate condition. To understand the "almost hazard rate condition," recall that the optimal inclusion rules for the relaxed problem (9), as characterized in Lemma 4, have the property that, for any agent with m high valuations, she either gets access to all goods or only her high-valuation goods. Allowing access to all goods is preferred in terms of social surplus, but may reduce the revenue. Specifically, with probability β_m, an agent has exactly m high valuations, and such an agent is willing to pay an extra $(M-m)l$ for access to her low-valuation goods. On the other hand, with probability $\sum_{j=m+1}^{M}\beta_j$, an agent's number of high-valuation goods exceeds m, and from such an agent the revenue is reduced by $(h-l)$. The "almost hazard rate condition" identified in Proposition 2 insures that $\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m(M-m)}\frac{(h-l)}{l}$ decreases in m. This implies that if it is optimal to give agents with m high valuations access to their low-valuation goods, then it must also be optimal to give agents with more than m high valuations access to their low-valuation goods. That is, the condition makes sure that the optimal mechanism is monotonic.

In general, $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ may be non-monotonic, because $\beta_m(M-m)$ may locally decrease faster than $\sum_{j=m+1}^{M}\beta_j$. However, the important special case where valuations are independent satisfies the "almost hazard rate condition," because under independence, the tail of the distribution over m is thin so that $\sum_{j=m+1}^{M}\beta_j$ decreases rapidly as m increases:

Remark 1 *A sufficient condition for $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ to be strictly decreasing in m on $\{0, ..., M-1\}$ is that the valuations for any goods j and $j' \neq j$ are independent.*

An immediately corollary of Proposition 2 is that the inclusion rule for low valuation goods as characterized in Lemma 4 takes a sharp threshold rule if $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ is strictly decreasing in m:

[16] The detailed proofs of these results are available in the online appendix.

Corollary 1 *Let* (ρ, η, t) *be a solution to (9). Suppose that* $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ *is strictly decreasing in* m *on* $\{0, ..., M-1\}$*. Then, there exists some* $\widetilde{m}$ *such that:*

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

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

3.6 The Main Result

Now we provide our main asymptotic result regarding the limit of the sequences of exact optimal solutions to (9) as the number of agents n goes to infinity.

Because our asymptotic characterization is a limit of exact solutions to the finite problem, it is a *valid approximation* of the solution for a large, yet finite, economy. This is one of the key advantages of our approach over the approach of directly considering a mechanism design problem with a continuum of agents. Another problematic issue related to mechanism design with a continuum of agents is that one is forced to assume that aggregate variables are independent of individual announcements; as a result, incentive constraints are much harder, if possible at all, to formulate under the continuum approach.[17]

Considering the asymptotic result as the number of the agents goes to infinity allows us to obtain a more easily interpretable characterization of the solutions when n is large. Note that our earlier characterizations of the optimal inclusion and provision rules for the finite case are still contingent on the multiplier on the resource constraint; the limiting characterization is easier to understand because it can be described without any reference to any endogenous multiplier. Conditions for the limiting case simplify for two reasons. First, terms such as $H^j(\theta, m)/n$ and $L^j(\theta, m)/n$ used in Lemma 5 to describe the optimal provision rule converge to their expectations by the Law of Large Number. The consequence is that provision probabilities for unusual realizations of θ can be largely ignored. Secondly, in a large economy, a version of the "Paradox of Voting" applies to the provision rules: an individual agent must have a negligible impact on provision decisions when n is large. This simplifies the analysis tremendously as the provision rule becomes almost constant in the relevant range when n is large enough.

Henceforth mechanisms are indexed by the number of agents n when needed. Using the central limit theorem one can establish that:

Lemma 6 *For each* n*, let* (ρ_n, η_n, t_n) *be a solution to problem (9). Then,* $\mathrm{E}[\rho_n^j(\theta)\,|\theta_i'] - \mathrm{E}[\rho_n^j(\theta)\,|\theta_i''] \to 0$ *as* $n \to \infty$ *for any* j *and any pair* $\theta_i', \theta_i'' \in \Theta$*.*

Lemma 6 immediately implies that $\mathrm{E}[\rho_n^j(\theta)\,|\theta_i'] - \mathrm{E}[\rho_n^j(\theta)] \to 0$. Hence, all conditional probabilities appearing in the incentive constraints may be approximated by the *ex ante* probability of provision, a fact used extensively in the proof of Proposition 3.

Our main result below is about the provision probabilities and inclusion rules in the limit as n goes to infinity:

Proposition 3 *Suppose that* $\frac{1}{M-m}\frac{\sum_{j=m+1}^{M}\beta_j}{\beta_m}$ *is strictly decreasing on* $\{0, ..., M-1\}$ *and let* $R(m)$ *be the revenue from the inclusion and pricing rules as defined in (19). Let* (ρ_n, η_n, t_n) *be a solution to the full problem (6) when there are* n *agents in the economy. Then,*

[17] For example, a Groves-Clarke mechanism cannot even be formulated, despite the fact that such a mechanism is applicable and generates an efficient outcome for *any* finite economy.

1. ***(Asymptotic Provision Probabilities)***

   (a) *If* $\max_m R(m) - cM < 0$, *then* $\lim_{n\to\infty} E\rho_n^j(\theta) = 0$ *for all* $j \in \mathcal{J}$;[18]

   (b) *If* $\max_m R(m) - cM > 0$, *then* $\lim_{n\to\infty} E\rho_n^j(\theta) = 1$ *for all* $j \in \mathcal{J}$.

2. ***(Asymptotic Inclusion Rules Conditional on Provision)*** *If* $\max_m R(m) - cM > 0$, *and* m^* *be the smallest* m *such that* $R(m) - cM > 0$, *then there exists* $N < \infty$ *such that if* $n \geq N$,

   (a) *for all type* θ_i *with* $m(\theta_i) \geq m^*$, $\eta_n^j(\theta_i) = 1$ *for all* $j \in \mathcal{J}$;

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

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

   for all j *such that* $\theta_i^j = l$;

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

The proof of Proposition 3 requires quite a bit of technical work. Yet, the key idea is rather simple. Provided that the "almost hazard rate condition" is satisfied, Corollary 1 ensures that the optimal inclusion rules have a threshold characterization. The condition on whether $\max_m R(m) - cM$ is positive or negative is therefore a condition on whether it is at all possible to generate sufficient revenues to cover the costs if all the public goods are provided with certainty. Most work in the proof for Part 1 goes into establishing that a large economy is approximately an economy where the provision decisions are made *ex ante*, not conditioning on θ. The intuition is that when agents have a small influence on provisions, the welfare loss from making the provision decision *ex ante* is negligible.

