Public Good Provision Re-Examined

Loren K. Fryxell*

May 30, 2024

(Latest Version Here)

Abstract

I write down the government's public good provision problem from first principles and, contrary to popular wisdom, find a solution. I call it the costsharing pivotal mechanism. Both the statement of the problem and the solution are new. The cost-sharing pivotal mechanism is strategy-proof, employs a utilitarian decision rule—generalizing the efficient decision rule by scaling each individual's monetary value by an arbitrary welfare weight—satisfies a new participation constraint, satisfies a new fairness principle, and is ex-post budget-balanced asymptotically in large populations. Moreover, I show that the common methodological simplification of taking values to be net of one's cost share is not without loss of generality, standard participation constraints are not well-suited for the government's public good provision problem, and the most well-known mechanism for public good provision, the Clarke mechanism, violates a basic fairness constraint: if nothing is produced, no one should pay.

1 Introduction

A government would like to decide whether or not to produce a public good whose value to each individual is privately known and whose cost of production is publicly known. Through some previously enacted democratic process, the government has decided what constitutes a fair share of the production cost for each individual as a

^{*}Department of Economics and Global Priorities Institute, University of Oxford (e-mail: loren.fryxell@economics.ox.ac.uk). I wish to thank Jeff Ely, Eddie Dekel, Meg Meyer, Alessandro Pavan, Alex Teytelboym, and Asher Wolinsky for their invaluable comments, discussions, and guidance. I also wish to thank Gustav Alexandrie, Florian Brandl, Alexis Ghersengorin, Joe Hardwick, Marcus Pivato, Jim Schummer, Teru Thomas, William Thomson, John Weymark, and seminar participants at Northwestern, Oxford, Bristol, SSCW 2018, and TTW 2023 for helpful comments and discussions. Finally, I wish to thank Luis Mota Freitas for superb research assistance. All errors are my own.

function of their observable characteristics, such as income and location.¹ A utilitarian government produces the public good if and only if the social welfare benefits outweigh the social welfare costs. The government asks each individual to report their willingess to pay for the public good and, for each individual, scales this report by an individual-specific welfare weight, which converts monetary units of value into interpersonally comparable welfare units. Note that the social welfare costs depend on the allocation of costs between, e.g., the rich and the poor, since while the total monetary cost remains the same, the total welfare cost depends on who pays.

The government would like to design a decision procedure that (a) provides incentives for each individual to truthfully report their willingness to pay for the public good, (b) provides incentives for each individual to participate in the process, (c) produces the welfare-maximizing quantity of the public good, (d) raises exactly the amount of revenue necessary to produce this quantity, and (e) does not ask anyone to pay what has been deemed an unfair amount. Call this the *government's public good provision problem*.

It is well known that no decision procedure can satisfy all five criteria at once— Green and Laffont (1979) show that no mechanism can satisfy (a), (c), and (d) alone. However, I show that all five can be satisfied when the population is large, and, moreover, that there is—up to vanishingly small perturbations—a unique solution. I call this the *cost-sharing pivotal mechanism*.

In particular, the cost-sharing pivotal mechanism satisfies (a), (b), (c), and (e) exactly and (d) arbitrarily closely as the population goes to infinity. Formally, these properties correspond to (a) strategy-proofness, (b) cost-sharing universal participation, (c) utilitarianism, (d) asymptotic ex-post budget-balance, and (e) the fair pricing principle. Strategy-proofness is standard in the literature and says that reporting one's private information truthfully should be a weakly dominant strategy for each individual. Asymptotic properties relating to budget-balance are also standard, though the precise specification of asymptotic ex-post budget-balance here is new. Cost-sharing universal participation and the fair pricing principle are each new concepts. Costsharing universal participation states that each individual should always prefer to participate in the mechanism rather than to not participate, receive the alternative that would have been chosen without her, and be taxed her fair share of its cost. The fair pricing principle states that each individual should never pay more than her monetary value for what is produced or her fair share of its cost, whichever is larger.

An efficient decision rule maximizes the sum of individuals' willingness to pay for the good minus its cost to produce. It is standard within mechanism design to require efficiency as an axiom or to select an optimal mechanism with respect to this objec-

¹For instance, those who earn more and live closer to the public good might be assigned a higher fair cost share.

tive. What I will refer to as a *utilitarian* decision rule maximizes a weighted sum of individuals' willingness to pay for the good minus a weighted sum of individuals' cost shares. The weights, called welfare weights, convert monetary units of value into interpersonally comparable welfare units. This idea has a long history in welfare economics with the study of unweighted and weighted benefit-cost analysis (BCA). An efficient decision rule is equivalent to an *unweighted* benefit-cost analysis, and a utilitarian decision rule is equivalent to a *weighted* benefit-cost analysis. Within welfare economics, the prominent view is that a weighted benefit-cost analysis is strongly preferred to an unweighted benefit-cost analysis.² Indeed, Fleurbaey and Abi-Rafeh (2016) write that "it is hard to defend unweighted BCA because it is connected to no good welfare economics." In recent years, several papers in mechanism design have begun to use a utilitarian objective.³ This paper imposes it as an axiom.

As a motivating example,⁴ suppose a city of 1,000 people would like to decide whether or not to construct a public good—say, a library—which costs \$550,000. The government has previously decided that each individual's fair share of the cost of the public good should be their proportionate share of the total cost by income.

#	income	welfare weight	WTP	fair cost share	welfare benefit	welfare cost
400	\$30,000	2	\$600	\$300	1,200	600
500	\$60,000	1	\$400	\$600	400	600
80	\$100,000	.6	\$200	\$1,000	120	600
20	\$250,000	.24	\$200	\$2,500	48	600
1,000			\$460,000	\$550,000	690,560	600,000

Figure 1: The leftmost column indicates how many individuals there are at each income level, and the bottom row indicates the total of each column, weighted by the number of individuals at each income level.

Notice that most of the monetary value comes from those with lower incomes, as those with higher incomes often purchase books and other media rather than borrow them. Moreover, this discrepancy is even more pronounced when we look at welfare values,

²See Blackorby and Donaldson (1990), Adler (2012), Adler (2016), Boadway (2016), Fleurbaey and Abi-Rafeh (2016), Fleurbaey et al. (2013), and Bressler and Heal (2022).

³See Dworczak, Kominers and Akbarpour (2021), Pai and Strack (2022), Akbarpour, Dworczak and Kominers (2023), Reuter and Groh (2023), and Akbarpour et al. (2024). It is worth mentioning that the purpose of these papers is to study a situation where the use of welfare weights gives importantly different results than without, where the results are already known or uninteresting. In contrast, the results of this paper are still novel and relevant when considering the special case of an efficient decision rule (indeed, the paper was originally written this way).

⁴This example is originally from Carroll (2023) and inspired my decision to use a utilitarian rather than efficient decision rule.

since a given willingness to pay corresponds to a higher welfare value the lower one's income. Indeed, by only looking at monetary values, we would conclude that the library should not be built—as the monetary costs (\$550,000) outweigh the monetary benefits (\$460,000). Standard in mechanism design, this is what the efficient decision rule prescribes. However, if we take into account that an additional dollar contributes more welfare to those with lower incomes than it does to those with higher incomes, then we see that the welfare benefits of constructing the library (690,560) exceed the welfare costs (600,000). So to maximize social welfare the library should be built, as the utilitarian decision rule prescribes.

Ideally, we would like each individual to report their willingness to pay truthfully, implement the utilitarian decision, and ask everyone to pay their fair cost share. However, this procedure has infinitely perverse incentives. If an individual's value for the good is greater than her fair share of its cost, it is weakly dominant for her to report an infinitely large value. If an individual's value for the good is less than her fair share of its cost, it is weakly dominant for her to report a zero value. Yet, a simple refinement to this procedure can correct these infinitely perverse incentives in a way that, in large populations, does not disturb the equilibrium outcome. That is, in large populations, everyone in fact pays their fair cost share.

A classic result in mechanism design is that a mechanism is strategy-proof and efficient if and only if it is a Groves mechanism: that is, if and only if each individual pays the monetary loss she imposes on society, excluding her own monetary benefits from the public good, plus a term that is constant in her report (and hence, does not affect her incentives).⁵ A simple generalization of this result implies that a mechanism is strategy-proof and *utilitarian* if and only if it is what I will call a *generalized* Groves mechanism: that is, if and only if each individual pays her monetary equivalent of the welfare loss she imposes on society, excluding her own welfare benefits from the public good, plus a term that is constant in her report.

This means that the only flexibility which remains in searching for a solution to the government's public good provision problem is the choice of this constant term. The answer is simple: in addition to the externality she imposes on society, an individual should pay her fair share of the cost of what would have been produced without her.

The cost-sharing pivotal mechanism employs a utilitarian decision rule and charges each individual her monetary equivalent of the welfare loss she imposes on society, excluding her own welfare benefits from the public good, plus her fair share of the cost of what would have been produced without her.

 $^{^{5}}$ This mechanism was first introduced by Groves (1973). This result was first shown by Green and Laffont (1977) and later generalized by Holmström (1979).

In the library example, the cost-sharing pivotal mechanism opts to construct the library and taxes each individual exactly her fair cost share. This is because no individual sways the decision with her report. Hence, the externality each individual imposes on society is zero,⁶ and what would have been produced without them is precisely what is produced with them. Indeed, the outcome in which no individual sways the decision occurs with arbitrarily high probability as the population gets large, and consequently the *generic* outcome of the cost-sharing pivotal mechanism is that everyone pays exactly their fair share of the cost of what is produced.

When everyone pays their fair cost share, cost-sharing universal participation and the fair pricing principle are immediately satisfied. But the cost-sharing pivotal mechanism also satisfies these conditions globally, no matter the reports of the individuals and no matter the outcome of the mechanism. Perhaps surprisingly, this mechanism—including in the special case of an efficient decision rule—is new. Why is it that such a simple mechanism had not been identified before?

The standard approach to modeling production costs in mechanism design is to exogenously assign cost shares to individuals, embed these cost shares into their value for the alternatives, and proceed with the analysis as if there were no production costs to begin with. Each individual is required to pay their share of the cost of whatever alternative is chosen, and they simply report their value for each alternative net of this amount. This quantity is called the individual's *net value*, and we may call this approach the *net value approach*. The idea is that it is without loss of generality to study environments without production costs, since production costs can always be baked into the values in this way.⁷ Indeed, Green and Laffont (1979, p. 31) even contend that "there is no real alternative to this approach." This paper shows otherwise. In particular, keeping values and costs separate provides a richer mathematical and conceptual structure which allows for a wider class of desiderata (e.g., cost-sharing universal participation and the fair pricing principle), mechanisms (e.g., the cost-sharing pivotal mechanism), and proof techniques (e.g., Theorem 4) than can be constructed when combining values and costs together.

The most well-known mechanism in the family of Groves mechanisms is the *pivotal* mechanism.⁸ A pivotal mechanism employs an efficient decision rule and charges each individual the monetary loss she imposes on society, excluding her own monetary benefits from the public good. Incorporating net values into the pivotal mechanism produces the *Clarke mechanism*.⁹ This has become the canonical mechanism for

⁶To see this, note that by removing any individual from the welfare benefits, the total welfare benefit is at least 689,360 which still exceeds the welfare cost of 600,000.

⁷See Green and Laffont (1979, p. 29). See also Moulin (1988, p. 205), Varian (1992, p. 426), and Mas-Colell, Whinston and Green (1995, p. 877).

⁸A pivotal mechanism is also known as a VCG mechanism after Vickrey (1961), Clarke (1971), and Groves (1973).

 $^{^{9}}$ The Clarke mechanism was first proposed by Clarke (1971). I discuss the Clarke mechanism

public good provision, and, within the class of strategy-proof mechanisms, there has been little exploration beyond it. Despite that, the Clarke mechanism violates costsharing universal participation, the fair pricing principle, and an even more basic condition I call *no-extortion*: if nothing is produced, no one should pay.

Example 1 (The Clarke mechanism violates cost-sharing universal participation, the fair pricing principle, and no-extortion). Suppose Alice and Benjamin are deciding whether to construct a public park. The cost of the park is \$4, and each individual's fair share of the cost is \$2. Alice values the park at \$0, and Benjamin values the park at \$3. Suppose they have equal welfare weights, so that the utilitarian and efficient decision rule coincide.

The Clarke mechanism is equivalent to embedding cost shares into individual values, removing all production costs, and running a pivotal mechanism. Alice values the park net of her cost share at -\$2, and Benjamin values the park net of his cost share at \$1. Taking these as their values, the efficient decision is not to construct the park (in which case Benjamin's total welfare is 0) and the efficient decision ignoring Alice's preferences is to construct the park (in which case Benjamin's total welfare is 1), so Alice's pivotal transfer is \$1.

This violates no-extortion (and hence also the fair pricing principle) since nothing is produced but Alice must still pay. This violates cost-sharing universal participation since Alice would strictly prefer to not participate, receive the alternative that would have been chosen without her (no public good), and pay her fair share of its cost (\$0), rather than to participate, receive no public good, and pay \$1. In the cost-sharing pivotal mechanism, the park is not constructed and Alice and Benjamin both pay zero. \Box

The fact that the Clarke mechanism violates such participation and fairness constraints is perhaps not surprising, as these criteria, and the solution that comes out of them, arise from thinking about values and costs separately, while the Clarke mechanism arises from combining values and costs into a single term and running a pivotal mechanism. Given the historic focus on modeling production costs with net values, and the prevailing view that there is no alternative to this approach, it is perhaps also not surprising that further exploration in the space of strategy-proof public good provision mechanisms has not been pursued, as there are few to no natural alternatives to the Clarke mechanism when using net values.

One notable property of the Clarke mechanism is that it never runs a budget deficit. Nevertheless, I argue that what really matters for practical large-scale mechanism design is that future changes to tax policy which balance the budget in the long run are sufficiently small that they do not affect individuals' incentives today. I call

and how it compares to the cost-sharing pivotal mechanism in Section 9.3.

this the *approximate budget-balance principle*.¹⁰ In other words, what really matters for practical large-scale mechanism design is precisely asymptotic ex-post budget-balance. Moreover, it turns out that no-deficit and fairness are fundamentally in conflict: no strategy-proof and utilitarian mechanism can satisfy both no-deficit and no-extortion.

The key takeaways of the paper are as follows. (i) A utilitarian, rather than merely efficient, decision rule ought to be considered for the problem of public good provision. (ii) Costs of producing the public good should be modeled explicitly rather than embedded in the individuals' values. In particular, the net value approach is not without loss. (iii) Standard participation constraints are not well-suited for the government's public good provision problem. A new participation constraint, which I call cost-sharing universal participation, should be used instead. (iv) Fairness constraints ought to be explicitly considered in public good provision, constraining what the government can fairly ask individuals to contribute to the public good. A new fairness principle, which I call the fair pricing principle, is a natural desiderata for these environments. (v) The canonical public good provision mechanism—the Clarke mechanism—violates cost-sharing universal participation, the fair pricing principle, and an even more basic constraint I call no-extortion: if nothing is produced, no one should pay. (vi) Contrary to popular wisdom surrounding the provision of public goods, a natural solution arises to the government's public good provision problem. I call it the *cost-sharing pivotal mechanism*. It is new and, modulo vanishingly small perturbations, it is unique.

The rest of the paper is organized as follows. Section 2 contains a review of the related literature on public good provision. Section 3 introduces the model and shows that, for any strategy-proof and utilitarian mechanism, only willingness to pay can be elicited and welfare weights and cost shares must be independent of willingness to pay (Theorem 1). Section 4 introduces the cost-sharing pivotal (CSP) mechanism. Section 5 introduces cost-sharing universal participation and shows that the CSP can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and cost-sharing universal participation (Theorem 2). Section 6 introduces the fair pricing principle and shows that the CSP can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and the fair pricing principle (Theorem 3). Section 7 shows that the CSP is asymptotically expost budget balanced (Theorem 4) and that no other satisfactory mechanism consistently comes closer to budget-balance than the CSP (Proposition 5). Section 8 shows that the CSP is the unique mechanism, up to vanishingly small perturbations, that satisfies strategy-proofness, utilitarianism, cost-sharing universal participation, the fair pricing principle, and asymptotic ex-post budget-balance (Theorem 5). Section 9 discusses the net value approach and the Clarke mechanism. Section 10 shows that no

 $^{^{10}}$ See Section 7.1 for discussion.

strategy-proof and utilitarian mechanism can satisfy both no-deficit and no-extortion (Proposition 6). Section 11 concludes.

2 Related Literature

In this section, I review some of the related literature on the public good provision problem. Mailath and Postlewaite (1990) consider binary public good environments and show that in any mechanism satisfying Bayesian incentive-compatibility, interim individual-rationality, and ex-ante no-deficit, if the cost of the public good increases linearly with population size or faster, the probability that the public good is produced goes to zero as the population size goes to infinity—even when the probability that it is efficient to produce the public good converges to one.¹¹ In a similar vein, Al-Najjar and Smorodinsky (2000) consider binary public good environments and show that any mechanism satisfying Bayesian incentive-compatibility, interim individualrationality, and no-small-contributors¹² cannot raise revenues that are unbounded as the population size goes to infinity.

Xi and Xie (2023) consider binary public good environments and propose a class of mechanisms which are strategyproof and ex-post individually rational. They show that if the cost of provision grows slower than the square root of population size, these mechanisms generate an ex-ante budget-surplus asymptotically and are asymptotically efficient, while if the cost grows faster than the square root of population size, any mechanism which is strategypoof, ex-post individually-rational, and generates an ex-ante budget-surplus asymptotically must have a provision probability converging to zero. Kuzmics and Steg (2017) consider binary public good environments and show that the welfare-maximizing mechanism among all strategy-proof, ex-post individually rational, and no-deficit mechanisms is what we might call a "unanimous split-the-cost mechanism," in which each player has a fixed cost share, the good is provided if and only if all individuals' values exceed their own cost share, and each player pays her cost share if the good is provided and zero otherwise. Notice that this is a special case of Mailath and Postlewaite (1990), and hence if the cost of the public good increases linearly with population size or faster, the probability that the public good is produced in such a mechanism goes to zero as the population size goes to infinity.

Serizawa (1999) considers continuous public good environments with monotonic and quasi-concave preferences and characterizes the class of strategy-proof and ex-post budget-balanced mechanisms which are also symmetric (minimax rule), anonymous (q-rule), or symmetric and ex-post individually-rational (minimum demand rule).

¹¹See also Hellwig (2003).

 $^{^{12}{\}rm This}$ condition requires that, if an individual's expected transfer is less than some threshold, the individual's expected transfer is zero.

Laffont and Maskin (1982) show that in binary decision problems with no production costs, if the population distribution of values is symmetric around zero, the welfaremaximizing budget-balanced mechanism has a sink agent—an agent who's preferences are ignored and who receives the transfers from the other agents—and a decision rule which is efficient among the remaining agents. Nath and Sandholm (2019) consider finite decision problems with three or more alternatives and no production costs and bound the asymptotic inefficiency of strategy-proof mechanisms using strategyproof and ex-post budget-balanced mechanisms with a sink agent and a decision rule which is efficient among the non-sink agents. Moreover, they show that any strategy-proof and ex-post budget-balanced mechanism has at least one a sink agent. Drexl and Kleiner (2018) consider binary decision problems with no production costs and show that the welfare-maximizing mechanism among all anonymous, strategyproof, universal participation, and no-deficit mechanisms is qualified majority voting with threshold k, where k is the ceiling of $\frac{-n\mathbb{E}(\theta_i | \theta_i \ge 0)}{\mathbb{E}(\theta_i | \theta_i \ge 0) - \mathbb{E}(\theta_i | \theta_i \le 0)}$ and θ_i is *i*'s value for the project. Note that the use of anonymous mechanisms rules out having a sink agent.

Rob (1982) considers binary decision environments with no production costs and shows that the pivotal mechanism is asymptotically efficient—i.e., the probability that any individual has a positive transfer goes to zero and the expected per capita transfer goes to zero as the population size goes to infinity.¹³ This immediately implies an analogous result for the Clarke mechanism in a binary public good environment with production costs (using the net value approach to embed costs into values) with equal cost shares and a cost of the public good which increases linearly with population size. By contrast, I show that the cost-sharing pivotal mechanism is asymptotically ex-post budget-balanced in a finite public good environment with arbitrary sequences of costs and arbitrary sequences of cost shares. See Sections 7.2 and 9.1 for a discussion of the importance of such robustness.

3 Model

A standard mechanism design environment with transfers is defined by $(I, Y, \Theta, \{v_i\}_{i \in I})$, where I is a set of n individuals, Y is a set of social alternatives, $\Theta = \Theta_1 \times \ldots \times \Theta_n$ is a type space, and $v_i : Y \times \Theta_i \to \mathbb{R}$ is individual *i*'s willingness to pay for each alternative y given her type θ_i . To analyze the government's public good provision problem, I add four elements to the standard environment: observable characteristics, production costs, fair cost shares, and welfare weights.

A public good provision environment with transfers is defined by

 $\mathcal{E} = (I, Y, \Theta, Z, \{v_i\}_{i \in I}, c, \{c_i\}_{i \in I}, \{\lambda_i\}_{i \in I}).$

¹³See also Green, Kohlberg and Laffont (1976), Green and Laffont (1979), and Mitsui (1983).

where $Z = Z_1 \times \ldots \times Z_n$ is a space of observable characteristics for each individual $i, c: Y \to \mathbb{R}$ is the production \cos^{14} for each alternative $y, c_i: Y \times Z \to \mathbb{R}$ is the amount society has agreed is fair for i to contribute if y is produced (where $\sum_{i \in I} c_i(y, z) = c(y)$ for all y, z), and $\lambda_i: Z \to \mathbb{R}_+$ is *i*'s welfare weight.

Theorem 1 elucidates why cost shares and welfare weights must depend only on observable characteristics and not on individuals' private information: any mechanism which is utilitarian and strategy-proof selects cost shares and welfare weights independently of any individual's private information. That being said, this separation is also pragmatic. It is important that a public good provision mechanism is simple and transparent. How fair cost shares and welfare weights are computed should be plain and easy to understand, and decided by a combination of democratically elected officials, researchers, and ordinary citizens. See Appendix A for some examples of how a government might construct fair cost shares from observable characteristics. Alternatively, by collecting individual cost shares through the tax system rather than by an explicit payment, what constitutes a fair cost share can simply be derived from the current tax system. See Section 11 for a discussion of this method.

An outcome is a social alternative $y \in Y$ and a vector of transfers $t \in \mathbb{R}^n$ representing the amount each individual *i* is asked to pay. Individual *i*'s preferences over outcomes are quasilinear and represented by $u_i(y, t, \theta) = v_i(y, \theta_i) - t_i$.

A direct revelation mechanism $f: \Theta \to Y \times \mathbb{R}^n$ is a mapping from type profiles to outcomes. Restricting attention to direct revelation mechanisms is without loss of generality by the revelation principle for dominant strategies.¹⁵ In these environments, f can be represented by a pair $f = (\alpha, \tau)$, where $\alpha : \Theta \to Y$ is a decision rule and $\tau : \Theta \to \mathbb{R}^n$ is a transfer rule. Each individual *i* reports their type θ_i to the mechanism. The mechanism implements social alternative $\alpha(\theta)$ at cost $c(\alpha(\theta))$ and collects transfers $\tau_i(\theta)$ from each *i*.

It is standard in mechanism design to consider an efficient decision rule $\alpha(\theta) \in \arg \max_{y \in Y} \sum_{i \in I} (v_i(y, \theta_i) - c_i(y, z))$. One justification for this is that any other decision is Pareto dominated by the efficient decision plus some redistributive transfers. However, in practice governments often do not (or cannot) implement such transfers and, in the absence of these transfers, the efficient decision rule looses its appeal.¹⁶ It is then important to consider a decision rule which accounts for the fact that a \$100 willingness to pay represents a higher welfare benefit to a poorer individual than to

¹⁴I might have called this an implementation cost, since, more generally, this is the monetary cost of implementing a decision (e.g., passing a bill, electing an official, or producing a public good). I call this a production cost for simplicity, since this is more intuitive when dealing with public goods.

 $^{^{15}}$ See, e.g., Gibbard (1973) or Mas-Colell, Whinston and Green (1995, p. 871)

¹⁶ "Checking that the unweighted sum of WTP_i is positive is equivalent to checking that the individuals who benefit from the change could compensate the losers. This approach has, however, been completely disqualified by welfare economists. In particular, the fact that compensation could be made is not a sufficient justification when it is not really made" (Fleurbaey et al., 2013).

a richer one and, similarly, a \$100 tax represents a higher welfare cost to a poorer individual than to a richer one.

One way to do this is to take a *weighted* sum of each individual's willingness to pay, i.e., $\alpha(\theta) \in \arg \max_{y \in Y} \sum_{i \in I} \lambda_i(z)(v_i(y, \theta_i) - c_i(y, z))$. The quasilinear utility representation $\lambda_i(z)(v_i(y, \theta_i) - t_i)$ can be understood as a first-order approximation to a more general representation with utility of money ϕ_i . That is,

$$\phi_i(v_i(y,\theta_i) - t_i + w) \approx \phi_i(w) + \phi'_i(w)(v_i(y,\theta_i) - t_i),$$

where ϕ_i is *i*'s interpersonally comparable utility of money and *w* is her current wealth level. Hence, we can think of *i*'s welfare weight λ_i as her marginal utility of income (optionally, with additional inequality aversion baked in) and α as a *utilitarian* decision rule.

Definition 1. A decision rule $\alpha : \Theta \to Y$ is *utilitarian* if, for all θ , $\alpha(\theta) \in A^*(\theta)$, where

$$A^*(\theta) = \underset{y \in Y}{\operatorname{arg\,max}} \sum_{i \in I} \lambda_i(z) (v_i(y, \theta_i) - c_i(y, z)).$$

A decision rule α is efficient if α is utilitarian and $\lambda_i(z) = \lambda_j(z)$ for all $i, j \in I$ and $z \in Z$. A mechanism $f = (\alpha, \tau)$ is utilitarian if α is utilitarian.

 $v_i(y, \theta_i)$ is the monetary value of y to i and, assuming i pays her fair share, $c_i(y, z)$ is the monetary cost of y to i. Scaling this by $\lambda_i(z)$ translates monetary units for i into interpersonally comparable welfare units. The government seeks to maximize aggregate social welfare. This captures the idea that, as in Figure 1, the welfare-maximizing decision and the efficient decision may come apart. In particular, a utilitarian government may seek to produce a good whose monetary costs exceed its monetary benefits (but whose welfare benefits exceed its welfare costs).

First, notice that the efficient decision rule is a special case of the utilitarian decision rule when the welfare weights are all equal to one. Second, notice that, unlike with the efficient decision rule, the choice of cost shares matters for the optimal decision. Costs borne by richer individuals cause less welfare loss than costs borne by poorer individuals.¹⁷ Third, notice that the utilitarian decision rule presupposes that individuals pay their fair share, even if in fact they may not. This is a simplification which is sensible when combined with conditions implying that such an outcome occurs with

¹⁷One might then think that the richest individual should bear the entire cost of the public good. While this would indeed inflict the smallest welfare cost on society, it is unlikely to be considered fair on the whole. An alternative which may better align with fairness principles is that each individual pays an equal share of their wealth. With log utility of money, this equalizes the welfare costs paid across individuals.

arbitrarily high probability, as is the case here.¹⁸ This is analogous to an efficient decision rule, which presupposes that a mechanism is budget-balanced, even if in fact it may not be. Again, this is a simplification which is sensible when combined with conditions implying that such an outcome occurs with arbitrarily high probability, e.g., asymptotic ex-post budget-balance.

We seek incentives for individuals to reveal their private information truthfully, regardless of their beliefs about the other players' actions or private information. A mechanism is strategy-proof if it is a dominant strategy for each individual i to report her type θ_i truthfully to the mechanism.

Definition 2. A mechanism $f = (\alpha, \tau)$ is strategy-proof if for all i, θ'_i, θ , $v_i(\alpha(\theta), \theta_i) - \tau_i(\theta) \ge v_i(\alpha(\theta'_i, \theta_{-i}), \theta_i) - \tau_i(\theta'_i, \theta_{-i}).$

We now proceed to state and prove Theorem 1. Let $\Theta_i(Y_0, Y_1)$ be the set of types θ_i such that if $y, y' \in Y_0$ or $y, y' \in Y_1$ then $v_i(y, \theta_i) = v_i(y', \theta_i)$.

Definition 3. An environment admits binary preferences if for each *i* there exists a partition $\{Y_0, Y_1\}$ of *Y* such that $\Theta_i(Y_0, Y_1)$ is non-degenerate, i.e., there exists $y_0 \in Y_0$, $y_1 \in Y_1$, and $\theta_i, \theta'_i \in \Theta_i(Y_0, Y_1)$ such that $v_i(y_1, \theta_i) - v_i(y_0, \theta_i) \neq v_i(y_1, \theta'_i) - v_i(y_0, \theta'_i)$.

Note that any environment which admits 1) a completely indifferent type and 2) a type which strictly prefers some alternative y_1 to another y_0 , where all other alternatives are either indifferent to y_1 or y_0 , admits binary preferences. Consider any environment which admits binary preferences. Theorem 1 states that, for any strategy-proof and utilitarian mechanism, only willingness to pay can be elicited and welfare weights and cost shares must be independent of willingness to pay.¹⁹

Theorem 1. Consider any standard mechanism design environment with transfers. Let θ_i^P represent private information to individual *i* which is sufficient to pin down her willingness to pay for each alternative $y \in Y$. Let θ_i^{-P} represent any additional information, private or not. Let α be a generalized affine maximizing decision rule of the form

$$\alpha(\theta^{P}, \theta^{-P}) \in \underset{y \in Y}{\arg \max} \sum_{i \in I} \lambda_{i}(\theta^{P}, \theta^{-P})v_{i}(y, \theta^{P}_{i}) + k(y, \theta^{P}, \theta^{-P})$$

¹⁸In particular, mechanisms which satisfy strategy-proofness, utilitarianism, asymptotic ex-post budget-balance, and at least one of cost-sharing universal participation and the fair pricing principle satisfy the property that individuals pay their fair share with arbitrarily high probability (Theorem 5).

¹⁹Roberts (1979) and Carbajal, McLennan and Tourky (2013) both characterize affine maximizing decision rules as the class of decision rules which are dominant-strategy implementable. However, neither result applies here because our domain of interest does not satisfy their richness assumptions. In particular, the assumption that values for the public good are non-decreasing violates Roberts' (1979) universal domain and Carbajal, McLennan and Tourky's (2013) flexibility.

for some $\lambda_i : \Theta^P \times \Theta^{-P} \to \mathbb{R}_{++}$ and $k : Y \times \Theta^P \times \Theta^{-P} \to \mathbb{R}$. For any environment that admits binary preferences, α is dominant-strategy implementable if and only if

- 1. Θ^{-P} is observable and
- 2. $\{\lambda_i\}_{i \in I}$ and k do not depend on θ^P .

That is,

$$\alpha(\theta^P, \theta^{-P}) \in \underset{y \in Y}{\operatorname{arg\,max}} \sum_{i \in I} \lambda_i(\theta^{-P}) v_i(y, \theta^P_i) + k(y, \theta^{-P})$$

for some $\lambda_i : \Theta_i^{-P} \to \mathbb{R}_{++}$ and $k : Y \times \Theta^{-P} \to \mathbb{R}$, where Θ^P is private information and Θ^{-P} is observable.

Proof. Suppose $\alpha(\Theta) = \{y_0, y_1\}$. By Carbajal, McLennan and Tourky (2013, Proposition 1), for any standard environment $\mathcal{E} \in \mathbb{E}$, α is dominant-strategy implementable if and only if for all $i \in I$ and all $\theta_{-i} \in \Theta_{-i}$, there exists a threshold $T_i(\theta_{-i}) \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that for all $\theta_i \in \Theta_i$, $v_i(y_1, \theta_i) - v_i(y_0, \theta_i) > T_i(\theta_{-i})$ implies $\alpha(\theta) = y_1$ and $v_i(y_1, \theta_i) - v_i(y_0, \theta_i) < T_i(\theta_{-i})$ implies $\alpha(\theta) = y_0$.

Consider any environment that admits binary preferences. Consider any individual i and any partition $\{Y_0, Y_1\}$ of Y such that $\Theta_i(Y_0, Y_1)$ is non-degenerate. The previous result implies that α is strategy-proof for i on $\Theta_i(Y_0, Y_1)$ only if for all $\theta_{-i} \in \Theta_{-i}$, there exists a threshold $T_i(\theta_{-i}) \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that for all $\theta_i \in \Theta_i(Y_0, Y_1), v_i(Y_1, \theta_i) - v_i(Y_0, \theta_i) >$ $T_i(\theta_{-i})$ implies $\alpha(\theta) \in Y_1$ and $v_i(Y_1, \theta_i) - v_i(Y_0, \theta_i) < T_i(\theta_{-i})$ implies $\alpha(\theta) \in Y_0$, where $v_i(Y_0, \theta_i) \equiv v_i(y_0, \theta_i)$ for any $y_0 \in Y_0$ and $v_i(Y_1, \theta_i) \equiv v_i(y_1, \theta_i)$ for any $y_1 \in Y_1$.