Part 2 in Proposition 3 provides a simple limiting characterization of the optimal inclusion rule. As we already know from the finite case, all agents get access to their high-valuation goods. Whether an agent gets access to her low-valuation goods, however, depends on whether she has at least m^* high-valuation goods in her reported profile, where m^* is determined as the smallest threshold for which the provider can make a positive profit if providing the good for sure. The reason for the mixing for types with $m^* - 1$ valuations is that if access is granted only to those with m^* high valuations and above, there will be a strict budget surplus when n is large. There is therefore room to give partial access to low-valuation goods to types with $m^* - 1$ high valuations, which increases social surplus.

4 Some Special Cases

In this section, we describe some special cases. We first consider the case with a single good. We then study the case with two goods, where we are able to provide a complete characterization also when the "almost hazard rate" condition in Proposition 3 is violated. We finally summarize our results in relation to the existing literature in Table 1.

[18] More precisely, for this case $\lim_{n\to\infty} E\rho_n^j(\theta) = 0$ for all $j \in \mathcal{J}$ holds for any sequence of *feasible* solutions to the full problem (6).

4.1 A Single Public Good

Let α be the probability that an agent has a high valuation for the public good.[19] With only a single good, the "almost hazard condition" in Proposition 3 is trivially satisfied, thus Proposition 3 is always applicable. Note from the definition of $R(m)$ in (19), we have:

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

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

The case with $l \geq c$ is trivial: if the low valuation exceeds the cost of provision, it is first best efficient to always provide and never exclude any consumers from access, which can be implemented by a uniform user fee of c.

When $l < c < \alpha h$, we know from Proposition 3 that, as $n \to \infty$, the probability for access for type l goods converges to

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

Moreover, $G_1(\Phi)$ as defined in (17) is equal to zero, thus the provision rule characterized in (21) simplifies to

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

which, interestingly, is exactly the same provision rule that would be chosen by a profit-maximizing monopolist. Of course, a a profit-maximizing monopolist will not sell access to low valuation agents. Thus for the one good case, the welfare loss from a for-profit monopolist relative to the constrained social optimum is only due to over-exclusion, not due to under-provision, even for finite n. With more than one good, there is typically also under-provision by a for-profit monopolistic provider when n is finite.

Separate Provisions of Many Excludable Public Goods. If there are $M > 1$ excludable public goods, but if the social planner is restricted to consider the provision of each of the public goods separately, then the optimal separate provision and exclusion rules for each public good is identical to the single good case described above, with $\alpha = \sum_{m=1}^{M} \frac{m}{M} \beta_m$ being the (marginal) probability that an agent has a high valuation for any particular good.[20]

It is also clear that the results for the one-good case applies to the case with multiple goods when valuations for different goods are perfectly correlated.

4.2 Two Excludable Public Goods

Here we describe the results for the two-good case.[21] For the two-good case, we can also characterize the optimal mechanism when the"almost hazard rate" regularity condition is violated. Most of the

[19]Note that when $M = 1$, $\alpha = \beta_1$ in our earlier notation. We introduce the new notation α so that our discussion for the one good case can be exactly mapped to the case of seperation provision of multiple public goods we will discuss below.

[20]Of course, this statement is true only under our maintained assumption that there are no complementarities in preferences and production costs.

[21]In an earlier working paper, Fang and Norman (2006a), we exclusively focused on the two-good case.

existing literature on multiple product mechanism design focused on the two-goods case, e.g. Armstrong (1996) and Armstrong and Rochet (1999). It turns out that when there are only two goods, the "almost hazard rate" regularity condition is reduced to a "not too positively correlated" condition (e.g. Armstrong and Rochet 1999).

4.2.1 When the "Almost Hazard Rate" Condition is Satisfied

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

$$\frac{\beta_1}{2} > \frac{\beta_0\beta_2}{1-\beta_0}, \tag{24}$$

which can be interpreted as saying that the valuation of the two public goods are "not too positively correlated." Furthermore, it can be verified that

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

where $\alpha = \frac{\beta_1}{2} + \beta_2$ is the marginal probability that an agent has a high valuation for any good. Ruling out the trivial case of $l \geq c$, we see from Proposition 3 that there are three possibilities:

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

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

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

$$\underbrace{\frac{\beta_1/2}{\beta_0 + \beta_1/2}}_{\substack{\text{prob of mixed type} \\ \text{given a low valuation}}} \times \underbrace{\frac{2\alpha h - 2c}{2\alpha h - (\beta_1 + \beta_2)(h + l)}}_{\substack{\text{access prob to the low-valuation} \\ \text{good for mixed type from (25)}}}. \tag{26}$$

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

4.2.2 When the "Almost Hazard Rate" Condition is Violated

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

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

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

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

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

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

Surprisingly, Proposition 4 shows that the solution is identical to the case when bundling is not allowed.[22] To understand why, recall that asymptotic provision or non-provision is related to whether the maximal revenue for a monopolistic provider of the goods – if provided – exceeds the costs. The revenue maximizing selling strategy for a monopolist, if both public goods are provided, is either to sell goods separately at price h, or sell the goods as a bundle at price $h + l$, or to charge l for each good. These selling strategies generate a revenue of $2\alpha h$, $(\beta_1 + \beta_2)(h + l)$, and $2l$ respectively. Since we are already assuming $l < c$, the question is thus whether $\max\{(\beta_1 + \beta_2)(h + l), 2\alpha h\}$ exceeds $2c$. For the first part of Proposition 4, if $\alpha h < c$ and $l < c$ are both satisfied, we have that

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

[22] This is *not* inconsistent with the results in McAfee, McMillan and Whinston (1989) and Jehiel, Meyer-ter-Vehn and Moldovanu (2007). They showed that for "*generic*" *continuous* valuation distributions, a monopolist seller's revenue will be higher under mixed bundling in posted price and auction settings respectively. Our distributions are discrete.

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

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

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

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

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

Some algebra along the lines discussed in connection with (26) shows that the *ex ante* probability of getting access and therefore also the social surplus is actually smaller using the bundling mechanism in this case. Ultimately, this is driven by the relative scarcity of mixed type agents.