Let α be a generalized affine maximizing decision rule, let θ_i^P represent private information to individual *i* which is sufficient to pin down her willingness to pay for each alternative, and let θ_i^{-P} represent any additional information, private or not. Let $V_{Y_0}^i(\theta^P, \theta^{-P})$ be the value of the affine maximizer within Y_0 when *i* is indifferent between all $y_0 \in Y_0$, i.e.,

$$V_{Y_0}^{i}(\theta^{P}, \theta^{-P}) = \max_{y_0 \in Y_0} \sum_{j \neq i} \lambda_j(\theta^{P}, \theta^{-P}) v_j(y_0, \theta_j^{P}) + k(y_0, \theta^{P}, \theta^{-P}),$$

and likewise for $V_{Y_1}^i(\theta^P, \theta^{-P})$. Then²⁰

$$T_i(\cdot) = \frac{1}{\lambda_i(\theta^P, \theta^{-P})} \Big(V_{Y_0}^i(\theta^P, \theta^{-P}) - V_{Y_1}^i(\theta^P, \theta^{-P}) \Big).$$

 α is strategy-proof only if, for every i, T_i does not depend on i's private information. Hence, λ_i , λ_j for every $j \neq i$, and k cannot depend on θ_i^P . But this holds for every i, so α

$$\max_{y_1 \in Y_1} \sum_{i \in I} \lambda_i(\theta^P, \theta^{-P}) v_i(y_1, \theta^P_i) + k(y_1, \theta^P, \theta^{-P}) > \max_{y_0 \in Y_0} \sum_{i \in I} \lambda_i(\theta^P, \theta^{-P}) v_i(y_0, \theta^P_i) + k(y_0, \theta^P, \theta^{-P}).$$

 $[\]overline{\frac{20 \text{To see this, note that } v_i(Y_1, \theta_i^P) - v_i(Y_0, \theta_i^P)} > \frac{1}{\lambda_i(\theta^P, \theta^{-P})} (V_{Y_0}^i(\theta^P, \theta^{-P}) - V_{Y_1}^i(\theta^P, \theta^{-P})) \text{ if and only if }$

is strategy-proof only if

$$\alpha(\theta^P, \theta^{-P}) \in \underset{y \in Y}{\operatorname{arg\,max}} \sum_{i \in I} \lambda_i(\theta^{-P}) v_i(y, \theta^P_i) + k(y, \theta^{-P})$$

for some $\lambda_i : \Theta^{-P} \to \mathbb{R}_{++}$ and $k : Y \times \Theta^{-P} \to \mathbb{R}$, where Θ^P is private information and Θ^{-P} is observable. To complete the proof, note that this α is an affine maximizer in the sense of Roberts (1979), so is dominant-strategy implementable.

In the upcoming sections, I suppress the dependence of the welfare weights λ_i and cost shares c_i on observable characteristics z, since this will not be necessary for the analysis. I reintroduce it in Section 7, where the dependence is key.

4 The Cost-Sharing Pivotal Mechanism

Suppose that the space of admissible values for each i, $\{v_i(\cdot, \theta_i) : \theta_i \in \Theta_i\}$, is convex. Moreover, suppose that for each i the fully indifferent type, denoted $0 \in \Theta_i$ where $v_i(y,0) = v_i(y',0)$ for all $y, y' \in Y$, is admissible and that $A^*(\theta)$ is non-empty for all θ . Let $\mathbb{E}_X \subset \mathbb{E}$ be the set of such convex public good environments.

For any convex environment, the class of Groves mechanisms²¹ fully characterizes the class of strategy-proof and efficient mechanisms (Holmström, 1979), and an immediate generalization of this result implies that the class of generalized Groves mechanisms fully characterizes the class of strategy-proof and *utilitarian* mechanisms.

Definition 4. A mechanism $f = (\alpha, \tau)$ is a generalized Groves mechanism if the decision rule is utilitarian and the transfer rule satisfies, for all i and θ ,

$$\tau_i(\theta) = g_i(\theta_{-i}) - \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta)) \Big),$$

for any function $g_i: \Theta_{-i} \to \mathbb{R}$.

Proposition 1 (Holmström, 1979). Given any convex environment $\mathcal{E}_X \in \mathbb{E}_X$, a mechanism is strategy-proof and utilitarian if and only if it is a generalized Groves mechanism.

Proof. The proof follows immediately from the proof in Holmström (1979) of the result that, given any convex environment $\mathcal{E}_X \in \mathbb{E}_X$, a mechanism is strategy-proof and efficient if and only if it is a Groves mechanism.

²¹A mechanism is a *Groves mechanism*, due to Groves (1973), if the decision rule is efficient and the transfer rule satisfies, for all *i* and θ , $\tau_i(\theta) = g_i(\theta_{-i}) - (\sum_{j \neq i} v_j(\alpha(\theta), \theta_j)) - \sum_{k \in I} c_k(\alpha(\theta))$ for some $g_i : \Theta_{-i} \to \mathbb{R}$.

Recall that multiplying by λ_i translates monetary units of value for *i* into interpersonally comparable welfare units and dividing by λ_i does the reverse. A generalized Groves mechanism subsidizes each individual *i* her monetary equivalent of the social welfare value of the public good, excluding her own welfare benefits, minus a term that is constant in her report.²² Since *i*'s total welfare consists of her welfare from the decision minus her welfare cost from her payment, her total welfare is precisely equal to social welfare minus a constant that does not depend on her report. That is,

$$\lambda_i(v_i(\alpha(\theta), \theta_i) - \tau_i(\theta)) = \lambda_i v_i(\alpha(\theta), \theta_i) + \sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta)) - \lambda_i g_i(\theta_{-i}),$$

and so *i*'s and society's incentives are perfectly aligned.

The most well-known mechanism within the Groves class—the pivotal mechanism chooses $g_i(\theta_{-i})$ to equal the social monetary value of the efficient decision made when ignoring her preferences, excluding her own monetary benefits from the public good.²³ This implies that an individual's transfer is always non-negative. Like with the Groves mechanisms, we can generalize the pivotal mechanism to employ a utilitarian, rather than merely efficient, decision rule as follows.

Definition 5. A mechanism $f = (\alpha, \tau)$ is a generalized pivotal mechanism if the decision rule is utilitarian and the transfer rule satisfies, for all i and θ ,

$$\tau_i(\theta) = \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i})) \Big) \\ - \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta)) \Big).$$

A generalized pivotal mechanism charges each individual i her monetary equivalent of the welfare loss she imposes on society, excluding her own welfare benefits from the public good. In particular, a generalized pivotal mechanism charges i her monetary equivalent of the difference in social welfare between the utilitarian decision made

²²Note that *i*'s (generalized) Groves transfer can be equivalently defined as a (generalized) pivotal transfer plus a term that is constant in her report, as I do in Section 1. Relative to how it is defined here, this simply bakes an additional constant into the constant term.

²³A mechanism is a *pivotal mechanism*, also known as a VCG mechanism, if the decision rule is efficient and the transfer rule satisfies, for all *i* and θ , $\tau_i(\theta) = \sum_{j \neq i} v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} c_k(\alpha(0, \theta_{-i})) - (\sum_{j \neq i} v_j(\alpha(\theta), \theta_j)) - \sum_{k \in I} c_k(\alpha(\theta))$. Note that in the literature, the pivotal mechanism is sometimes defined differently in different contexts. Indeed, it is often taken to be equivalent to the Clarke mechanism (see, e.g., Mas-Colell, Whinston and Green (1995, p. 878)), which is *not* equivalent to the pivotal mechanism as defined here. I disambiguate these two mechanisms in Appendix B.

when ignoring her preferences and the utilitarian decision made when considering her preferences, excluding her own welfare benefits from the public good.

The pivotal mechanism earns its name from the fact that it only charges individuals a non-zero amount if they are pivotal, i.e., if the decision changes when taking their preferences into account.

Definition 6. Given a mechanism $f = (\alpha, \tau)$ and type profile $\theta \in \Theta$, an individual *i* is *pivotal* if $\alpha(0, \theta_{-i}) \neq \alpha(\theta)$.

As we will see in Section 7.3, the probability that *any* individual is pivotal goes to zero as the population gets large (Proposition 4). Hence, the generalized pivotal mechanism raises no revenue in large populations. We would like to raise exactly the amount of revenue necessary to implement the chosen alternative. Recall that the only flexibility—if we are to design a strategy-proof and utilitarian mechanism—is in choosing $g_i(\theta_{-i})$, which does not depend on *i*'s report. This sets the stage for a new mechanism, which I call the cost-sharing pivotal mechanism.

The cost-sharing pivotal mechanism adds an additional term to *i*'s transfer in the generalized pivotal mechanism. It adds $c_i(\alpha(0, \theta_{-i}))$ —*i*'s fair share of the cost of what would have been produced without her. As a result, when no one is pivotal, the cost-sharing pivotal mechanism charges each individual precisely her fair share of the cost of what is produced and is ex-post budget-balanced in large populations (Theorem 4).

In fact, the cost-sharing pivotal mechanism is the unique solution to the public good provision problem as laid out in Section 1. That is, the cost-sharing pivotal mechanism is the unique mechanism—up to vanishingly small perturbations—which satisfies strategy-proofness, utilitarianism, cost-sharing universal participation, the fair pricing principle, and asymptotic ex-post budget-balance (Theorem 5).

Definition 7. A mechanism $f = (\alpha, \tau)$ is a cost-sharing pivotal mechanism (CSP) if the decision rule is utilitarian and the transfer rule satisfies, for all i and θ ,

$$\tau_i(\theta) = \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i})) \Big) \\ - \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta)) \Big) + c_i(\alpha(0, \theta_{-i})).$$

The cost-sharing pivotal mechanism employs a utilitarian decision rule and charges each individual her monetary equivalent of the welfare loss she imposes on society, excluding her own welfare benefits from the public good, plus her fair share of the cost of what would have been produced without her. Figure 2 depicts the cost-sharing pivotal mechanism graphically for the case of a continuous public good y. $\text{MSB} = \frac{1}{\lambda_i} \sum_k \lambda_k v'_k(y, \theta_k)$ is the monetary equivalent to *i* of the marginal social benefit of the public good. $\text{MSB}_{-i} = \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j v'_j(y, \theta_j)$ is the monetary equivalent to *i* of the marginal social benefit of the public good excluding *i*. $\text{MSC} = \frac{1}{\lambda_i} \sum_{k \in I} \lambda_k c'_k(y)$ is the marginal social cost of the public good when everyone pays their fair share of its cost. c'_i is *i*'s marginal fair cost share. $\alpha(\theta)$ is the utilitarian decision. $\alpha(0, \theta_{-i})$ is the utilitarian decision without *i*.



Figure 2

The cost-sharing pivotal mechanism implements $\alpha(\theta)$ and charges *i* the welfare loss she imposes on society excluding her own welfare benefits from the public good (the area in red), plus her fair share of the cost of what would have been produced without her (the area in green). The generalized pivotal mechanism, by comparison, charges *i* only the former (the area in red).

5 Participation

The second desiderata in the government's public good provision problem is participation, i.e., that the mechanism provides incentives for each individual to participate in the process. In this section, I introduce a new participation constraint called *costsharing universal participation*, which states that each individual should always prefer to participate in the mechanism rather than to not participate, receive the alternative that would have been chosen without her, and be taxed her fair share of its cost. I show that the cost-sharing pivotal mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and cost-sharing universal participation (Theorem 2). Participation constraints require that each individual prefers to participate in the mechanism rather than to not participate and receive some outside option. Hence, what distinguishes these constraints from each other is the outside option. The most commonly used participation constraint in mechanism design is individual rationality, where the outside option is some fixed outcome with a payoff of zero—i.e., to consume nothing and to pay nothing.

Definition 8. A mechanism $f = (\alpha, \tau)$ is *individually-rational* if for all i, θ ,

$$v_i(\alpha(\theta), \theta_i) - \tau_i(\theta) \ge 0.$$

Individual rationality says that an individual should prefer to participate in the mechanism rather than to consume nothing and pay nothing. This is fitting, for example, in an auction setting, where the outcome from not participating is, in fact, to consume and pay nothing. However, it is not fitting in a public goods setting where—since public goods are by definition non-rival and non-excludable—each individual consumes what the mechanism implements no matter if they participate or not.

Another participation constraint used in the literature is universal participation, which is designed to capture the non-rivalry and non-excludability of public goods.²⁴ Under universal participation, the outside option is precisely that the individual consumes whatever the mechanism would have implemented without her.

Definition 9. A mechanism $f = (\alpha, \tau)$ satisfies universal participation if for all θ and i,

$$v_i(\alpha(\theta), \theta_i) - \tau_i(\theta) \ge v_i(\alpha(0, \theta_{-i}), \theta_i).$$

Universal participation says that an individual should prefer to participate in the mechanism rather than to consume what would have been produced without her and pay nothing. This assumes that the implementer has no ability to tax individuals who do not participate in the mechanism. While this is true in some public good settings—for instance, a group of friends deciding which couch to buy for their apartment—this is not true for a government's public good provision problem. Indeed, in practice many public goods are funded by tax dollars which violate universal participation (e.g., my tax dollars may go to fund a park for which I have zero value).

I propose a new participation constraint, which I call cost-sharing universal participation, in which the outside option is that the individual consumes whatever the mechanism would have implemented without her *and* is taxed her fair share of its cost.

 $^{^{24}}$ See Green and Laffont (1979, Chapter 6) for a similar discussion of individual rationality and universal participation. Universal participation is sometimes called *no-free-ride* (see, e.g., Moulin (1986)).

Definition 10. A mechanism $f = (\alpha, \tau)$ satisfies cost-sharing universal participation (CS-UP) if for all θ and i,

$$v_i(\alpha(\theta), \theta_i) - \tau_i(\theta) \ge v_i(\alpha(0, \theta_{-i}), \theta_i) - c_i(\alpha(0, \theta_{-i})).$$

Cost-sharing universal participation says that an individual should prefer to participate in the mechanism rather than to consume what would have been produced without her and be taxed her fair share of its cost. This captures the idea that, by not participating, an individual avoids neither her consumption of the good that is eventually produced nor the taxes the government levies to fund it. What she does avoid is any alterations to her tax payment the government may impose as a function of her report to the mechanism.

Indeed, one useful way to implement a public good provision mechanism in practice is to collect each individual's fair cost share through the existing tax system, rather than to charge each individual an explicit payment through the mechanism itself. I discuss this further in Section 11.

The cost-sharing pivotal mechanism satisfies cost-sharing universal participation. In fact, the cost-sharing pivotal mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and cost-sharing universal participation. In a similar fashion, the generalized pivotal mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and universal participation.

Theorem 2. Given any convex environment $\mathcal{E}_X \in \mathbb{E}_X$,

- 1. a mechanism maximizes ex-post revenue among all mechanisms which satisfy strategy-proofness, utilitarianism, and universal participation if and only if it is a generalized pivotal mechanism; and
- 2. a mechanism maximizes ex-post revenue among all mechanisms which satisfy strategy-proofness, utilitarianism, and cost-sharing universal participation if and only if it is a cost-sharing pivotal mechanism.

Proof. Part I. Given any convex environment $\mathcal{E}_X \in \mathbb{E}_X$, a mechanism f is strategy-proof and utilitarian if and only if it is a generalized Groves mechanism by Proposition 1. Let $f = (\alpha, \tau)$ be a generalized Groves mechanism with $\tau_i(\theta) = g_i(\theta_{-i}) - \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta))]$ for some $g_i : \Theta_{-i} \to \mathbb{R}$. We would like to construct $g_i(\theta_{-i})$ to maximize revenue subject to universal participation. In particular,

$$g_i(\theta_{-i}) = \inf_{\theta_i \in \Theta_i} v_i(\alpha(\theta), \theta_i) + \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta)) \Big) - v_i(\alpha(0, \theta_{-i}), \theta_i).$$

Plugging $\theta_i = 0$ into the objective function, we have $\frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i}))]$, which is indeed the minimum since by definition of α , for all θ_i ,

$$\lambda_{i}v_{i}(\alpha(\theta),\theta_{i}) + \sum_{j\neq i}\lambda_{j}v_{j}(\alpha(\theta),\theta_{j}) - \sum_{k\in I}\lambda_{k}c_{k}(\alpha(\theta))$$

$$\geq \lambda_{i}v_{i}(\alpha(0,\theta_{-i}),\theta_{i}) + \sum_{j\neq i}\lambda_{j}v_{j}(\alpha(0,\theta_{-i}),\theta_{j}) - \sum_{k\in I}\lambda_{k}c_{k}(\alpha(0,\theta_{-i})),\theta_{j}$$

which holds if and only if for all θ_i ,

$$v_{i}(\alpha(\theta), \theta_{i}) + \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\alpha(\theta), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\alpha(\theta)) \Big) - v_{i}(\alpha(0, \theta_{-i}), \theta_{i})$$

$$\geq \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\alpha(0, \theta_{-i}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\alpha(0, \theta_{-i})) \Big).$$

Hence, $g_i(\theta_{-i}) = \frac{1}{\lambda_i} \left[\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta)) \right]$ as desired. \Box **Part II.** The proof follows analogously as in Part I. \Box

As we will see in the upcoming sections, in addition to satisfying cost-sharing universal participation, the cost-sharing pivotal mechanism is the unique mechanism (up to vanishingly small perturbations) that satisfies strategy-proofness, utilitarianism, the fair pricing principle, and asymptotic ex-post budget-balance alone. Given this, another way to think about the content of this section is the following. We use the cost-sharing pivotal mechanism because it is the unique mechanism which satisfies the previous four desiderata, and, in order to give individuals dominant-strategy incentives to participate, we *design* what the mechanism will do if an individual does not participate.