4.2.3 Relationship to the Existing Literature

Table 1 summarizes our results on the asymptotic provision probabilities under different bundling and exclusion scenarios for the two-good case, and contrasts our results with those in the literature.[23]

Bundling＼Exclusion	No Exclusion	Exclusion	
No Bundling	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite 1990)	$\mathrm{E}\rho_n^{j*} \to 0$, if $\alpha h < c$ $\mathrm{E}\rho_n^{j*} \to 1$, if $\alpha h > c$ (Norman 2004)	
Bundling Allowed	$\mathrm{E}\rho_n^{j*} \to 0$ (Mailath and Postlewaite 1990)	Case 1: $\frac{\beta_1}{2} > \frac{\beta_0\beta_2}{1-\beta_0}$ $\mathrm{E}\rho_n^{j*} \to 0$, if $\max_m R(m) < 2c$ $\mathrm{E}\rho_n^{j*} \to 1$, if $\max_m R(m) > 2c$	Case 2: $\frac{\beta_1}{2} \leq \frac{\beta_0\beta_2}{1-\beta_0}$ $\mathrm{E}\rho_n^{j*} \to 0$, if $\alpha h < c$ $\mathrm{E}\rho_n^{j*} \to 1$, if $\alpha h > c$
		(This Paper)	

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

5 Conclusion and Discussion

This paper studies the welfare maximizing provision mechanism for multiple excludable public goods when agents' valuations are private information. For a parametric class of problems with M goods whose valuations take binary values, we fully characterize the optimal mechanism and demonstrate that it involves bundling if a regularity condition, akin to a hazard rate condition, on the distribution of valuations is satisfied. Bundling improves the allocation in two ways. First, it may increase the asymptotic

[23] Mailath and Postlewaite (1990) considered a single-dimensional problem without use exclusion. However, the probabilities of provision in a multidimensional setting allowing for bundling can be bounded from above by a single-dimensional problem, where the valuation is the maximum of the individual good valuations.

provision probability of socially efficient public goods from zero to one. Second, the extent of use exclusion is decreased under bundling. For the case of two goods, we also show that if the regularity condition is violated, then the optimal solution replicates the separate provision outcome.

By studying the optimal, instead of revenue-maximizing, mechanism for the provision of excludable public goods, our paper highlights the potential importance of bundling on *social welfare* for a variety of markets where the goods are non-rival but potentially excludable in use. Thus our analysis can be viewed as a *positive* theory of bundling in these markets. It can also be viewed as a step toward the development of a useful *normative* benchmark for bundling for markets of excludable public goods, or, more generally, goods with large fixed cots. Our model is highly stylized, but it still has enough flexibility to generate a non-trivial trade-off with potentially important anti-trust implications. Our analysis shows that it is sometimes possible that bundling is *required* for the monopolist provider of the public goods to break even. In such cases the profit-maximizing outcome with bundling is better for the consumer than the welfare maximizing outcome without bundling; thus a requirement to "unbundle" is strictly worse for the consumers because the public goods would simply not be provided when bundling is not allowed.[24]

Our analysis may be relevant for the recent regulations considered by U.S. Congress and the Federal Communications Commission (FCC) requiring *à la carte* pricing of cable channels. The few existing economics studies, e.g., Crawford (2008), Cullen and Crawford (2007), Crawford and Yorokoglu (2009) and Yorokoglu (2009), on the welfare effect of bundling versus *à la carte* pricing of cable channels all assumed that the content quality of the channels is not affected by the price regulations. To the extent that television programming is well approximated as excludable public goods (due to its large fixed cost, but almost negligible cost of serving additional consumers), our paper cautions that forcing *à la carte* pricing *might* lead to the deterioration of the quality, or even elimination, of some channels.

References

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

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

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

[4] Armstrong, Mark and Jean-Charles Rochet (1999). "Multi-Dimensional Screening: A User's Guide." *European Economic Review*, Vol. 43: 959-979..

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

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

[24] This line of reasoning was an important part of the motivation in the decision by the Office of Fair Trading (2003) in the U.K. on alleged anti-competitive mixed bundling by the British Sky Broadcasting Limited.

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

[8] Bolton, Patrick and Mathias Dewatripont (2005). *Contract Theory.* MIT Press: Cambridge, MA.

[9] Borgers, Tilman and Peter Norman (2008). "A Note on Budget Balance under Interim Participation Constraints: The Case of Independent Types." Forthcoming, *Economic Theory.*

[10] Crawford, Gregory S. (2008). "The Discriminatory Incentives to Bundle in the Cable Television Industry." *Quantitative Marketing and Economics*, Vol. 6, No. 1, 41-78.

[11] Crawford, Gregory S. and Ali Yorokoglu (2008). "The Welfare Effects of Bundling in Multichannel Television." mimeo, University of Arizona and New York University.

[12] Cullen, Joseph and Gregory S. Crawford (2007). "Bundling, Product Choice, and Efficiency: Should Cable Television Networks be Offered a La Carte?" *Information Economics and Policy,* Vol. 19, No. 3-4, 379-404.

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

[14] Fang, Hanming and Peter Norman (2006a). "Optimal Provision of Multiple Excludable Public Goods." Cowles Foundation Discussion Paper No. 1441R.

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

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

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

[18] Jehiel, Philippe, Moritz Meyer-ter-Vehn and Benny Moldovanu (2007). "Mixed Bundling Auctions." *Journal of Economic Theory*, Vol. 134, 494-512.

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

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

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

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

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

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

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

[26] Vincent, Daniel R. and Alejandro Manelli. (2006). "Bundling as an Optimal Selling Mechanism for a Multiple-Good Monopolist." *Journal of Economic Theory*, Vol. 127, No. 1, 1-35.

[27] Yorokoglu, Ali (2008). "Bundling and Vertical Relationships in Multichannel Television." mimeo, New York University.

APPENDIX

PROOF OF LEMMA 3: Let j be some good for which $\theta_i^j = h$. The optimality conditions for problem (9) with respect to $\eta_i^j(\theta_i)$ are:

$$\begin{aligned}\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}(\theta)\rho^j(\theta)h + \lambda(m)m\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta)h & \\ -\lambda(m+1)(M-m)\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta)h + \gamma_i^j(\theta_i) - \phi_i^j(\theta_i) &= 0;\end{aligned} \tag{A1}$$

$$\gamma_i^j(\theta_i)\eta_i^j(\theta_i) = 0 \text{ and } \phi_i^j(\theta_i)\left[1-\eta_i^j(\theta_i)\right] = 0, \tag{A2}$$

where $\gamma_i^j(\theta_i) \geq 0$ and $\phi_i^j(\theta_i) \geq 0$ are respectively the multipliers for the constraints $\eta_i^j(\theta_i) \geq 0$ and $1-\eta_i^j(\theta_i) \geq 0$. Since by assumption $\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta_i,\theta_{-i}) > 0$ and since

$$\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}(\theta)\rho^j(\theta)h = \beta(\theta_i)\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta)h, \tag{A3}$$

we can simplify (A1) to:

$$\beta(\theta_i)h + \lambda(m)mh - \lambda(m+1)(M-m)h + \frac{\gamma_i^j(\theta_i) - \phi_i^j(\theta_i)}{\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta)} = 0. \tag{A4}$$

Together with the complementary slackness conditions (A2), (A4) implies that

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

However, (14) implies that

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

Hence $\eta_i^j(\theta_i) = 1$ for all $\theta_i \in \Theta$ such that $\theta_i^j = h$. ■

PROOF OF LEMMA 4: Consider $\theta_i \in \Theta$ with $m(\theta_i) = m$ and let j be some good for which $\theta_i^j = l$. The optimality conditions for problem (9) with respect to $\eta_i^j(\theta_i)$ are:

$$\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}(\theta)\rho^j(\theta)l + \lambda(m)m \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^j(\theta)l \tag{A5}$$

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

$$\gamma_i^j(\theta_i)\eta_i^j(\theta_i) = 0 \text{ and } \phi_i^j(\theta_i)\left(1-\eta_i^j(\theta_i)\right) = 0, \tag{A6}$$

where $\gamma_i^j(\theta_i) \geq 0$ and $\phi_i^j(\theta_i) \geq 0$ are the multipliers for the constraints $\eta_i^j(\theta_i) \geq 0$ and $1-\eta_i^j(\theta_i) \geq 0$ respectively. Using (A3) as in the proof of Lemma 3, we may rearrange (A5) as:

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

Note that:

$$\begin{aligned}
& \beta(\theta_i)l + \lambda(m)ml - \lambda(m+1)\left[(M-m-1)l+h\right] \\
= \;& \beta(\theta_i)l(1+\Lambda) - \lambda(m+1)(h-l) \\
= \;& \beta(\theta_i)l(1+\Lambda) - (h-l)\frac{m!\,(M-(m+1))!}{M!}\Lambda\sum_{j=m+1}^{M}\beta_j \\
= \;& \frac{m!\,(M-(m+1))!}{M!}\left[\beta_m(M-m)l(1+\Lambda) - (h-l)\Lambda\sum_{j=m+1}^{M}\beta_j\right] \\
= \;& \frac{m!\,(M-(m+1))!}{M!}(1+\Lambda)G_m(\Phi)
\end{aligned} \tag{A8}$$

for $\Phi = \frac{\Lambda}{1+\Lambda}$, where the first equality uses (14), the second uses Lemma 2 and the third uses the formula (12) for β_m. Substituting (A8) into (A7) and using the complementary slackness conditions in (A6), we immediately have (18), as asserted in the lemma.

PROOF OF LEMMA 5: Using the notation $H^j(\theta,m)$ and $L^j(\theta,m)$ introduced in Section 3.4, we can write the optimality conditions for problem (9) with respect to $\rho^j(\theta)$ as:

$$\begin{aligned}
& \boldsymbol{\beta}(\theta)\left[\sum_{i=1}^{n}\eta_i^j(\theta_i)\theta_i^j - cn\right] + \sum_{m=0}^{M}\lambda(m)m\left[H^j(\theta,m)\boldsymbol{\beta}_{-i}(\theta_{-i})h + L^j(\theta,m)\boldsymbol{\beta}_{-i}(\theta_{-i})\eta(m)l\right] \\
& -\sum_{m=0}^{M-1}\lambda(m+1)\left\{H^j(\theta,m)\boldsymbol{\beta}_{-i}(\theta_{-i})(M-m)h + L^j(\theta,m)\boldsymbol{\beta}_{-i}(\theta_{-i})\eta(m)\left[(M-m)l+(h-l)\right]\right\} \\
& -\Lambda\beta(\theta)cn + \gamma^j(\theta) - \phi^j(\theta) \;=\; 0,
\end{aligned} \tag{A9}$$

together with the complementary slackness conditions, where $\eta(m)$ denotes the probability that an agent with m high valuation goods gets access to her low valuation goods valuations (as characterized in Lemma 4), $\gamma^j(\theta) \geq 0$ and $\phi^j(\theta) \geq 0$ are the multipliers associated with the boundary conditions for

$\rho^j(\theta) \in [0,1]$.[25] Using (A3) and the fact that

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

we can rewrite (A9), after collecting terms, as:

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

which we can further simplify, after using (14), to:

$$\sum_{m=0}^{M} H^j(\theta,m)(1+\Lambda)\,h + \sum_{m=0}^{M} L^j(\theta,m)\left\{\eta(m)\,l + \eta(m)\left[\Lambda l - \frac{\lambda(m+1)}{\beta(\theta_i)}(h-l)\right]\right\} - (1+\Lambda)\,cn + \frac{\gamma^j(\theta) - \phi^j(\theta)}{\boldsymbol{\beta}(\theta)} = 0. \tag{A10}$$

From (13) and (12), we have:

$$\frac{\lambda(m+1)}{\beta(\theta_i)} = \frac{1}{\beta_m(M-m)}\Lambda \sum_{j=m+1}^{M} \beta_j. \tag{A11}$$

Hence,

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

Substituting this into (A10) gives us

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

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

PROOF OF PROPOSITION 2: First note that (18) is irrelevant when $m = M$. Moreover, it is immediate from (21) that the probability of provision given $(\mathbf{h}, \theta_{-i})$ is always weakly higher than under (θ_i, θ_{-i}) for any $\theta_{-i} \in \Theta_{-i}$. Finally, from (18) and (21), we know that a sufficient condition for monotonicity is that $\frac{1}{\beta_m(M-m)}\max\{0, G_m(\Phi)\}$ increases in m on $\{0, ...M-1\}$, which follows if $\frac{(h-l)}{\beta_m(M-m)}\sum_{j=m+1}^{M}\beta_j$ strictly decreases in m since

$$\frac{G_m(\Phi)}{\beta_m(M-m)} = l - \Phi\left[\frac{(h-l)}{\beta_m(M-m)}\sum_{j=m+1}^{M}\beta_j\right]. \quad ■ \tag{A12}$$

[25] The notation is somewhat unsatisfactory in that $\boldsymbol{\beta}_{-i}(\theta_{-i})$ would be more appropriately denoted by $\boldsymbol{\beta}_{-i}(\theta|\theta_i)$ where θ_i would describe a type that would enter in the particular term (and therefore change with each term).