Suppose first that the government can implement any alternative and tax any individual any amount regardless of if they participate. Then it could make sense to tax those who do not participate the transfer they would have been charged had they reported the average value among those who did participate (optionally, with similar observable characteristics) and to implement this alternative. Since the mechanism is strategy-proof, this gives them dominant-strategy incentives to participate, and in the event that they don't participate due to unforeseen circumstances or just by random chance, this estimates their value based on the data from the individuals who did participate and implements the utilitarian alternative given this estimate. Of course, if individuals are not missing at random (e.g., because those with lower monetary values, conditional on observables, do not participate at a higher rate), this estimate will be biased.

Suppose next that the government can only tax individuals who do not participate their fair share of the cost of what is implemented without them. This could be, for example, because individuals are taxed their fair cost share through the tax system itself (see Section 11). Then the question becomes: what decision rule $\alpha_{-i} : \Theta_{-i} \to Y$ should the government use if an individual *i* does not participate? We know that the cost-sharing pivotal mechanism satisfies cost-sharing universal participation (Theorem 2), so $\alpha_{-i}(\theta_{-i}) = \alpha(0, \theta_{-i})$ will always give individuals a dominant-strategy to participate. But there are other α_{-i} that work too.²⁵ That said, most don't, and in particular the decision rule which takes *i* to have the average value among those who did participate, as mentioned in the previous paragraph, does not.

6 Fairness

The fifth desiderata in the government's public good provision problem is fairness, i.e., that the mechanism does not ask anyone to pay what has been deemed an unfair amount. In this section, I introduce a new fairness principle called *the fair pricing principle*, which states that each individual should never pay more than her monetary value for what is produced or her fair share of its cost, whichever is larger. I show that the cost-sharing pivotal mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and the fair pricing principle (Theorem 3).

Suppose that the space of social alternatives is unidimensional and continuous. That is, let $Y = \mathbb{R}_+$ represent the quantity and/or quality of the public good to be produced. Moreover, suppose that for each i and θ_i , v_i is absolutely continuous and non-decreasing in y with $v_i(0, \theta_i)$ normalized to 0 for all θ_i . Likewise, suppose that costs c and fair shares c_i for each i are absolutely continuous and non-decreasing with c(0) normalized to $0.^{26}$ Let $\mathbb{E}_C \subset \mathbb{E}_X$ be the set of such continuous public good environments.

In addition to giving individuals incentives to participate, a participation constraint can also be interpreted as requiring that a mechanism satisfy some notion of fairness. The most commonly used participation constraint in mechanism design is individual rationality. Interpreted as a fairness constraint, individual rationality can be rewritten as follows.

²⁶Hence, $v_i(y, \theta_i) = \int_0^y v'_i(z, \theta_i) dz$ for all $i, c(y) = \int_0^y c'(z) dz$, and $c_i(y) = \int_0^y c'_i(z) dz$ for all i.

²⁵For a given $\alpha_{-i}(\theta_{-i})$, suppose there exists $\theta_i^* \in \Theta_i$ such that $\alpha(\theta_i^*, \theta_{-i}) = \alpha_{-i}(\theta_{-i})$. Then if *i* is of type θ_i^* , the payoff of reporting θ_i^* must be no less than the payoff of not participating. Hence, one α_{-i} that gives dominant-strategy incentives to participate is one which produces the same or more of the public good than would have been produced without *i*, $\alpha_{-i}(\theta_{-i}) = \alpha(0, \theta_{-i}) + \delta$ for $\delta \geq 0$, as long as *i*'s pivotal payment from reporting such a θ_i^* is no larger than the increase in her fair cost share, $c_i(\alpha_{-i}(\theta_{-i})) - c_i(\alpha(0, \theta_{-i}))$. Clearly, $\delta = 0$ always works, but for some θ_{-i} , $\delta > 0$ also works.

Definition 11. A mechanism $f = (\alpha, \tau)$ is *individually-rational* if for all i, θ ,

$$\tau_i(\theta) \le v_i(\alpha(\theta), \theta_i) = \int_0^{\alpha(\theta)} v'_i(y, \theta_i) \, \mathrm{d}y.$$

Individual rationality states that it is fair for an individual to pay up to her total monetary value for what is produced and no more. This is arguably too demanding for public good environments. Such environments are commonly characterized by a sense of community, wherein each individual understands that everyone needs to chip in—but not unreasonably so—for the greater good, even if that means paying more than their monetary value for the public good. Although I may not value the new park or library in the neighborhood, I understand that if my community decides it is worthwhile to produce, I will be asked—and I ought—to chip in my fair share.

I propose the following principle, which I call the fair pricing principle, to capture this sentiment. The fair pricing principle states that it is fair for an individual to pay up to her total monetary value for what is produced or her fair share of its cost, whichever is larger.

Definition 12. A mechanism $f = (\alpha, \tau)$ satisfies the fair pricing principle (FPP) if for all i, θ ,

$$\tau_i(\theta) \le \max\left\{v_i(\alpha(\theta), \theta_i), c_i(\alpha(\theta))\right\} = \max\left\{\int_0^{\alpha(\theta)} v_i'(y, \theta_i) \, \mathrm{d}y, \int_0^{\alpha(\theta)} c_i'(y) \, \mathrm{d}y\right\}.$$

In other words, the fair pricing principle states that it is fair to ask an individual to pay up to her fair share of the cost of what is produced—and it is also fair to ask an individual to pay more, as long as this amount is less than her total monetary value for what is produced.

I now define two progressively weaker notions of fairness. The first I call the fair pricing principle per unit. The second I call no-extortion. The fair pricing principle per unit captures the same sentiment as the fair pricing principle, but unit by unit.

Definition 13. A mechanism $f = (\alpha, \tau)$ satisfies the fair pricing principle per unit *(FPP per unit)* if for all i, θ ,

$$\tau_i(\theta) \le \int_0^{\alpha(\theta)} \max\left\{ v'_i(y,\theta_i), \, c'_i(y) \right\} \, \mathrm{d}y.$$

The fair pricing principle per unit says that for each unit of the public good, it is fair for an individual to pay up to her marginal value for that unit or her fair share of its marginal cost, whichever is larger. Since the maximum is taken unit by unit, this is a strictly weaker condition than the fair pricing principle, as this allows, for instance, that an individual be taxed her fair share for the first marginal unit of the good and taxed her marginal value for the second—which may total more than the maximum of her fair share and her value for both units of the good (see Example 3).

The second I call no-extortion. No-extortion is a minimal standard of fairness in public good environments. It applies the fair pricing principle only to the case when no public good is produced. It says that if nothing is produced, no one should pay.²⁷

Definition 14. A mechanism $f = (\alpha, \tau)$ satisfies *no-extortion (NE)* if when nothing is produced, no one pays. That is, $\alpha(\theta) = 0$ implies $\tau_i(\theta) \leq 0$ for all *i*.

It is clear that individual rationality implies the fair pricing principle which implies the fair pricing principle per unit which implies no-extortion. At this point, one might be wondering if cost-sharing universal participation also implies no-extortion. The answer is yes, but with a minor qualification. This qualification turns out to be important for fairness in general. I discuss it now.

In a continuous public good environment, it is intuitive that if *i*'s marginal value increases everywhere, all else fixed, a utilitarian decision rule would never select a strictly smaller amount of the public good. This intuition is nearly correct, but for a technical reason not entirely so. It is true if the set of utilitarian decisions is always a singleton, but it is not true in general. What is true is that increasing an individual's marginal value increases the set of utilitarian decisions in the strong set order.²⁸

Proposition 2. Given any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$, if $v'_i(y, \hat{\theta}_i) \ge v'_i(y, \theta_i)$ for all i, y, and θ_i , then $A^*(\hat{\theta}) \ge A^*(\theta)$ in the strong set order.

Proof. For any θ , $y^* \in A^*(\theta)$ if and only if for any $y' \leq y^* \leq y''$,

$$\int_{y'}^{y^*} \sum_{i \in I} \lambda_i v_i'(y, \theta_i) - \lambda_i c_i'(y) \, \mathrm{d}y \ge 0 \ge \int_{y^*}^{y''} \sum_{i \in I} \lambda_i v_i'(y, \theta_i) - \lambda_i c_i'(y) \, \mathrm{d}y$$

²⁷One might think that it is perfectly reasonable to charge a payment to an individual who sways the decision towards something they prefer. With negative monetary values, it is possible to sway the decision from producing the good to not. One could then define no-extortion with the prerequisite that *i* weakly prefers having any amount of the good to not. That is, a mechanism satisfies *noextortion* if, for any *i* and θ_i such that $v_i(y, \theta_i) \ge v_i(0, \theta_i)$ for all $y, \alpha(\theta) = 0$ implies $\tau_i(\theta) \le 0$. In this and all subsequent sections, I consider the case of non-negative (indeed, non-decreasing) values. Hence, for simplicity I define no-extortion without this prerequisite. It is immediate that all the results in this paper hold identically for either definition.

²⁸For any $A, B \subseteq \mathbb{R}$, $A \leq B$ in the strong set order if for any $a \in A$ and $b \in B$, min $\{a, b\} \in A$ and max $\{a, b\} \in B$. See Milgrom and Shannon (1994).

That is, y^* is utilitarian if and only if moving from any $y' \leq y^*$ to y^* weakly increases total welfare and moving from y^* to any $y'' \geq y^*$ weakly decreases total welfare.

Suppose $v'_i(y, \hat{\theta}_i) \ge v'_i(y, \theta_i)$ for all i, y, and $\theta_i, y^* \in A^*(\theta), \hat{y}^* \in A^*(\hat{\theta})$, and $\hat{y}^* \le y^*$. Then

$$\begin{split} \int_{\hat{y}^*}^{y^*} \sum_{i \in I} \lambda_i v_i'(y, \theta_i) - \lambda_i c_i'(y) \, \mathrm{d}y &\geq 0 \\ &\geq \int_{\hat{y}^*}^{y^*} \sum_{i \in I} \lambda_i v_i'(y, \hat{\theta}_i) - \lambda_i c_i'(y) \, \mathrm{d}y \geq \int_{\hat{y}^*}^{y^*} \sum_{i \in I} \lambda_i v_i'(y, \theta_i) - \lambda_i c_i'(y) \, \mathrm{d}y \end{split}$$

where the first inequality follows since $y^* \in A^*(\theta)$, the second since $\hat{y}^* \in A^*(\hat{\theta})$, and the third since $v'_i(y, \hat{\theta}_i) \ge v'_i(y, \theta_i)$ for all i, y, and θ_i . Hence, $y^* \in A^*(\hat{\theta})$ and $\hat{y}^* \in A^*(\theta)$.

Because a utilitarian decision rule can select *any* utilitarian decision, nothing prevents it from selecting a strictly smaller amount of the public good when marginal values increase, provided this remains in the set of utilitarian decisions. A trivial example is if an increase in *i*'s marginal value doesn't change the set of utilitarian decisions. In such a case, a utilitarian decision rule may very well select a strictly smaller amount. It turns out that such decision rules will cause problems for fairness, in the sense of violating no-extortion or any of the stronger notions above (Example 2). This motivates the following definition, which aligns utilitarian decision rules with our intuition.

Definition 15. A decision rule $\alpha : \Theta \to \mathbb{R}_+$ is monotone if $v'_i(y, \hat{\theta}_i) \ge v'_i(y, \theta_i)$ for all y and i implies $\alpha(\hat{\theta}) \ge \alpha(\theta)$.

A decision rule is monotone if increasing any individual's value function pointwise never decreases the decision. With a *non*-monotone decision rule, the cost-sharing pivotal mechanism violates no-extortion.

Example 2 (With a non-monotone decision rule, the cost-sharing pivotal mechanism violates no-extortion). Let $I = \{i, j\}$ and $Y = \{0, 1, 2\}$, representing no park, a small park, and a large park, respectively. Let $\lambda_i = \lambda_j = 1$, so that the utilitarian and efficient decision rule coincide. The cost of the small park is 10 and the cost of the large park is 20, and each individual's fair share is 5 for the small park and 10 for the large park. Individual *i* values the small park at 0 and the large park at 2, and individual *j* values the small park at 10 and the large park at 15. Given this, the set of utilitarian decisions is $\{0, 1\}$. Both no park and the small park result in a total welfare of zero, which is optimal. Suppose α selects 0—no park.

Ignoring *i*'s preferences, the set of utilitarian decisions remains unchanged. Suppose α is non-monotonic and for this type profile selects 1—the small park. The costsharing pivotal mechanism charges *i* the welfare loss imposed on society, excluding her own benefits from the public good, when taking her preferences into account (0, since social welfare, minus i's, is zero both with no park and the small park) plus her fair share of the cost of what would have been produced without her (5, her fair cost share of the small park), violating no-extortion. \Box

With a monotone decision rule, cost-sharing universal participation indeed implies no-extortion.

Fact 1. Given any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$, if the decision rule is monotone, cost-sharing universal participation implies no-extortion.

Proof. Suppose $\alpha(\theta) = 0$. Since α is monotone, $\alpha(0, \theta_{-i}) = 0$. Then cost-sharing universal participation implies $\tau_i(\theta) \leq 0$ as desired.

With a monotone decision rule, the cost-sharing pivotal mechanism satisfies noextortion and the fair pricing principle per unit, but it violates the fair pricing principle (Example 3). That said, the cost-sharing pivotal mechanism satisfies the fair pricing principle under the natural assumption that marginal values are non-increasing and marginal costs are non-decreasing.

Assumption 1. $v'_i(y, \theta_i)$ is non-increasing in y for all i and θ_i , and c'(y) and $c'_i(y)$ are non-decreasing for all i.

Note that in the special case of a binary public good environment—where a public good can either be produced or not—Assumption 1 always holds.

Example 3 (Without Assumption 1, the cost-sharing pivotal mechanism violates the fair pricing principle). Let $I = \{i, j\}$ and $Y = \{0, 1, 2\}$, representing no park, a small park, and a large park, respectively. Let $\lambda_i = \lambda_j = 1$, so that the utilitarian and efficient decision rule coincide. The cost of the small park is 10 and the cost of the large park is 20, and each individual's fair share is 5 for the small park and 10 for the large park. Individual *i* values the small park at 0 and the large park at 11 (violating Assumption 1), and individual *j* values the small park at 11 and the large park at 11. Total welfare is 0 with no park, 0 + 11 - 10 = 1 with the small park, and 11 + 11 - 20 = 2 with the large park, so the large park is the utilitarian decision.

Ignoring *i*'s preferences, the small park would be produced for a total welfare, minus i's, of 1. With *i*, the large park is produced for a total welfare, minus *i*'s, of 11-20 = -9. The cost-sharing pivotal mechanism charges *i* the welfare loss imposed on society, excluding her own benefits from the public good, when taking her preferences into account (1 - (-9) = 10) plus her fair share of the cost of what would have been produced without her (5), for a total of 15. This violates the fair pricing principle, which requires that *i* pay no more than the maximum of her value and her fair share

of the cost of what is produced (11 and 10, respectively).²⁹ \Box

In Theorem 2, we showed that the cost-sharing pivotal mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and cost-sharing universal participation. Perhaps surprisingly, it is *also* true that the cost-sharing pivotal mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, utilitarianism, and the fair pricing principle.

Theorem 3. Consider any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$ and any utilitarian and monotone decision rule α . A mechanism $f = (\alpha, \tau)$ maximizes ex-post revenue among all mechanisms which satisfy strategy-proofness, utilitarianism, and the fair pricing principle per unit if and only if it is a cost-sharing pivotal mechanism. Under Assumption 1, the same holds replacing the fair pricing principle per unit with the fair pricing principle.

Comments on Proof. The proof proceeds similarly as in Theorem 2, albeit with several more cases.

7 Budget-Balance

The fourth desiderata in the government's public good provision problem is budgetbalance, i.e., that the mechanism raises exactly the amount of revenue necessary to produce the desired quantity of the public good. In this section, I show that the cost-sharing pivotal mechanism is *asymptotically ex-post budget balanced*. That is, the probability of ex-post budget-balance goes to one and the expected distance from ex-post budget-balance per capita goes to zero as the population size goes to infinity (Theorem 4). Moreover, I show that no other mechanism satisfying the desired criteria consistently comes closer to budget-balance than the cost-sharing pivotal mechanism (Proposition 5).

The rest of the section is organized as follows. First, I discuss why asymptotic ex-post budget-balance is an appropriate desiderata for the government's public good provision problem. Second, I discuss two important features of the asymptotic analysis. Third, I present the formal model and results.

7.1 The Approximate Budget-Balance Principle

It is common in the literature on public good provision to restrict attention to mechanisms which never run a budget deficit. This is often simply termed *feasibility*. I want

²⁹On the other hand, the fair pricing principle per unit requires that i pay no more than the maximum of her marginal value and her marginal fair cost share for each marginal unit of what is produced—i.e., she can pay up to 5 for the first unit and up to 11 for the second (up to 16 total).

to argue that in public good provision environments with large populations, this is an inappropriate notion of feasibility. In most such environments, surpluses and deficits ultimately trickle down to the citizens. What matters is that their incentives are preserved in light of this fact. This inspires what I call the *approximate budget-balance principle*.