PROOF OF PROPOSITION 3: [Part 1(a)] Suppose, for a contradiction, that there is a subsequence of *feasible* mechanisms with $\lim_{n\to\infty} \mathrm{E}\left[\rho_n^j(\theta)\right] = \rho^*$ and $\lim_{n\to\infty} \eta_n(\tilde{m}) = \eta^*$, where $\tilde{m}$ is the threshold number of high valuation goods in the sense of Lemma 1, which without loss of generality is taken to be constant along the sequence by the choice of a converging subsequence. We introduce the following notations:

- $\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right]$ denotes the conditional provision probability for good j perceived by agent i given that $\theta_i^j = h$ and $\theta_i^k = h$ with $k \in \mathcal{J}$ for exactly m goods (including j).
- $\mathrm{E}\left[\rho_n^k(\theta)\mid m, \theta_i^k = l\right]$ denotes the conditional provision probability for good j perceived by agent i given that $\theta_i^k = l$ and $\theta_i^j = h$ with $j \in \mathcal{J}$ for exactly m goods.

Notice that, by Proposition 1 of Fang and Norman (2006a), neither of the conditional probabilities above depends on j; moreover, Propositions 1 and 2 in Fang and Norman (2006a) jointly imply that for every $m \in \{0,..,M\}$, there exists some $t_n(m)$ such that the transfer $t_n(m)$ is paid by every agent i with m high valuations. The utility from truth-telling for an agent with m high-valuation goods is,

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

Hence, the participation constraints for type-θ_i agents with $m(\theta_i) = m < \tilde{m}$ high valuations imply that:[26]

$$t_n(m) \le mh\mathrm{E}\left[\rho_n^j(\theta)\,|m, \theta_i^j = h\right] \text{ if } m(\theta_i) = m < \tilde{m}. \tag{A13}$$

Similarly, the participation constraint for agents with $\tilde{m}$ high valuations implies that:

$$t_n(\tilde{m}) \le \tilde{m}h\mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m}, \theta_i^j = h\right] + (M-\tilde{m})\eta_n(\tilde{m})\,\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m}, \theta_i^k = l\right] l \text{ if } m(\theta_i) = \tilde{m}. \tag{A14}$$

The downward adjacent incentive constraint for an agent with type θ_i with $m(\theta_i) = \tilde{m}+1$ against misreporting as type $\hat{\theta}_i$ (which differs from θ_i only in that one of the high valuations in θ_i is reported as low valuation) requires that

$$\begin{aligned} & (\tilde{m}+1)h\mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m}+1, \theta_i^j = h\right] + [M-(\tilde{m}+1)]\,\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m}+1, \theta_i^k = l\right] l - t_n(\tilde{m}+1) \\ \ge\;& \tilde{m}h\mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m}, \hat{\theta}_i^j = h\right] + \eta_n(\tilde{m})\,\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m}, \hat{\theta}_i^k = l\right]\{[M-(\tilde{m}+1)]\,l + h\} - t_n(\tilde{m}) \\ \ge\;& \eta_n(\tilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m}, \theta_i^k = l\right](h-l), \end{aligned}$$

where the first inequality is required by the downward adjacent incentive constraint, and the second inequality follows from (A14). Thus,

$$\begin{aligned} t_n(\tilde{m}+1) \;\le\;& (\tilde{m}+1)h\mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m}+1, \theta_i^j = h\right] + (M-\tilde{m}-1)\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m}+1, \theta_i^k = l\right] l \\ & -\eta_n(\tilde{m})\,\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m}, \theta_i^k = l\right](h-l). \end{aligned} \tag{A15}$$

[26] Strictly speaking, we only impose the participation constraint on type-**l**. However, as we mentioned in the text, the downward adjacent incentive constraints together with the participation constraint for type-**l** imply that all participation constraints hold.

Finally, for every type θ_i with $m(\theta_i) = m > \tilde{m}+1$, the incentive constraint that type θ_i will not find it profitable to mis-report as type $\hat{\theta}_i$ with $m\left(\hat{\theta}_i\right) = \tilde{m}+1$ implies that

$$\begin{aligned}
& mh\mathrm{E}\left[\rho_n^j(\theta)\,|m,\theta_i^j = h\right] + (M-m)\,\mathrm{E}\left[\rho_n^k(\theta)\,|m,\theta_i^k = l\right] l - t_n(m) \\
\geq\ & (\tilde{m}+1)h\mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m}+1,\hat{\theta}_i^j = h\right] + \mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m}+1,\hat{\theta}_i^k = l\right]\left[(M-m)\,l + (m-\tilde{m}-1)\,h\right] - t_n(\tilde{m}+1) \\
\geq\ & \mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m}+1,\theta_i^j = l\right]\left[(m-\tilde{m}-1)\,(h-l)\right] + \eta_n(\tilde{m})\,\mathrm{E}\left[\rho_n^j(\theta)\,|\tilde{m},\theta_i^k = l\right](h-l)\,,
\end{aligned}$$

where the last inequality follows from (A15). Hence, for all $m > \tilde{m}+1$,

$$\begin{aligned}
t_n(m) \leq\ & mh\mathrm{E}\left[\rho_n^j(\theta)\,|m,\theta_i^j = h\right] + (M-m)\mathrm{E}\left[\rho_n^j(\theta)\,|m,\theta_i^k = l\right] l \\
& -\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m}+1,\theta_i^k = l\right](m-\tilde{m}-1)\,(h-l) - \eta_n(\tilde{m})\,\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m},\theta_i^k = l\right](h-l)\,.
\end{aligned} \tag{A16}$$

From the "Paradox of Voting"-like Lemma 6, we have

$$\lim_{n\to\infty}\mathrm{E}\left[\rho_n^j(\theta)\,|m,\theta_i^j = h\right] = \lim_{n\to\infty}\mathrm{E}\left[\rho_n^k(\theta)\,|m,\theta_i^k = l\right] = \lim_{n\to\infty}\mathrm{E}\left[\rho_n^k(\theta)\,|\tilde{m},\theta_i^k = l\right] = \lim_{n\to\infty}\mathrm{E}\left[\rho_n^j(\theta)\right] = \rho^*. \tag{A17}$$