I propose that governments can finance small deficits and distribute small surpluses through future changes to tax policy that are sufficiently inconsequential so as not to affect the incentives of their citizens. That is, violations of budget-balance in either direction are feasible, as long as they are *small*. One way to understand this is that governments maintain a fund capable of absorbing small deficits and surpluses. Over time, any accumulated surplus is repaid through tax cuts, while any accumulated deficit is replenished through tax increases. If these adjustments are small and indirect,³⁰ it is reasonable to assume that individuals do not perceive them as part of the mechanism.

Approximate Budget-Balance Principle. Sufficiently small and indirect changes to tax policy are not perceived by individuals as part of their transfer and hence do not affect their incentives. Approximate budgetbalance is thus as good in practice as exact budget-balance.

Under the approximate budget-balance principle, if the expected distance from expost budget-balance per capita is sufficiently small, the mechanism is feasible. It is not particularly important that a mechanism never runs a budget deficit. Arguably, this is even somewhat of a red herring. In fact, never running a budget deficit and fairness turn out to be fundamentally in conflict: no strategy-proof and utilitarian mechanism can satisfy both no-deficit and no-extortion (Proposition 6).

7.2 Discussion of Analysis

To analyze the asymptotic behavior of the cost-sharing pivotal mechanism as the population gets large, I consider sequences of public good environments with increasing population size n, where the private information and observable characteristics of each individual are drawn from some joint distribution which is unknown to the government. Two important features of this analysis are 1) that the cost of the public goods can vary arbitrarily with n and 2) that welfare weights and cost shares can depend arbitrarily on observable characteristics, and observable characteristics are correlated arbitrarily with types.³¹ I discuss each below.

³⁰Changes can be indirect in several ways. For example, they can be indirect in context (changes to tax policies not related to public good provision mechanisms) and indirect in time (changes occur sufficiently far into the future).

 $^{^{31}}$ Moreover, I discuss in Section 9.1 how a common methodological simplification in mechanism design, the net-value approach, cannot accommodate either.

The cost of the public good can vary arbitrarily with population size. It is common that asymptotic results involving sequences of public good environments depend on how the sequence of cost functions grows with n.³² Importantly, the results here hold no matter how costs grow with n. To see the relevance of this distinction, consider for simplicity a binary public good environment in which a public good can either be produced or not. We do not know the population distribution of values. We would like to talk about what happens in large populations for any such distribution. That is, we would like to learn about what happens as we increase the number of i.i.d. draws from any distribution holding all else constant. But what does it mean to hold all else constant? In particular, what about the relationship between the population size and the cost of the public good, if anything, should remain constant as we increase n?

If we assume costs grow very slowly or not at all, then as the population grows large, the utilitarian decision is almost surely to produce the good, making the public good provision problem trivial, no matter the underlying population distribution. If it is optimal to produce the good with arbitrarily high probability, we do not need to elicit individuals' preferences—we can simply produce the good and tax everyone their fair share. Similarly, if we assume costs grow sufficiently quickly, then as the population grows large, the utilitarian decision is almost surely *not* to produce the good, again trivializing the problem—we can simply not produce the good. To avoid this situation in which the chosen sequence of cost functions plays a significant (and artificial) role in the analysis, we would like our results to be robust to *any* such sequence.³³

Welfare weights and cost shares can vary arbitrarily with observables. So far, the fact that welfare weights and cost shares can depend arbitrarily on observable characteristics has not played a role in the analysis. But this will no longer be the case, as we will need to make statements about deviations from budget-balance, and this depends on individuals' welfare weights and cost shares—and, in particular, on how they correlate with individuals' types.

Notably, the commonly assumed special case of equal cost shares, $c_i(y) = c(y)/n$ for all *i*, assumes away an important dimension of fair public good provision—that fair cost shares may depend on observables, and observables may be correlated with types. To gain some intuition, consider a cost-sharing pivotal mechanism with equal cost shares. It might be that the expected distance from ex-post budget-balance is small because almost everyone is not pivotal and so pays their equal share, the sum of which approaches the full cost as the population gets large. However, now suppose an individual's fair share is determined by her income, which is highly correlated with

 $^{^{32}}$ See, e.g., Mailath and Postlewaite (1990) and Xi and Xie (2023).

³³Arguably, if anything should be held constant, it is precisely the probability that the utilitarian decision is to produce the good. That is, we would like $\mathbb{P}(\alpha^n(\theta^n) = y)$ to be constant in *n* for each *y*. Computing sequences of cost functions with this property is not easy. Luckily, we do not have to compute such a sequence, as the result is robust across all sequences.

her value for the public good. In particular, suppose that those who tend to be pivotal (due to high values for the good) tend to have much larger fair shares, leaving much smaller fair shares for everyone else. Now, even if almost everyone is not pivotal, they will pay a much smaller sum in total, and this amount may not approach the full cost. In particular, the cost-sharing pivotal mechanism might be asymptotically ex-post budget-balanced in the former case but not the latter.

Encouragingly, the results herein are robust to each of the aforementioned concerns. Theorem 4 shows that the cost-sharing pivotal mechanism is asymptotically ex-post budget-balanced no matter how the cost of the public goods vary with n, no matter how welfare weights and fair cost shares are constructed from observables, and no matter how types and observables are correlated.

7.3 Formal Analysis

Suppose that the space of social alternatives is unidimensional and finite. That is, let $Y = \{0, 1, \ldots, \bar{y}\}$ for some $\bar{y} \in \mathbb{N}$ represent the quantity and/or quality of the public good to be produced. For each i and θ_i , v_i is non-decreasing in y with $v_i(0, \theta_i)$ normalized to 0 for all θ_i . For notational convenience, let $\Theta_i \subseteq \mathbb{R}^{\bar{y}+1}_+$ and $v_i(y, \theta_i) =$ $\theta_i(y)$ for all i and y, where $\theta_i(y)$ is the yth component of θ_i . Costs c and fair shares c_i for each i are non-decreasing with c(0) normalized to 0. Let $\mathbb{E}_F \subset \mathbb{E}_C$ be the set of such finite public good environments.³⁴

A sequence of public good provision environments with transfers is defined by

$$\mathcal{E}^{n} = (I^{n}, Y^{n}, \Theta^{n}, Z^{n}, \{v_{i}^{n}\}_{i \in I}, c^{n}, \{c_{i}^{n}\}_{i \in I}, \{\lambda_{i}^{n}\}_{i \in I}).$$

For each $n \in \mathbb{N}$, let there be n individuals with the same type space $\Theta_0 \subseteq \mathbb{R}^{\bar{y}}_+$ and the same arbitrary set of observable characteristics Z_0 . Let the set of alternatives be fixed at $Y_0 = \{0, \ldots, \bar{y}\}$ for some $\bar{y} \in \mathbb{N}$. Let the sequence of cost functions c^n be arbitrary (i.e., the cost function can change arbitrarily with n). Let the sequence of cost share functions c_i^n , for each $i = 1, \ldots, n$, be arbitrary (i.e., how the cost shares are computed from observable characteristics can change arbitrarily with n). Finally, let the welfare weight function $\lambda_i^n : Z_i \to \mathbb{R}_+$ be the same for each individual i and depend only on i's observable characteristics z_i . Let $\mathbb{E}_F^{\mathbb{N}}$ denote the set of all such sequences of finite public good environments.³⁵

³⁴To see that $\mathbb{E}_F \subset \mathbb{E}_C$, note that every finite environment can be equivalently represented by a continuous environment in which we restrict attention to decision rules α with range $Y = \{0, 1, \ldots, \bar{y}\}$ and v_i, c , and c_i for all i are piecewise linear on $\{[0, 1), [1, 2), \ldots, [\bar{y} - 1, \bar{y}), [\bar{y}, \infty)\}$ and constant on $[\bar{y}, \infty)$.

³⁵Formally, for each $n \in \mathbb{N}$, let $I^n = \{1, \ldots, n\}$, $Y^n = Y_0$, $\Theta^n = (\Theta_0)^n$, $Z^n = (Z_0)^n$, $v_i^n(y, \theta_i^n) = \theta_i^n(y)$, $c^n : Y_0 \to \mathbb{R}_+$ be any non-decreasing function, $\{c_i^n\}_{i \in I}$ be any set of functions $c_i^n : Y_0 \times Z^n \to \mathbb{R}_+$ such that each $c_i^n(y, z^n)$ is non-decreasing in y and $\sum_{i=1}^n c_i^n(y, z^n) = c^n(y)$, and $\{\lambda_i^n\}_{i \in I}$ be any set of identical functions $\lambda_i^n : Z_i \to \mathbb{R}_+$.

In order make probabilistic statements about mechanisms in these environments, we will define a population distribution over types and observable characteristics jointly from which individuals are drawn. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and (θ_i, z_i) be a sequence of i.i.d. random vectors defined on it, where $\theta_i : \Omega \to \Theta_0$ is individual *i*'s type and $z_i : \Omega \to Z_0$ is individual *i*'s observable characteristics.

Let $\theta^n = (\theta_i)_{i=1}^n$ denote the sequence of types up to n and $z^n = (z_i)_{i=1}^n$ the sequence of observable characteristics up to n. The realization of z^n determines the welfare weights $\lambda_i^n(z_i^n)$ and the cost shares $c_i^n(\cdot, z^n)$ for each individual i. Given a sequence of mechanisms (α^n, τ^n) , the realization of θ^n and z^n together determine the public good provision level $\alpha^n(\theta^n, z^n)$ and each individual's transfer $\tau_i^n(\theta^n, z^n)$.

We are now ready to define asymptotic ex-post budget-balance.

Definition 16. Consider any sequence of i.i.d. random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}((\theta_i(y_1) - \theta_i(y_0))^2) < \infty$ for all $y_0, y_1 \in Y$, and consider any induced sequence of finite environments $\mathcal{E}_F^n \in \mathbb{E}_F^{\mathbb{N}}$. A sequence of mechanisms (α^n, τ^n) is asymptotically ex-post budget-balanced (AEPBB) if

1. the probability of ex-post budget-balance goes to one as n goes to infinity, i.e.,

$$\mathbb{P}\Big(\sum_{i=1}^{n}\tau_{i}^{n}(\theta^{n}, z^{n}) = c^{n}(\alpha^{n}(\theta^{n}, z^{n}))\Big) \to 1 \quad \text{as} \quad n \to \infty, \quad \text{and}$$

2. the expected distance from ex-post budget-balance per capita goes to zero as n goes to infinity, i.e.,

$$\frac{1}{n}\mathbb{E}\Big(\Big|\sum_{i=1}^n\tau_i^n(\theta^n,z^n)-c^n(\alpha^n(\theta^n,z^n))\Big|\Big)\to 0 \quad \text{as} \quad n\to\infty.$$

A mechanism is asymptotically ex-post budget-balanced if the probability of ex-post budget-balance goes to one and the expected distance from ex-post budget-balance per capita goes to zero as the population size goes to infinity. Assume that each individual's welfare weight is bounded away from zero and infinite.

Assumption 2. For all *i* and *n*, $\lambda_i^n \in [\lambda_L, \lambda_H]$, where $0 < \lambda_L \leq \lambda_H < \infty$.

The following two propositions are the driving force behind Theorem 4, which shows that the cost-sharing pivotal mechanism is asymptotically ex-post budget-balanced. They are also useful results in their own right. The first says that the total CSP payment is bounded below by zero and above by $\frac{\lambda_H}{\lambda_L}c(\alpha(\theta)) + c(\bar{y})$, and hence that the distance from ex-post budget balance in a CSP is bounded above by $\frac{\lambda_H}{\lambda_L}c(\bar{y})$.³⁶

 $[\]overline{\frac{^{36}}{\text{To see this, note that } \frac{\lambda_H}{\lambda_L} c(\alpha(\theta)) + c(\bar{y}) - c(\alpha(\theta))} = \frac{\lambda_H - \lambda_L}{\lambda_L} c(\alpha(\theta)) + c(\bar{y}) \leq \frac{\lambda_H - \lambda_L}{\lambda_L} c(\bar{y}) + c(\bar{y}) = \frac{\lambda_H - \lambda_L}{\lambda_L} c(\bar{y}).$

The second says that the probability that any individual is pivotal goes to zero as n goes to infinity.

Proposition 3. Consider any finite environment $\mathcal{E}_F \in \mathbb{E}_F$ and suppose Assumption 2. For any $\theta \in \Theta$, the total generalized pivotal payment can be no less than zero and no more than $\frac{\lambda_H}{\lambda_L}c(\alpha(\theta))$, and the total CSP payment can be no less than zero and no more than $\frac{\lambda_H}{\lambda_L}c(\alpha(\theta)) + c(\bar{y})$. That is,

- 1. $0 \leq \sum_{i=1}^{n} t_i(\theta) \leq \frac{\lambda_H}{\lambda_L} c(\alpha(\theta))$ and
- 2. $0 \leq \sum_{i=1}^{n} \tau_i(\theta) \leq \frac{\lambda_H}{\lambda_L} c(\alpha(\theta)) + c(\bar{y}),$

where $t_i(\theta)$ is i's transfer in a generalized pivotal mechanism and $\tau_i(\theta) = t_i(\theta) + c_i(\alpha(0, \theta_{-i}))$ is i's transfer in a CSP mechanism. If the decision rule is monotone, the upper bound on the total CSP payment can be lowered to $\frac{\lambda_H + \lambda_L}{\lambda_L} c(\alpha(\theta))$.

Proof Sketch. With only two alternatives $\{0, 1\}$, the maximum revenue in a generalized pivotal mechanism occurs when every individual is exactly pivotal, the entire cost share is borne by the individual with the highest welfare weight (the poorest individual), and the entire value is provided by the individual with the lowest welfare weight (the richest individual). This results in a revenue equal to the cost of the good (over no good) multiplied by $\frac{\lambda_H}{\lambda_L}$.

Now, consider running a sequence of binary generalized pivotal mechanisms on $\{0, \ldots, \bar{y}\}$, where the selected alternative from $\{0, 1\}$ is run against 2, the selected alternative from that mechanism is run against 3, and so on.³⁷ Lemma 1 shows that the transfer in such a sequence of binary generalized pivotal mechanisms is always larger than the transfer in a single generalized pivotal mechanism on $\{0, \ldots, \bar{y}\}$.³⁸

Hence, the total revenue in a generalized pivotal mechanism on $\{0, \ldots, \bar{y}\}$ is less than the revenue from a sequence of binary generalized pivotal mechanisms, each of which has a maximum revenue equal to the difference in cost between the two alternatives considered multiplied by $\frac{\lambda_H}{\lambda_L}$, so the generalized pivotal mechanism raises no more than $\frac{\lambda_H}{\lambda_L}c(\alpha(\theta))$ in revenue. In a CSP, the total fair-share payment $\sum_{i=1}^n c_i(\alpha(0, \theta_{-i}))$ can be no more than $c(\bar{y})$, and with a monotone decision rule, it can be no more than $c(\alpha(\theta))$.

Proposition 4. Consider any sequence of i.i.d. random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}((\theta_i(y_1) - \theta_i(y_0))^2) < \infty$ for all $y_0, y_1 \in Y$, and consider any induced sequence of finite environments $\mathcal{E}_F^n \in \mathbb{E}_F^{\mathbb{N}}$ satisfying Assumption 2. The probability that there is at least one pivotal player in \mathcal{E}_F^n goes to zero as n goes to infinity.

³⁷Viewed as a static mechanism, this mechanism is utilitarian and can be made to have the same decision rule as any generalized pivotal mechanism on $\{0, \ldots, \bar{y}\}$, though it is not strategy-proof.

³⁸Roughly, you have more opportunities to accumulate generalized pivotal payments in a sequence of binary generalized pivotal mechanisms than in a single overall generalized pivotal mechanism.

Comments on Proof. We first show that the probability that any individual is pivotal goes to zero in a binary public good environment. This involves showing that the probability that the maximum value θ_n^* in a sequence $\{\theta_i\}_{i=1}^n$ of non-negative i.i.d. random variables is greater than the distance between $\sum_{i=1}^n \theta_i$ and any sequence of non-negative real numbers c_n goes to zero. From here, we can extend the result to finite public good environments by noting that if an individual is pivotal in a finite public good environment, then there exists a binary public good environment in which they are also pivotal. Since there are finitely many $(\bar{y}(\bar{y}+1)/2)$ pairings of alternatives in a finite public good environment, the probability that any individual is pivotal in any of them goes to zero.

We may now show that the cost-sharing pivotal mechanism is asymptotically ex-post budget-balanced, and hence by the approximate budget-balance principle, that the cost-sharing pivotal mechanism is feasible in large populations.

Theorem 4. Consider any sequence of *i.i.d.* random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}((\theta_i(y_1) - \theta_i(y_0))^2) < \infty$ for all $y_0, y_1 \in Y$, and consider any induced sequence of finite environments $\mathcal{E}_F^n \in \mathbb{E}_F^{\mathbb{N}}$ satisfying Assumption 2. Any sequence of cost-sharing pivotal mechanisms (α^n, τ^n) is asymptotically ex-post budget-balanced.