Combining (A17) with (A13)-(A16), it follows that, for every $\varepsilon > 0$ there is some N such that when $n \geq N$,

$$t_n(m) \leq \begin{cases} \rho^* mh + \varepsilon & \text{if } m < \tilde{m} \\ \rho^*\left[\tilde{m}h + (M-\tilde{m})\eta^* l\right] + \varepsilon & \text{if } m = \tilde{m} \\ \rho^*\left[(\tilde{m}+1)h + (M-\tilde{m}-1)l - \eta^*(h-l)\right] + \varepsilon & \text{if } m \geq \tilde{m}+1. \end{cases}$$

Now summing over m, we find that the expected per capita transfer revenue must satisfy:

$$\begin{aligned}
\sum_{m=0}^{M}\beta_m t_n(m) \leq\ & \rho^*\left[h\sum_{m=0}^{\tilde{m}}\beta_m m + \left(\sum_{m=\tilde{m}+1}^{M}\beta_m\right)\{(\tilde{m}+1)h + (M-\tilde{m}-1)l\}\right] \\
& +\rho^*\eta^*\left[\beta_{\tilde{m}}(M-\tilde{m})l - \left(\sum_{m=\tilde{m}+1}^{M}\beta_m\right)(h-l)\right] + \varepsilon \\
=\ & \rho^*(1-\eta^*)\left\{\left(\sum_{m=\tilde{m}+1}^{M}\beta_m\right)\left[(\tilde{m}+1)\,h + (M-\tilde{m}-1)\,l\right] + h\sum_{m=0}^{\tilde{m}}\beta_m m\right\} \\
& +\rho^*\eta^*\left\{\left(\sum_{m=\tilde{m}}^{M}\beta_m\right)\left[\tilde{m}h + (M-\tilde{m})\,l\right] + h\sum_{m=0}^{\tilde{m}-1}\beta_m m\right\} + \varepsilon \\
=\ & \rho^*\left[(1-\eta^*)\,R(\tilde{m}+1) + \eta^* R(\tilde{m})\right] + \varepsilon.
\end{aligned} \tag{A18}$$

Moreover, for any $\varepsilon > 0$, we can find N such that when $n > N$, $M\sum_{\theta\in\Theta^n}\boldsymbol{\beta}(\theta)\,\rho_n^j(\theta)\,c \geq \rho^* Mc - \varepsilon$. Together with (A18), this implies that

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{\theta_i\in\Theta}\beta(\theta_i)\,t_{in}(\theta_i) - \sum_{\theta\in\Theta^n}\boldsymbol{\beta}(\theta)\sum_{j=1}^{M}\rho_n^j(\theta)\,cn =\ & n\sum_{m=0}^{M}\beta_m t_n(m) - M\sum_{\theta\in\Theta^n}\boldsymbol{\beta}(\theta)\,\rho_n^j(\theta)\,c \\
\leq\ & \rho^*\left[(1-\eta^*)\,R(\tilde{m}+1) + \eta^* R(\tilde{m}) - cM\right] + 2\varepsilon.
\end{aligned} \tag{A19}$$

Under the hypothesis that $\max_m R(m) < cM$, we know that $(1-\eta^*)\,R(\tilde{m}+1) + \eta^* R(\tilde{m}) - cM < 0$. Thus, for any $\rho^* > 0$, there exists $\varepsilon > 0$ such that the right hand side of (A19) is negative. Hence, the budget constraint (5) is violated for large n along any sequence with positive provision probability. Thus $\lim_{n\to\infty}\mathrm{E}\left[\rho_n^j(\theta)\right] = 0$ for any convergent subsequence.

[Part 1(b) and Part 2] Let m^* be the smallest m such that $R(m) - cM > 0$. Consider the sequence of mechanisms $(\widehat{\rho}_n, \widehat{\eta}_n, \widehat{t}_n)$ where $\widehat{\rho}_n^j(\theta) = 1$ for every n and θ, where all agents are given access to their high valuation goods and the inclusion rule for low valuation goods is

$$\widehat{\eta}_n(m) = \begin{cases} 0 & \text{if } m \leq m^* - 2 \\ \eta^* = \frac{R(m^*)-cM}{R(m^*)-R(m^*-1)} & \text{if } m = m^* - 1 \\ 1 & \text{if } m \geq m^*, \end{cases}$$

and where the transfer tule is

$$\widehat{t}_n(m) = \begin{cases} mh & \text{if } m \leq m^* - 2 \\ (m^*-1)h + (M - m^* + 1)\eta^* l & \text{if } m = m^* - 1 \\ m^*h + (M-m^*)l - \eta^*(h-l) & \text{if } m \geq m^*. \end{cases}$$

It can be verified that under the above mechanism, truth-telling is incentive compatible and individually rational. Also, the per capita expected transfer under the above mechanism is

$$\begin{aligned} \sum_{m=0}^{M} \beta_m \widehat{t}_n(m) &= h \sum_{m=0}^{m^*-2} \beta_m m + \beta_{m^*-1}\left[(m^*-1)h + (M-m^*+1)\eta^* l\right] \\ &\quad + \left(\sum_{m=m^*}^{M} \beta_m\right)\left[m^*h + (M-m^*)l - \eta^*(h-l)\right] \\ &= (1-\eta^*)R(m^*+1) + \eta^* R(m^*) = cM, \end{aligned} \tag{A20}$$

thus the balanced budget constraint also exactly holds. Hence $(\widehat{\rho}_n, \widehat{\eta}_n, \widehat{t}_n)$ is incentive feasible for every n. The associated *per capita* social surplus with the above mechanism is

$$\frac{S_n^*(\theta)}{n} = \sum_{j=1}^{M}\left\{h\sum_{m=1}^{M}\frac{H^j(\theta,m)}{n} + l\left[\eta^*\frac{L^j(\theta,m^*)}{n} + \sum_{m=m^*}^{M}\frac{L^j(\theta,m)}{n}\right]\right\} - cM, \tag{A21}$$

where the notations $H^j(\theta,m)$ and $L^j(\theta,m)$ were explained in Section 3.4. Noting that $\lim_{n\to\infty}\frac{H^j(\theta,m)}{n} = \frac{m}{M}\beta_m$ and $\lim_{n\to\infty}\frac{L^j(\theta,m)}{n} = \frac{M-m}{M}\beta_m$, we have

$$\lim_{n\to\infty}\frac{S_n^*(\theta)}{n} = h\sum_{m=1}^{M} m\beta_m + l\left[\eta^*(M-m^*)\beta_{m^*} + \sum_{m=m^*+1}^{M}(M-m)\beta_m\right] - cM > \varepsilon = 0 \tag{A22}$$

Now we show that, as $n \to \infty$, there is no other incentive feasible mechanism that yields higher per capita social surplus than the above mechanism. Invoking Corollary 1, suppose that the surplus maximizing mechanism is characterized by an inclusion rule with threshold $\tilde{m}^*$ and inclusion probability $\tilde{\eta}^*$ given $\tilde{m}^*$, and a provision rule $\tilde{\rho}_n^j$ that must be characterized by (21).

Since by definition m^* is the smallest m such that $R(m) - cM > 0$, we know that there are three possibilities: (1). $R(\tilde{m}^*) - cM < 0$; (2). $R(\tilde{m}^*) - cM > 0$ but $\tilde{m}^* > m^*$; or (3). $\tilde{m}^* = m^*$. If $R(\tilde{m}^*) - cM < 0$, then an argument identical with that in the proof of Part 1(a) establishes that $\lim_{n\to\infty} \mathrm{E}\left[\tilde{\rho}_n^j(\theta)\right] \to 0$; thus the social welfare is lower than that obtained from $(\widehat{\rho}_n, \widehat{\eta}_n, \widehat{t}_n)$. If $\tilde{m}^* > m^*$ but $R(\tilde{m}^*) - cM > 0$, then $\lim_{n\to\infty} \mathrm{E}\left[\rho_n^j(\theta)\right] \to 1$. The social surplus must therefore converge towards the same expression as in (A22), but with m^* replaced by $\tilde{m}^*$. But, if $\tilde{m}^* > m^*$, the social surplus is smaller because there is more exclusion. Finally, suppose that $\tilde{m}^* = m^*$. If $\widehat{\eta}_n(m^*) \to \tilde{\eta}^* > \eta^*$ an argument identical with Part 1(a) establishes that $\lim_{n\to\infty} \mathrm{E}\left[\tilde{\rho}_n^j(\theta)\right] \to 0$; if $\widehat{\eta}_n(m^*) \to \tilde{\eta}^* < \eta^*$ we have that $\lim_{n\to\infty} \mathrm{E}\left[\tilde{\rho}_n^j(\theta)\right] \to 1$ and again social surplus converges towards the same expression as in (A22), but with η^* replaced by $\tilde{\eta}^*$. Again, the social surplus is smaller if $\tilde{\eta}^* < \eta^*$ since there is more exclusion.■

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

$$0 \leq \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i},\mathbf{h})\, h - t_i(\mathbf{h}) - \left[\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i}) \sum_{j=1}^{M} \rho^j(\theta_{-i},\mathbf{l})\, \eta_i^j(\mathbf{l})\, h - t_i(\mathbf{l})\right], \quad \text{(A23)}$$

to program (9). Let $\psi \geq 0$ denote the multiplier for constraint (A23). Similar to (14), with the additional constraint (A23) the optimality conditions with respect to $t(\theta_i)$ for types with 2, 1 and 0 high valuation goods are, respectively,

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

We now proceed with a sequence of intermediate results.

Claim A1 *Suppose that $\frac{\beta_1}{2} < \frac{\beta_0\beta_2}{1-\beta_0}$. Then (A23) binds.*

To see this, suppose that constraint (A23) does not bind. Then we know the solution with constraint (A23) is unchanged from (9). The inclusion rule (18) for the special case with $M = 2$ is simplified to:

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

and the provision rule (21) is simplified to:

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

where $G_1(\Phi) = \beta_1 l - \Phi(h-l)\beta_2$ and $G_0(\Phi) = \beta_0 2l - \Phi(h-l)(\beta_1+\beta_2)$. Notice that

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

Thus, we must have $\eta(1) \leq \eta(0)$ and $\rho^j(\theta_{-i},\theta_i) \leq \rho^j(\theta_{-i}, ll)$ if $\theta_i^j = l$ and $\theta_i^k = h$ for $k \neq j$.

Let $t(0), t(1)$ and $t(2)$ denote transfers for types with 0, 1 and 2 high-valuation goods respectively. Now we show that type-(h,h) will have a strict incentive to mis-report as type-(l,l), a contradiction. To see this, recall from the discussion prior to Proposition 1 that in the solution to the original problem (9), the downward adjacent incentive constraints for type-(h,h) and types (h,l) or (l,h) must bind if the postulated mechanism is not ex post efficient (which is clearly the case). Let $U\left(\tilde{\theta}_i|\theta_i\right)$ denote the expected utility for type-θ_i agent from truth-telling and from mis-reporting as type-$\tilde{\theta}_i$. We have:

$$\begin{aligned}
& U(hh|hh) = \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},hh)+\rho^2(\theta_{-i},hh)\right]h - t(2) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},hl)h+\rho^2(\theta_{-1},hl)\eta(1)h\right] - t(1) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},ll)\eta(0)h+\rho^2(\theta_{-i},ll)\eta(0)l\right] - t(0) + \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^2(\theta_{-i},hl)\eta(1)(h-l) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},ll)\eta(0)h+\rho^2(\theta_{-i},ll)\eta(0)h\right] - t(0) \\
& + \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^2(\theta_{-i},hl)\eta(1)-\rho^2(\theta_{-i},ll)\eta(0)\right](h-l) \\
\leq & \; U(ll|hh)
\end{aligned}$$

where the first inequality follows from binding downward adjacent incentive constraint for type-(h,h), the second equality follows from binding downward adjacent incentive constraint for type-(h,l), the third equality follows from re-arranging and the last inequality follows from the result above that $\eta(1)\leq\eta(0)$ and $\rho^j(\theta_{-i},\theta_i)\leq\rho^j(\theta_{-i},ll)$. Thus, constraint (A23) binds or is violated, a contradiction.