Proof Sketch. Condition 1 of AEPBB follows by Proposition 4. To show Condition 2, consider a binary public good environment for simplicity (the full proof considers any finite public good environment). The distance from EPBB per capita in a CSP is bounded by $\frac{\lambda_H}{\lambda_L} \cdot \frac{c^n(1)}{n}$ (Proposition 3). First, suppose $\frac{c^n(1)}{n} \not\rightarrow \infty$ as $n \rightarrow \infty$. The CSP is EPBB if no individual is pivotal. The probability that any individual is pivotal goes to zero as n goes to infinity (Proposition 4). Hence, the distance from EPBB per capita is bounded by a constant and the probability of not EPBB goes to zero. Next, suppose $\frac{c^n(1)}{n} \rightarrow \infty$ as $n \rightarrow \infty$. The CSP is EPBB if no good is produced. The probability that the utilitarian decision is to produce the good goes to zero faster than $\frac{c^n(1)}{n}$ goes to infinity by Chebyshev's inequality.

By itself, asymptotic ex-post budget-balance does not imply that there is not another satisfactory mechanism f^* that gets closer to budget-balance than the cost-sharing pivotal mechanism for each n (even if they perform similarly in the limit). Proposition 5 shows that this is, in fact, the case—i.e., that no other mechanism satisfying the desired criteria consistently comes closer to budget-balance than the cost-sharing pivotal mechanism.³⁹

Proposition 5. Consider any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$. Suppose that for any *i* and θ_{-i} , max $A^*(0, \theta_{-i})$ exists and there exists $\theta_i \in \Theta_i$ such that

$$v'_{i}(y,\theta_{i}) = \begin{cases} \nu'_{i}(y) & \text{if } y \leq \max A^{*}(0,\theta_{-i}) \\ 0 & \text{if } y > \max A^{*}(0,\theta_{-i}) \end{cases},$$
(1)

³⁹Note that this result holds for any continuous public good environment, for which any finite public good environment is a special case.

where $\nu'_i(y) \geq \max_{j \neq i} v'_j(y, \theta_j)$ and $\nu'_i(y) > 0$ for all y.⁴⁰ For any utilitarian and monotone decision rule α , there is no strategy-proof and cost-sharing universal participation mechanism that is no farther from ex-post budget-balance than a cost-sharing pivotal mechanism for every θ . The same holds replacing cost-sharing universal participation with the fair pricing principle per unit. Under Assumption 1, the same holds replacing the fair pricing principle per unit with the fair pricing principle.

Proof Sketch. For any θ_{-i} , if θ_i satisfies (1) then every agent's pivotal payment is zero, so the CSP is either exactly budget-balanced or runs a deficit at θ^{41} By Proposition 1, a mechanism f is strategy-proof and utilitarian if and only if i's transfer can be expressed as the sum of her generalized pivotal payment and a term that depends only on her opponents' reports $h_i(\theta_{-i})$. If $h_i(\theta_{-i}) < c_i(\alpha(0, \theta_{-i}))$ the mechanism runs a strictly larger budget-deficit at θ than a CSP, and if $h_i(\theta_{-i}) > c_i(\alpha(0, \theta_{-i}))$ the mechanism violates cost-sharing universal participation by Theorem 2 and the fair pricing principle by Theorem 3.

8 Uniqueness

We have shown that the cost-sharing pivotal mechanism satisfies strategy-proofness, utilitarianism, cost-sharing universal participation (Theorem 2), the fair pricing principle (Theorem 3), and asymptotic ex-post budget-balance (Theorem 4). Because the last property only constrains a mechanism's limiting behavior, there will be many other mechanisms which also satisfy these five criteria. In particular, we can construct mechanisms that deviate from the cost-sharing pivotal mechanism by increasingly small amounts, e.g., by providing rebates that vanish in magnitude or probability with n, that satisfy all five.

For example, consider a mechanism which is equivalent to a cost-sharing pivotal mechanism, except that it provides a \$1 rebate to any individual *i* if all of *i*'s opponents report a particular type $\bar{\theta}_0$. This rebate depends only on the reports of *i*'s opponents, so the mechanism is strategy-proof. This transfer is always less than that of the cost-sharing pivotal mechanism, so it satisfies cost-sharing universal participation and the fair pricing principle. The probability that any *i* receives a rebate goes to zero as *n* goes to infinity, as does the expected distance from ex-post budget-balance per capita, so it is asymptotically ex-post budget-balanced.⁴²

⁴⁰This is a weak richness condition on the domain of values which says that each individual may value marginal units of the public good as much as any other individual and that this marginal value may drop to zero at any point—namely, at max $A^*(0, \theta_{-i})$.

⁴¹If the utilitarian decision is unique for each θ , this simplifies to the following. For any θ_{-i} , if θ_i satisfies (1) then no agent is pivotal, so the CSP is EPBB at θ .

⁴²To see this, let $p = \mathbb{P}(\theta_i = \bar{\theta}_0) < 1$. Then, $\mathbb{P}(\text{EPBB}) \leq \mathbb{P}(\text{EPBB}$ in the CSP) + $\mathbb{P}(\exists i \in I, \forall j \neq i, \theta_j = \bar{\theta}_0)$ which goes to zero since $\mathbb{P}(\exists i \in I, \forall j \neq i, \theta_j = \bar{\theta}_0) \leq \sum_{i=1}^n \mathbb{P}(\forall j \neq i, \theta_j = \bar{\theta}_0) = np^{n-1} \to 0$ as n goes to infinity. Moreover, conditional on at least one pivotal player, the total rebate is no larger than n. Hence, $\frac{1}{n}\mathbb{E}(\text{distance from EPBB}) \leq \frac{1}{n}\mathbb{E}(\text{distance from EPBB} \text{ in the CSP}) + \frac{1}{n}n^2p^{n-1} \to 0$ as n goes to infinity.

Theorem 5 shows that these are the *only* kinds of mechanisms which satisfy these five criteria—i.e., that all such mechanisms are simply perturbations of the costsharing pivotal mechanism whose deviations disappear in the limit. In particular, given any sequence of utilitarian and monotone decision rules, any asymptotically ex-post budget-balanced sequence of mechanisms satisfying strategy-proofness and at least one of cost-sharing universal participation and the fair pricing principle must have a transfer rule equal to that of the cost-sharing pivotal mechanism plus a term h_i^n which is 1) non-negative, 2) zero with probability one in the limit, and 3) zero in expectation, averaged across individuals, in the limit. In this sense, the cost-sharing pivotal mechanism is the *unique* solution to the government's public good provision problem.

Theorem 5. Consider any sequence of i.i.d. random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}((\theta_i(y_1) - \theta_i(y_0))^2) < \infty$ for all $y_0, y_1 \in Y$, and consider any induced sequence of finite environments $\mathcal{E}_F^n \in \mathbb{E}_F^{\mathbb{N}}$ satisfying Assumption 2. Let α^n be a sequence of utilitarian and monotone decision rules and τ^n be a sequence of CSP transfer rules. A sequence of utilitarian and monotone mechanisms $(\alpha^n, \hat{\tau}^n)$ satisfies asymptotic ex-post budget-balance, strategy-proofness for each n, and cost-sharing universal participation for each n if and only if, for all i,

$$\hat{\tau}_i^n(\theta^n, z^n) = \tau_i^n(\theta^n, z^n) + h_i^n(\theta_{-i}^n, z^n),$$

for some sequence of functions $h_i^n: \Theta_0^{n-1} \times Z_0^n \to \mathbb{R}$ such that

- 1. $h_i^n(\theta_{-i}^n, z^n) \leq 0$ for all $\theta_{-i}^n \in \Theta_0^{n-1}$, $z^n \in Z_0^n$, *i*, and *n*,
- 2. $\lim_{n\to\infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \text{ for all } i) = 1, \text{ and}$
- 3. $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E}(h_i^n(\theta_{-i}^n, z^n)) = 0.$

The same holds replacing cost-sharing universal participation with the fair pricing principle per unit. Under Assumption 1, the same holds replacing the fair pricing principle per unit with the fair pricing principle.

Proof Sketch. 1. Follows from Proposition 1, Theorem 2, and Theorem 3.

- 2. The probability that no agent is pivotal goes to one (Proposition 4), and the CSP is EPBB when no agent is pivotal. Hence, by (1), $\mathbb{P}(\text{EPBB}) \to 1 \iff \mathbb{P}(\sum_{i=1}^{n} h_i^n(\theta_{-i}^n, z^n) = 0 | \text{ no one pivotal}) \to 1 \iff \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \text{ for all } i | \text{ no one pivotal}) \to 1 \iff \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \text{ for all } i | no \text{ one pivotal}) \to 1 \iff \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \text{ for all } i) \to 1.$
- 3. Follows from Theorem 4.

9 The Net-Value Approach and the Clarke Mechanism

In this section, I discuss a common methodological simplification used in mechanism design, where values are taken to be net of cost shares, and argue that it is not without loss. I call this the net value approach. I then discuss the canonical mechanism in public good provision, the Clarke mechanism, and contrast it with the cost-sharing pivotal mechanism.

The rest of the section is organized as follows. First, I provide an overview of the net value approach and discuss why it is not without loss. Second, I discuss the participation constraint that arises when plugging net values into universal participation (net-value universal participation). Third, I discuss the mechanism that arises when plugging net values into the pivotal mechanism (the Clarke mechanism). Because the pivotal and Clarke mechanisms are defined with an efficient decision rule, the analysis in this section assumes all welfare weights are equal to one.

9.1 Overview of the Net Value Approach

The standard approach in the literature to modeling production costs is to embed them into the individuals' values and proceed with the analysis as if there were no production costs to begin with. Instead of using an individual *i*'s value v_i for some alternative *y*—say, a park—we use instead an individual *i*'s net value \tilde{v}_i for the construction of the park and paying a tax equal to her fair share of its cost $c_i(y)$. In particular, if *i* has type θ_i , her net value for *y*, including a tax of $c_i(y)$, is $\tilde{v}_i(y, \theta_i) = v_i(y, \theta_i) - c_i(y)$, and her net transfer, the additional payment beyond the tax $c_i(\alpha(\theta))$, is $\tilde{\tau}_i(\theta) = \tau_i(\theta) - c_i(\alpha(\theta))$.

Mechanically, this means that when facing an environment with production costs, we may simply specify the taxes for each alternative, ask individuals to report their values net of this tax, and utilize these values as if they were the individuals' actual values and there were no production costs. I call this the *net value approach*.

Indeed, the prevailing view in the literature is that it is without loss of generality to study environments without production costs, since production costs can always be baked into the values in this way.⁴³ Green and Laffont (1979, p. 31) even contend that "there is no real alternative to this approach." This paper shows otherwise. Keeping values and costs separate provides a richer mathematical framework which allows for a wider class of desiderata, mechanisms, and proof techniques than can be constructed when combining values and costs together.

⁴³See Green and Laffont (1979, p. 29), Moulin (1988, p. 205), Varian (1992, p. 426), and Mas-Colell, Whinston and Green (1995, p. 877).

For example, cost-sharing universal participation, the fair pricing principle, and the cost-sharing pivotal mechanism are all naturally defined in terms of individuals' cost shares. Additionally, Theorem 4 cannot be proved using the net value approach. To see this, consider a binary public good environment in which a public good can either be produced or not. We would like to fix a population distribution of values for the public good and consider sequences of i.i.d. random draws from this distribution. Under the net value approach, we would embed cost shares into values and hence assume that these *net* values are drawn i.i.d. from some distribution. But, while individuals' values are i.i.d., individuals' *net* values depend on their cost share, which depends on the total cost of the public good and the observable characteristics of the others, and hence are in general not i.i.d. Indeed, assuming net values are i.i.d. requires the underlying assumptions that 1) cost shares are split equally across individuals and 2) costs increase linearly with population size, which, as discussed in Section 7.2, I do not assume.

Another challenge when using the net value approach is the reinterpretation of properties with and without net values. Using the net value approach, a researcher studies properties and mechanisms without production costs and presumes the results generalize to the case with production costs. However, it is not always clear what some desiderata mean when interpreted with net values rather than with intrinsic values. An important case of this is universal participation, as I discuss in the following section.

9.2 Net-Value Universal Participation

The pivotal mechanism satisfies strategy-proofness, efficiency, and universal participation, which is a desirable set of criteria for public good provision with no production costs. What does this imply for the pivotal mechanism with net values—i.e., the Clarke mechanism? Strategy-proofness with net values is equivalent to strategyproofness without net values, and efficiency with net values is equivalent to efficiency without net values.⁴⁴ However, universal participation with net values is not equivalent to universal participation without net values.

As we will see, this is because the "zero" value changes meaning when we replace values with net values. Recall that $0 \in \Theta_i$ is a type for which $v_i(y,0) = 0$ for all y—i.e., *i*'s value for each alternative is zero. Define $\tilde{0} \in \Theta_i$ to be a type for which $\tilde{v}_i(y, \tilde{0}) = 0$ for all y—i.e., *i*'s net value for each alternative is zero. Given type $\tilde{0}$, *i*'s value for y is then $v_i(y, \tilde{0}) = \tilde{v}_i(y, \tilde{0}) + c_i(y) = c_i(y)$. In other words, if *i* has a net value for y of zero, she in fact has a positive value for y itself—a value exactly equal to her fair share of its cost. Henceforth, assume $\tilde{0}$ is an admissible type for each *i*.

 $[\]overline{ \tau_i(\theta) \geq v_i(\alpha(\theta'_i, \theta_{-i}), \theta_i) - \tau_i(\theta'_i, \theta_{-i}) } = \tilde{v}_i(\alpha(\theta'_i, \theta_{-i}), \theta_i) - \tilde{\tau}_i(\theta'_i, \theta_{-i})$ if and only if $v_i(\alpha(\theta), \theta_i) - \tau_i(\theta) \geq v_i(\alpha(\theta'_i, \theta_{-i}), \theta_i) - \tau_i(\theta'_i, \theta_{-i})$ and $\arg \max_y \sum_{i \in I} \tilde{v}_i(y, \theta_i) = \arg \max_y \sum_{i \in I} v_i(y, \theta_i) - c_i(y).$

Because universal participation is a desirable constraint in public good provision environments without production costs, it is often simply presumed that a mechanism which satisfies universal participation will continue to be desirable when adding production costs via the net value approach. Indeed, to my knowledge, no one has taken the step of explicitly spelling out what universal participation states when net values are employed. Since doing so produces a distinct constraint, I give it its own name—*net-value universal participation*. It states the following.

Definition 17. A mechanism $f = (\alpha, \tau)$ satisfies *net-value universal participation* if for all θ and i,

$$v_i(\alpha(\theta), \theta_i)) - \tau_i(\theta)) \ge v_i(\alpha(\tilde{0}, \theta_{-i}), \theta_i) - c_i(\alpha(\tilde{0}, \theta_{-i})),$$

or equivalently, $\tilde{v}_i(\alpha(\theta), \theta_i) - \tilde{\tau}_i(\theta) \geq \tilde{v}_i(\alpha(\tilde{0}, \theta_{-i}), \theta_i).$

The second line makes clear that net-value universal participation is simply universal participation with net values.⁴⁵ The first line expresses this condition in terms of actual values v_i and transfers τ_i , making clear what it states about the underlying fundamentals: each individual should always prefer to participate in the mechanism rather than to not participate, receive the alternative chosen had she valued each alternative at precisely her fair share of its cost, and be taxed her fair share of the cost of that alternative. In my opinion, this is an unnatural desiderata for these environments.

The difference between cost-sharing universal participation and net-value universal participation is what the mechanism purports to do when an individual does not participate. In the former, the mechanism selects the optimal decision supposing the individual had a value of zero for each good and charges them their fair share of its cost. In the latter, the mechanism selects the optimal decision supposing the individual had a value exactly equal to their fair share of the cost of each good, and charges them their fair share of the cost of that good. Notice that cost-sharing universal participation retains the idea from universal participation that the decision made if an individual does not participate is equivalent to the decision made had they never existed. Net-value universal participation does not.

Finally, each condition can be seen as an alternative way to generalize universal participation to environments with production costs. Cost-sharing universal participation adds to universal participation that, in addition to receiving what is produced without her, an individual is also taxed her fair share of its cost. Net-value universal participation simply reinterprets universal participation with net values.

⁴⁵Expressed in terms of net values, cost-sharing universal participation states that $\tilde{v}_i(\alpha(\theta), \theta_i) - \tilde{\tau}_i(\theta) \geq \tilde{v}_i(\alpha(0, \theta_{-i}), \theta_i)$ for all θ and i.

9.3 The Clarke Mechanism

As noted in the preceding section, cost-sharing universal participation and net-value universal participation can be seen as alternative ways to generalize universal participation to environments with production costs. Exactly the same is true for the cost-sharing pivotal mechanism and the Clarke mechanism. Each mechanism can be seen as an alternative way to generalize the pivotal mechanism to environments with production costs. The cost-sharing pivotal mechanism appends to the transfers in the pivotal mechanism that, in addition to paying the externality she imposes on society, an individual is also taxed her fair share of the cost of what would have been produced without her. The Clarke mechanism simply reinterprets the pivotal mechanism with net values.⁴⁶

Definition 18. A mechanism $f = (\alpha, \tau)$ is a *Clarke mechanism* if the decision rule is efficient and the transfer rule satisfies, for all *i* and θ ,

$$\begin{aligned} \tau_i(\theta) &= \Big(\sum_{j \neq i} v_j(\alpha(\tilde{0}, \theta_{-i}), \theta_j) - \sum_{j \neq i} c_j(\alpha(\tilde{0}, \theta_{-i}))\Big) \\ &- \Big(\sum_{j \neq i} v_j(\alpha(\theta), \theta_j) - \sum_{j \neq i} c_j(\alpha(\theta))\Big) + c_i(\alpha(\theta)), \end{aligned}$$

or equivalently, if the net transfer rule satisfies, for all i and θ ,

$$\tilde{\tau}_i(\theta) = \sum_{j \neq i} \tilde{v}_j(\alpha(\tilde{0}, \theta_{-i}), \theta_j) - \sum_{j \neq i} \tilde{v}_j(\alpha(\theta), \theta_j)$$

The second line makes clear that the Clarke mechanism is simply a pivotal mechanism with net values and no production costs. The first line expresses actual transfers τ_i in terms of actual values v_i , making clear what the mechanism states about the underlying fundamentals: each individual pays the loss imposed on society—excluding her own benefits from the public good and supposing each good were cheaper by precisely her fair share of its cost—when taking both her preferences and her share of the cost into account for the decision, plus her fair share of the cost of what is produced. Compare this to the cost-sharing pivotal mechanism, in which each individual pays the loss imposed on society—excluding her own benefits from the public good—when taking her preferences into account for the decision, plus her fair share of the cost of what share of the cost of what would have been produced without her.

The definition of a Clarke mechanism in terms of the underlying fundamentals is clunky. In a pivotal mechanism, each individual pays the loss she imposes on society

⁴⁶The terms pivotal and Clarke mechanism are often used interchangeably in the literature (see, e.g., Mas-Colell, Whinston and Green (1995, p. 878). See Appendix B for a disambiguation of these two mechanisms.

when taking her preferences into account for the decision, excluding her own benefits from the public good. This is simple and intuitive. But plugging net values into the pivotal mechanism and being explicit about the result is much less so. As with net-value universal participation, to my knowledge, this is the first time the Clarke mechanism has been expressed in this way.

We know from Theorem 2 that the pivotal mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, efficiency, and universal participation. Reinterpreting this with net values immediately implies that the Clarke mechanism can be characterized as the mechanism which maximizes ex-post revenue among all mechanisms satisfying strategy-proofness, efficiency, and net-value universal participation.

Corollary 1. Given any convex environment $\mathcal{E}_X \in \mathbb{E}_X$, a mechanism maximizes expost revenue among all mechanisms which satisfy strategy-proofness, efficiency, and net-value universal participation if and only if it is a Clarke mechanism.

Figure 3 depicts the cost-sharing pivotal mechanism and the Clarke mechanism graphically for the case of a continuous public good y. $v' = \sum_k v'_k(y, \theta_k)$ is the marginal social benefit of the public good. $v'_{-i} = \sum_{j \neq i} v'_j(y, \theta_j)$ is the marginal social benefit of the public good excluding i. c' is the marginal social cost of the public good. c'_i is i's marginal fair cost share. $c'_{-i} = c'(y) - c'_i(y)$ is the marginal social cost of the public good minus i's marginal fair cost share. $\alpha(\theta)$ is the efficient decision (where v' and c'cross). $\alpha(0, \theta_{-i})$ is the efficient decision without i (where v'_{-i} and c' cross). $\tilde{\alpha}(\theta_{-i})$ is the efficient decision ignoring i's preferences and i's share of the cost (where v'_{-i} and c'_{-i} cross).⁴⁷

The Clarke mechanism implements $\alpha(\theta)$ and charges *i* the loss imposed on society excluding her own benefits from the public good and supposing each good were cheaper by precisely her fair share of its cost—when taking both her preferences and her share of the cost into account for the decision (the area in red on the bottom panels), plus her fair share of the cost of what is produced (the area in green on the bottom panels). The cost-sharing pivotal mechanism implements $\alpha(\theta)$ and charges *i* the loss imposed on society—excluding her own benefits from the public good—when taking her preferences into account for the decision (the area in red on the top panels), plus her fair share of the cost of what would have been produced without her (the area in green on the top panels).

Notice that, as the right panels of Figure 3 show, the Clarke mechanism incorporates a somewhat unusual notion of what it means for an individual to be pivotal. Suppose i has zero value for the public good. She is *not* pivotal, as the efficient decision is the same with or without her. The pivotal mechanism charges i zero, and the cost-sharing

⁴⁷Formally, $\tilde{\alpha}(\theta_{-i}) \in \arg\max_{y \in Y} \sum_{j \neq i} v_j(y, \theta_j) - c_j(y)$ and $\alpha(\theta) \in \arg\max_{y \in Y} \sum_{i \in I} v_i(y, \theta_i) - c_i(y)$. Note that while $\tilde{\alpha}(\theta_{-i}) > \alpha(\theta)$ in Figure 3, it may also be that $\tilde{\alpha}(\theta_{-i}) \leq \alpha(\theta)$.



Figure 3

pivotal mechanism charges *i* her fair share of the cost of what is produced (the green triangle in the top right panel). The Clarke mechanism, on the other hand, charges *i* her fair share of the cost of what is produced plus a modified conception of a "pivotal" payment (the red triangle in the bottom right panel). Ignoring *i*'s preference does not change the optimal decision ($\alpha(0, \theta_{-i}) = \alpha(\theta)$), but ignoring *i*'s share of the cost makes the good cheaper, and if the good is cheaper, the others want to produce more of it ($\tilde{\alpha}(\theta_{-i}) > \alpha(\theta)$). The red triangle captures this "loss" to society—the welfare loss of not getting to produce more of the good if it were cheaper by precisely *i*'s cost share. In the Clarke mechanism, an individual is "pivotal" if the efficient decision changes when excluding her preferences and her fair share of the cost.

This distinctive feature of the Clarke mechanism ultimately leads it to violate desirable participation constraints and fairness principles. In particular, the Clarke mechanism violates cost-sharing universal participation, the fair pricing principle, and no-extortion. Violations of all three can be seen in Example 1. Violations of the first two can also be seen in the bottom right panel of Figure 3, where i has a zero value for the good but is asked to pay more than her fair share for what is produced. By contrast, the cost-sharing pivotal mechanism satisfies all three.

Although cost-sharing universal participation neither implies nor is implied by netvalue universal participation, it turns out that the Clarke mechanism always raises more revenue than the cost-sharing pivotal mechanism.

Fact 2. Given any convex environment $\mathcal{E}_X \in \mathbb{E}_X$ with all welfare weights equal to one, the transfer for each individual *i* in a Clarke mechanism is no less than that in a cost-sharing pivotal mechanism.

Proof. $\tau_i^{\text{Clarke}}(\theta) \geq \tau_i^{\text{CSP}}(\theta)$ if and only if

$$\sum_{j \neq i} v_j(\alpha(\tilde{0}, \theta_{-i}), \theta_j) - c_j(\alpha(\tilde{0}, \theta_{-i})) = \max_{y \in Y} \sum_{j \neq i} v_j(y, \theta_j) - c_j(y)$$
$$\geq \sum_{j \neq i} v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} c_k(\alpha(0, \theta_{-i})) + c_i(\alpha(0, \theta_{-i})).$$

In fact, one of the main selling points of the Clarke mechanism in the literature is that it never runs a budget deficit.

Definition 19. A mechanism $f = (\alpha, \tau)$ satisfies *no-deficit* if $\sum_{i \in I} \tau_i(\theta) \ge c(\alpha(\theta))$ for all θ .

Nevertheless, as discussed in Section 7.1, the approximate budget-balance principle argues that what really matters for the government's public good provision problem is that future changes to tax policy which balance the budget in the long run do not distort individual incentives—and what really matters for *this* is that the expected distance from budget-balance per capita is small. It is not particularly important that a mechanism never runs a budget deficit. Arguably, this is even somewhat of a red herring. In fact, it turns out that never running a budget deficit and fairness are fundamentally in conflict with each other. No strategy-proof and utilitarian mechanism can satisfy both no-deficit and no-extortion, as I show in the next section.

10 Impossibility of No-Deficit and No-Extortion

In this section, I show that no strategy-proof and utilitarian mechanism can satisfy both no-deficit (the mechanism never runs a budget deficit) and no-extortion (if nothing is produced, no one pays). No-deficit constrains transfers from below, no-extortion constrains transfers from above, and, alongside strategy-proofness and utilitarianism, they are incompatible. Let $\mathbb{E}_B \subset \mathbb{E}_F$ denote the set of binary public good environments in which a public good can either be produced or not (i.e., the set of finite public good environments with $\bar{y} = 1$). **Proposition 6.** There does not exist any mechanism that satisfies strategy-proofness, utilitarianism, no-extortion, and no-deficit in all binary public good environments (and hence, in all finite, continuous, and convex public good environments).

Proof. A potential project-starter is an individual for whom the project may or may not be produced, depending on her report. In particular, a player i is a potential project-starter for θ_{-i} if $\alpha(0, \theta_{-i}) = 0$ and there exists $\theta_i \in \Theta_i$ such that $\alpha(\theta) > 0$.

We first show that a mechanism satisfying strategy-proofness, utilitarianism, and no-extortion can charge potential project-starters no more than their generalized pivotal payment. Given any convex environment $\mathcal{E}_X \in \mathbb{E}_X$, a mechanism f is strategy-proof and utilitarian if and only if it is a generalized Groves mechanism by Proposition 1. Let $f = (\alpha, \tau)$ be a generalized Groves mechanism with $\tau_i(\theta) = \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i}))] - \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i}))] + h_i(\theta_{-i})$ for some $h_i : \Theta_{-i} \to \mathbb{R}$. If i is a potential project-starter for θ_{-i} , then no-extortion requires $\tau_i(0, \theta_{-i}) = \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(0, \theta_j) - \sum_{k \in I} \lambda_k c_k(0)] - \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(0, \theta_j) - \sum_{k \in I} \lambda_k c_k(0)] + h_i(\theta_{-i}) = h_i(\theta_{-i}) \leq 0.$

We now show the result. Consider a binary public good environment with three individuals each with equal welfare weights and value 3 for the public good which costs 8 to produce. The utilitarian decision is to produce the public good. Each player is a potential project-starter, so in order to satisfy no-extortion their transfer must be no larger than their pivotal payment by the previous argument, which is 2. Hence, total revenue can be no more than 6, violating no-deficit.

This result forces us to be careful about what we seek. No-deficit is an intuitive criterion, but arguably isn't what really matters for practical large-scale mechanism design. By the approximate budget-balance principle (small changes to future tax policy do not affect individual incentives), the appropriate criterion for large-scale mechanisms is asymptotic ex-post budget-balance. Moreover, while no-extortion can be satisfied alongside asymptotic ex-post budget-balance, it cannot be satisfied alongside asymptotic ex-post budget-balance.

11 Conclusion

In this paper, I lay out the government's public good provision problem from first principles and identify its solution. The cost-sharing pivotal mechanism (a) provides incentives for each individual to truthfully report their willingness to pay for the public good (strategy-proofness), (b) provides incentives for each individual to participate in the process (cost-sharing universal participation), (c) produces the welfaremaximizing quantity of the public good (utilitarianism), (d) raises exactly the amount of revenue necessary to produce this quantity in large populations (asymptotic ex-post budget-balance), and (e) does not ask anyone to pay what has been deemed an unfair amount (the fair pricing principle). I conclude with a constructive interpretation about how to implement this mechanism in practice.

Taken literally, the cost-sharing pivotal mechanism asks each individual to report their willingness to pay for the public good(s), and then, upon receiving everyone's reports, implements the utilitarian decision and requests a payment from each individual. The generic outcome in the cost-sharing pivotal mechanism is that each individual is asked to pay precisely her fair share of what is produced, and so, upon hearing such a decision, each individual would be expected to transfer this amount to the government.

But instead of thinking of an individual's fair cost share as a predetermined amount, agreed upon in advance by a group of democratically elected officials, researchers, and ordinary citizens (as I proposed in Section 3), we can *also* think of an individual's fair cost share as being whatever the existing tax system would ultimately collect from her if the government needed to raise additional funds to finance the public good. That is, what constitutes an individual's fair cost share is *built into* the existing tax system. Indeed, this is how public goods are normally funded.⁴⁸

Under such an interpretation, the only explicit payments an individual needs to make or receive are the adjustments to their transfer that occur if and only if they sway the decision.⁴⁹ If an individual is pivotal (which occurs almost never), then today she pays her pivotal payment minus a discount equal to the difference between her fair share of the cost of what is produced and her fair share of the cost of what would have been produced without her. If an individual is not pivotal (which occurs generically), her payment today is zero. In either case, the tax system ultimately collects her fair share of the cost of what is produced over time.

References

- Adler, Matthew D. 2012. Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis. Oxford University Press.
- Adler, Matthew D. 2016. "Benefit–Cost Analysis and Distributional Weights: An Overview." *Review of Environmental Economics and Policy*, 10(2).

$$\frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i})) \Big) \\ - \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta)) \Big) - \Big(c_i(\alpha(\theta)) - c_i(\alpha(0, \theta_{-i})) \Big).$$

⁴⁸Note that this means fair cost shares cannot depend on observable characteristics which relate specifically to the public good, like an individual's distance from the public good.

 $^{^{49}}$ In particular, each individual *i*'s payment today is the difference between her transfer in the mechanism and her fair share of the cost of what is produced,

- Akbarpour, Mohammad, Eric Budish, Piotr Dworczak, and Scott Duke Kominers. 2024. "An Economic Framework for Vaccine Prioritization." Quarterly Journal of Economics, 139(1): 359–417.
- Akbarpour, Mohammad, Piotr Dworczak, and Scott Duke Kominers. 2023. "Redistributive Allocation Mechanisms."
- Al-Najjar, Nabil I., and Rann Smorodinsky. 2000. "Provision of a public good with bounded cost." *Economics Letters*, 67(3): 297–301.
- Blackorby, Charles, and David Donaldson. 1990. "A Review Article: The Case against the Use of the Sum of Compensating Variations in Cost-Benefit Analysis." *Canadian Journal of Economics*, 23(3): 471–494.
- **Boadway, Robin.** 2016. "Cost-Benefit Analysis." In *The Oxford Handbook of Well-Being and Public Policy*., ed. Matthew D. Adler and Marc Fleurbaey, Chapter 3. Oxford University Press.
- Bressler, R. Daniel, and Geoffrey Heal. 2022. "Valuing Excess Deaths Caused by Climate Change."
- Carbajal, Juan Carlos, Andrew McLennan, and Rabee Tourky. 2013. "Truthful implementation and preference aggregation in restricted domains." *Journal of Economic Theory*, 148(3): 1074–1101.
- Carroll, Gabriel. 2023. "Concepts of Fairness in Economics." Lecture Notes.
- Clarke, Edward H. 1971. "Multipart Pricing of Public Goods." *Public Choice*, 11: 17–33.
- Drexl, Moritz, and Andreas Kleiner. 2018. "Why Voting? A Welfare Analysis." American Economic Journal: Microeconomics, 10(3): 253–271.
- **Dworczak, Piotr, Scott Duke Kominers, and Mohammad Akbarpour.** 2021. "Redistribution Through Markets." *Econometrica*, 89(4): 1665–1698.
- Fleurbaey, Marc, and Rossi Abi-Rafeh. 2016. "The Use of Distributional Weights in Benefit–Cost Analysis: Insights from Welfare Economics." *Review of Environmental Economics and Policy*, 10(2): 286–307.
- Fleurbaey, Marc, Stéphane Luchini, Christophe Muller, and Erik Schokkaert. 2013. "Equivalent Income and Fair Evaluation of Health Care." *Health Economics*, 22: 711–729.
- Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica*, 41(4): 587–601.

- Green, Jerry R., and Jean-Jacques Laffont. 1977. "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods." *Econometrica*, 45(2): 427–438.
- Green, Jerry R., and Jean-Jacques Laffont. 1979. Incentives in Public Decision-Making. Vol. 1, North-Holland Publishing Company.
- Green, Jerry R., Elon Kohlberg, and Jean-Jacques Laffont. 1976. "Partial Equilibrium Approach to the Free-Rider Problem." *Journal of Public Economics*, 6(4): 375–394.
- Groves, Theodore. 1973. "Incentives in Teams." Econometrica, 41(4): 617–631.
- Hellwig, Martin F. 2003. "Public-Good Provision with Many Participants." *Review* of Economic Studies, 70(3): 589–614.
- Holmström, Bengt. 1979. "Groves' Scheme on Restricted Domains." *Econometrica*, 47(5): 1137–1144.
- Kuzmics, Christoph, and Jan-Henrik Steg. 2017. "On public good provision mechanisms with dominant strategies and balanced budget." *Journal of Economic Theory*, 170: 56–69.
- Laffont, Jean-Jacques, and Eric S. Maskin. 1982. "The Theory of Incentives: An Overview." In *Advances in Economic Theory*. Chapter 2, 31–94. Cambridge University Press.
- Mailath, George J., and Andrew Postlewaite. 1990. "Asymmetric Information Bargaining Problems with Many Agents." *Review of Economic Studies*, 57(3): 351– 367.
- Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. Microeconomic Theory. Oxford University Press.
- Milgrom, Paul, and Chris Shannon. 1994. "Monotone Comparative Statics." *Econometrica*, 62(1): 157–180.
- Mitsui, Toshihide. 1983. "Asymptotic Efficiency of the Pivotal Mechanism with General Project Space." Journal of Economic Theory, 31: 318–331.
- Moulin, Hervé J. 1986. "Characterizations of the Pivotal Mechanism." Journal of Public Economics, 53–78.
- Moulin, Hervé J. 1988. Axioms of Cooperative Decision Making. Cambridge University Press.