Claim A2 *Suppose that $\frac{\beta_1}{2}<\frac{\beta_0\beta_2}{1-\beta_0}$. Then the downward adjacent constraints for type-(h,h) and type-(h,l) or type-(l,h) must both bind in the optimal solution.*

To see this, first note that it could not be the case that both the downward adjacent constraints for type-(h,h) and type-(h,l) are slack in the optimal solution. If so, then it will immediately follow that type-(h,h) will have strict incentive to mis-report as type-(l,l), contradicting Claim A1. So suppose that the downward adjacent incentive constraint for type-(h,l) [or type-(l,h)] binds, but the downward adjacent incentive constraint for type-(h,h) is slack, i.e.,

$$\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},hh)+\rho^2(\theta_{-i},hh)\right]h-t(2) > \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i},hl)h+\rho^2(\theta_{-i},hl)\eta(1)h\right]-t(1). \tag{A24}$$

Given (A24) and a binding (A23), the optimality conditions with respect to $\eta_i^2(hl)$ are the complementary slackness conditions together with the first order condition:

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

where $\gamma_i^2(hl)$ and $\phi_i^2(hl)$ are the multipliers associated with $\eta_i^2(hl)\geq 0$ and $\eta_i^2(hl)\leq 1$ respectively. This immediately implies that $\eta_i^2(hl)=\eta_i^1(lh)=\eta(1)=1$ in optimum.

Similarly, given (A24) and a binding (A23) the optimality conditions with respect to $\rho^1(\theta)$ are the complementary slackness conditions together with the first order condition

$$\begin{aligned}
& \left\{h\left[H^1(\theta,1)+H^1(\theta,2)\right]+l\left[L^1(\theta,1)+\eta(0)L^1(\theta,0)\right]-cn\right\}+\lambda(1)\frac{2}{\beta_1}\left[H^1(\theta,1)h+L^1(\theta,1)\eta(0)l\right] \\
& +\psi\frac{1}{\beta_2}H^1(\theta,2)h-\psi\frac{1}{\beta_0}\boldsymbol{\beta}_{-i}(\theta_{-i})\eta(0)H^1(\theta,2)-\lambda(1)\frac{\eta(0)}{\beta_0}\left[H^1(\theta,1)h+L^1(\theta,1)l\right]+\lambda(0)\frac{1}{\beta_0}L^1(\theta,0)l-\Lambda cn \\
& +\frac{\gamma^1(\theta)-\phi^1(\theta)}{\boldsymbol{\beta}(\theta)}=0,
\end{aligned} \tag{A25}$$

where $\gamma^1(\theta) \geq 0$ and $\phi^1(\theta) \geq 0$ are respectively the multipliers associated with $\rho^1(\theta) \geq 0$ and $1-\rho^1(\theta) \geq 0$. Now consider the first order condition (A25) for profile $\theta' = (\theta_{-i}, lh)$ and $\theta'' = (\theta_{-i}, ll)$ for any $\theta_{-i} \in \Theta^{n-1}$, we have $L^1(\theta', 0) = L^1(\theta'', 0) - 1$, $H^1(\theta', 1) = H^1(\theta'', 1)$, $L^1(\theta', 1) = L^1(\theta'', 1) + 1$, $H^1(\theta', 2) = H^1(\theta'', 2)$. Thus the terms in the first two lines in (A25) is higher under θ'. Thus it must be the case that $\rho^1(\theta') \geq \rho^1(\theta'')$, i.e., $\rho^1(\theta_{-i}, hl) \geq \rho^1(\theta_{-i}, ll)$ for every θ_{-i}. Together with the fact that $\eta(1) = 1 \geq \eta(0)$ we established above, we thus have:

$$\sum_{\theta_{-i}\in\Theta_{-i}} \beta(\theta_{-i})\rho^1(\theta_{-i}, lh)\eta(1) \geq \sum_{\theta_{-i}\in\Theta_{-i}} \beta(\theta_{-i})\rho^1(\theta_{-i}, ll)\eta(0), \tag{A26}$$

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

$$\begin{aligned}
U(hh|hh) &= \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, hh) + \rho^2(\theta_{-i}, hh)\right]h - t(2) \\
&= \overbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, ll)\eta(0)h + \rho^2(\theta_{-i}, ll)\eta(0)h\right] - t(0)}^{U(ll|hh)} \\
&= \overbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, ll)\eta(0)l + \rho^2(\theta_{-1}, ll)\eta(0)h\right] - t(0)}^{U(ll|lh)} + (h-l)\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^1(\theta_{-i}, ll)\eta(0) \\
&= \overbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, lh)\eta(1)l + \rho^2(\theta_{-1}, lh)h\right] - t(1)}^{U(lh|lh)} + (h-l)\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^1(\theta_{-i}, ll)\eta(0) \\
&= \overbrace{\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\left[\rho^1(\theta_{-i}, lh)\eta(1)h + \rho^2(\theta_{-1}, lh)h\right] - t(1)}^{U(lh|hh)} \\
&\quad + (h-l)\left[\sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^1(\theta_{-i}, ll)\eta(0) - \sum_{\theta_{-i}\in\Theta_{-i}} \boldsymbol{\beta}_{-i}(\theta_{-i})\rho^1(\theta_{-i}, lh)\eta(1)\right] \\
&\leq U(lh|hh),
\end{aligned} \tag{A27}$$

where the second equality follows from the binding (A23), third and fifth equality are simply rearrangement of terms, and the fourth equality follows from the postulated downward adjacent incentive constraint for type-(l, h) and the last inequality follows from (A26). Inequality (A27) contradicts the postulated (A24).

The statement that the downward adjacent incentive constraint for type-(h, l), or type-(l, h) must bind can be proved analogously.

We thus conclude from Claims A1 and A2 that all the incentive constraints bind for this case. Examining these incentive constraints and using the key "Paradox of Voting"-like Lemma 6, we have that in the limit as $n \to \infty$, these constraints imply:

$$\begin{aligned}
2h - t(2) &= h + \eta(1)h - t(1) \\
2h - t(2) &= 2h\eta(0) - t(0) \\
h + \eta(1)l - t(1) &= (h+l)\eta(0) - t(0)
\end{aligned}$$

where the first equality is the limit of the binding constraint (A23) and the second is the limit of the binding downward adjacent incentive constraints (10) for types (h, h) and (h, l) respectively. These three equalities immediately imply that in the limit as $n \to \infty$, $\eta(1) = \eta(0)$. This inclusion rule immediately implies that bundling is not used in the optimal mechanism in the limit (as well as when n is sufficiently large). We thus conclude that the limit provision and inclusion rules for this case is identical to the single good case analyzed in Section 4.1. The desired results follow. ■