- Nath, Swaprava, and Tuomas Sandholm. 2019. "Efficiency and budget balance in general quasi-linear domains." *Games and Economic Behavior*, 113: 673–693.
- Pai, Mallesh, and Philipp Strack. 2022. "Taxing Externalities Without Hurting the Poor."
- Reuter, Marco, and Carl-Christian Groh. 2023. "Mechanism Design for Unequal Societies."
- Roberts, Kevin. 1979. "The Characterization of Implementable Choice Rules." In Aggregation and Revelation of Preferences., ed. Jean-Jacques Laffont, Chapter 18, 321–348. North-Holland Publishing Company.
- **Rob, Rafael.** 1982. "Asymptotic Efficiency of the Demand Revealing Mechanism." Journal of Economic Theory, 28: 207–220.
- Serizawa, Shigehiro. 1999. "Strategy-Proof and Symmetric Social Choice Functions for Public Good Economies." *Econometrica*, 67(1): 121–145.
- van der Vaart, A. W. 1998. Asymptotic Statistics. Cambridge University Press.
- Varian, Hal R. 1992. Microeconomic Analysis. . 3rd ed., Norton & Company, Inc.
- Vickrey, William. 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders." Journal of Finance, 16(1): 8–37.
- Xi, Jin, and Haitian Xie. 2023. "Strength in numbers: robust mechanisms for public goods with many agents." Social Choice and Welfare.

A Fair Cost Shares: Some Examples

In this appendix, I present some examples of how a government might construct fair cost shares from observable traits. A simple notion of fairness is that of equal cost shares, which says that it is fair for everyone to pay an equal share of the cost: $c_i(y,z) = c(y)/n$ for all y and z. But as discussed in Section 3, fair shares may depend on anything which is observable to the government, including income level and distance from the public good.

Suppose income levels w_i are observable (i.e., w_i is a component of z_i). An appealing notion of fairness might then be that it is fair for everyone to pay an equal share of their *income*:

$$c_i(y,z) = c(y) \frac{w_i}{\sum_j w_j}$$

for all y and z. In fact, a natural generalization of this idea is that it is fair for everyone to pay an equal share of their duty:

$$c_i(y, z) = c(y) \frac{\delta(z_i)}{\sum_j \delta(z_j)},$$

where $Z_i = Z_0$ for all i and $\delta : Z_0 \to \mathbb{R}_+$ is an index measuring i's relative duty to pay for the public good. If distance d_i from the public good is observable, an appealing index of duty might be one in which duty is constant within a particular radius r and is inversely proportional to distance beyond r,

$$\delta(d_i) = \begin{cases} 1/r & \text{if } 0 \le d_i \le r \\ 1/d_i & \text{if } d_i > r \end{cases}$$

If both income w_i and distance d_i are observable, an appealing index of duty might be

$$\delta(w_i, d_i) = \begin{cases} w_i/r & \text{if } 0 \le d_i \le r \\ w_i/d_i & \text{if } d_i > r \end{cases},$$

so that doubling one's income doubles one's duty and doubling one's distance halves one's duty beyond radius r.

Alternatively, what constitutes a fair cost share can simply be derived from the current tax system. See Section 11 for a discussion.

B Disambiguating the Pivotal and Clarke Mechanisms

The pivotal mechanism is generally defined in a standard mechanism design environment (without production costs). In such environments, it is defined as follows.

Definition 20. A mechanism $f = (\alpha, \tau)$ is a pivotal mechanism (without production costs) if $\alpha(\theta) \in \arg \max_{y \in Y} \sum_{i \in I} v_i(y, \theta_i)$ and

$$\tau_i(\theta) = \sum_{j \neq i} v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{j \neq i} v_j(\alpha(\theta), \theta_j)$$

In words, the pivotal mechanism employs an efficient decision rule and charges each individual the loss imposed on society—excluding her own benefits from the public good—when taking her preferences into account for the decision.

How then does one define the pivotal mechanism in environments with production costs? Production costs change the efficient decision rule from maximizing $\sum_{i \in I} v_i(y, \theta_i)$ to maximizing $\sum_{i \in I} v_i(y, \theta_i) - c(y)$. In this paper, I define the pivotal mechanism identically as before, except with this updated decision rule. **Definition 21.** A mechanism $f = (\alpha, \tau)$ is a pivotal mechanism (with production costs) if $\alpha(\theta) \in \arg \max_{y \in Y} \sum_{i \in I} v_i(y, \theta_i) - c(y)$ and

$$\tau_i(\theta) = \Big(\sum_{j \neq i} v_j(\alpha(0, \theta_{-i}), \theta_j) - c(\alpha(0, \theta_{-i}))\Big) - \Big(\sum_{j \neq i} v_j(\alpha(\theta), \theta_j) - c(\alpha(\theta))\Big).$$

Notice that, with this generalized definition, the natural language description of the mechanism remains unchanged. The pivotal mechanism employs an efficient decision rule and charges each individual the loss imposed on society—excluding her own benefits from the public good—when taking her preferences into account for the decision.

Still, it is standard to do something different in the literature. To generalize the pivotal mechanism to environments with production costs, instead of updating the notion of efficiency and leaving everything else unchanged, it is standard to use the net value approach, embedding costs into individual values and otherwise ignoring production costs. This produces the same decision rule as above, but a different transfer rule. I call this mechanism a Clarke mechanism, though it is common in the literature to refer to it as a pivotal mechanism as well. That is, it is common in the literature to consider the pivotal and Clarke mechanisms to be equivalent (see, for instance, Mas-Colell, Whinston and Green (1995, p. 878)). Plugging net values into the pivotal mechanism without production costs gives rise to the following, reprinted from Section 9.3.

Definition 18. A mechanism $f = (\alpha, \tau)$ is a *Clarke mechanism* if the decision rule is efficient and the transfer rule satisfies, for all *i* and θ ,

$$\begin{aligned} \tau_i(\theta) &= \left(\sum_{j \neq i} v_j(\alpha(\tilde{0}, \theta_{-i}), \theta_j) - \sum_{j \neq i} c_j(\alpha(\tilde{0}, \theta_{-i}))\right) \\ &- \left(\sum_{j \neq i} v_j(\alpha(\theta), \theta_j) - \sum_{j \neq i} c_j(\alpha(\theta))\right) + c_i(\alpha(\theta)), \end{aligned}$$

or equivalently, if the net transfer rule satisfies, for all i and θ ,

$$\tilde{\tau}_i(\theta) = \sum_{j \neq i} \tilde{v}_j(\alpha(\tilde{0}, \theta_{-i}), \theta_j) - \sum_{j \neq i} \tilde{v}_j(\alpha(\theta), \theta_j).$$

Unlike before, the natural language description of this mechanism is quite different than that of the original pivotal mechanism. The Clarke mechanism employs an efficient decision rule and charges each individual the loss imposed on society—excluding her own benefits from the public good and supposing each good were cheaper by precisely her fair share of its cost—when taking both her preferences and her share of the cost into account for the decision, plus her fair share of the cost of what is produced.

C Proofs Omitted from Main Body

Theorem 3. Consider any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$ and any utilitarian and monotone decision rule α . A mechanism $f = (\alpha, \tau)$ maximizes ex-post revenue among all mechanisms which satisfy strategy-proofness, utilitarianism, and the fair pricing principle per unit if and only if it is a cost-sharing pivotal mechanism. Under Assumption 1, the same holds replacing the fair pricing principle per unit with the fair pricing principle.

Proof. Given any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$, a mechanism f is strategy-proof and utilitarian if and only if it is a generalized Groves mechanism by Proposition 1. Let $f = (\alpha, \tau)$ be a generalized Groves mechanism with a monotone decision rule α and $\tau_i(\theta) = g_i(\theta_{-i}) - \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta))]$ for some $g_i : \Theta_{-i} \to \mathbb{R}$.

Part I. We would like to show that $g_i(\theta_{-i}) = \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i}))] + c_i(\alpha(0, \theta_{-i}))$ maximizes ex post revenue subject to the FPP per unit. The FPP per unit requires that for all *i* and θ , $\int_0^{\alpha(\theta)} \max\{v'_i(y, \theta_i), c'_i(y)\} dy + \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(\theta))] \ge g_i(\theta_{-i})$, so to maximize ex-post revenue subject to the FPP per unit, set

$$g_i(\theta_{-i}) = \inf_{\theta_i \in \Theta_i} \left\{ \int_0^{\alpha(\theta)} \max\{v_i'(y,\theta_i), c_i'(y)\} + \frac{1}{\lambda_i} \left(\sum_{j \neq i} \lambda_j v_j'(y,\theta_j) - \sum_{k \in I} \lambda_k c_k'(y) \right) \, \mathrm{d}y \right\}.$$

We would like to show that for any θ , increasing $v'_i(y, \theta_i)$ pointwise weakly increases the objective function, and hence that $0 \in \Theta_i$ is a minimizer. Suppose $v'_i(y, \hat{\theta}_i) \ge v'_i(y, \theta_i)$ for all $y \ge 0$. Since α is monotonic, $\alpha(\hat{\theta}_i, \theta_{-i}) \ge \alpha(\theta)$. We would like to show

$$\begin{split} \int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} \max\{v_{i}'(y,\hat{\theta}_{i}), c_{i}'(y)\} + \frac{1}{\lambda_{i}} \Big(\sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y)\Big) \, \mathrm{d}y \\ \geq \int_{0}^{\alpha(\theta)} \max\{v_{i}'(y,\theta_{i}), c_{i}'(y)\} + \frac{1}{\lambda_{i}} \Big(\sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y)\Big) \, \mathrm{d}y, \end{split}$$

or equivalently,

$$\begin{split} \int_{0}^{\alpha(\theta)} \max\left\{v_{i}'(y,\hat{\theta}_{i}), c_{i}'(y)\right\} &- \max\left\{v_{i}'(y,\theta_{i}), c_{i}'(y)\right\} \,\mathrm{d}y \\ &+ \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} \max\left\{v_{i}'(y,\hat{\theta}_{i}), c_{i}'(y)\right\} + \frac{1}{\lambda_{i}} \Big(\sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y)\Big) \,\mathrm{d}y \geq 0. \end{split}$$

The first term is non-negative, and by definition of α ,

$$\begin{split} \int_{0}^{\alpha(\theta_{i},\theta_{-i})} v_{i}'(y,\hat{\theta}_{i}) &+ \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}'(y) \Big) \, \mathrm{d}y \\ &\geq \int_{0}^{\alpha(\theta)} v_{i}'(y,\hat{\theta}_{i}) + \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}'(y) \Big) \, \mathrm{d}y \end{split}$$

$$\iff \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v_{i}'(y,\hat{\theta}_{i}) + \frac{1}{\lambda_{i}} \Big(\sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y) \Big) \, \mathrm{d}y \ge 0$$
$$\implies \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} \max \big\{ v_{i}'(y,\hat{\theta}_{i}), \, c_{i}'(y) \big\} + \frac{1}{\lambda_{i}} \Big(\sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y) \Big) \, \mathrm{d}y \ge 0.$$

Hence, $0 \in \Theta_i$ is a minimizer and for all θ_{-i} ,

$$g_i(\theta_{-i}) = \int_0^{\alpha(0,\theta_{-i})} \max\left\{ v_i'(y,0), c_i'(y) \right\} + \frac{1}{\lambda_i} \left(\sum_{j \neq i} \lambda_j v_j'(y,\theta_j) - \sum_{k \in I} \lambda_k c_k'(y) \right) \, \mathrm{d}y$$
$$= \frac{1}{\lambda_i} \left(\sum_{j \neq i} \lambda_j v_j(\alpha(0,\theta_{-i}),\theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0,\theta_{-i})) \right) + c_i(\alpha(0,\theta_{-i})).$$

Part II. Suppose Assumption 1. We would like to show that $g_i(\theta_{-i}) = \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v_j(\alpha(0, \theta_{-i}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\alpha(0, \theta_{-i}))] + c_i(\alpha(0, \theta_{-i}))$ maximizes ex-post revenue subject to the FPP. The FPP requires that for all *i* and θ , max $\{v_i(\alpha(\theta), \theta_i), c_i(\alpha(\theta))\} + \frac{1}{\lambda_i} [\sum_{j \neq i} \lambda_j v'_j(\alpha(\theta), \theta_j) - \sum_{k \in I} \lambda_k c'_k(\alpha(\theta))] \ge g_i(\theta_{-i})$, so to maximize ex-post revenue subject to the FPP, set

$$g_i(\theta_{-i}) = \inf_{\theta_i \in \Theta_i} \left\{ \max\left\{ \int_0^{\alpha(\theta)} v_i'(y,\theta_i) \, \mathrm{d}y, \int_0^{\alpha(\theta)} c_i'(y) \, \mathrm{d}y \right\} + \frac{1}{\lambda_i} \int_0^{\alpha(\theta)} \sum_{j \neq i} \lambda_j v_j'(y,\theta_j) - \sum_{k \in I} \lambda_k c_k'(y) \, \mathrm{d}y \right\}$$

We would like to show that for any θ , increasing $v'_i(y, \theta_i)$ pointwise weakly increases the objective function, and hence that $0 \in \Theta_i$ is a minimizer. Suppose $v'_i(y, \hat{\theta}_i) \ge v'_i(y, \theta_i)$ for all $y \ge 0$. Since α is monotonic, $\alpha(\hat{\theta}_i, \theta_{-i}) \ge \alpha(\theta)$. We would like to show

$$\max\left\{\int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v_{i}'(y,\hat{\theta}_{i}) \,\mathrm{d}y, \int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} c_{i}'(y) \,\mathrm{d}y\right\} + \frac{1}{\lambda_{i}} \int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} \sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y) \,\mathrm{d}y$$
$$\geq \max\left\{\int_{0}^{\alpha(\theta)} v_{i}'(y,\theta_{i}) \,\mathrm{d}y, \int_{0}^{\alpha(\theta)} c_{i}'(y) \,\mathrm{d}y\right\} + \frac{1}{\lambda_{i}} \int_{0}^{\alpha(\theta)} \sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y) \,\mathrm{d}y$$

which is equivalent to

$$\begin{split} \max\Big\{\int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v_{i}'(y,\hat{\theta}_{i}) \,\mathrm{d}y, \int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} c_{i}'(y) \,\mathrm{d}y\Big\} - \max\Big\{\int_{0}^{\alpha(\theta)} v_{i}'(y,\theta_{i}) \,\mathrm{d}y, \int_{0}^{\alpha(\theta)} c_{i}'(y) \,\mathrm{d}y\Big\} \\ &+ \frac{1}{\lambda_{i}} \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} \sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y) \,\mathrm{d}y \ge 0. \end{split}$$

Notice that, since $\int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v_{i}'(y,\hat{\theta}_{i}) + \frac{1}{\lambda_{i}} [\sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y)] \,\mathrm{d}y \geq \int_{0}^{\alpha(\theta)} v_{i}'(y,\hat{\theta}_{i}) + \frac{1}{\lambda_{i}} [\sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y)] \,\mathrm{d}y,$

$$\int_{\alpha(\theta)}^{\alpha(\theta_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) + \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v'_j(y,\theta_j) - \sum_{k \in I} \lambda_k c'_k(y) \Big) \, \mathrm{d}y \ge 0.$$

Hence, it is sufficient to show

$$\max\left\{\int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v_{i}'(y,\hat{\theta}_{i}) \,\mathrm{d}y, \int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} c_{i}'(y) \,\mathrm{d}y\right\} - \max\left\{\int_{0}^{\alpha(\theta)} v_{i}'(y,\theta_{i}) \,\mathrm{d}y, \int_{0}^{\alpha(\theta)} c_{i}'(y) \,\mathrm{d}y\right\} \\ \ge \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v_{i}'(y,\hat{\theta}_{i}) \,\mathrm{d}y. \quad (2)$$

Case 1. Suppose $\int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v'_{i}(y,\hat{\theta}_{i}) dy < \int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} c'_{i}(y) dy \text{ and } \int_{0}^{\alpha(\theta)} v'_{i}(y,\theta_{i}) dy < \int_{0}^{\alpha(\theta)} c'_{i}(y) dy$. Then (2) becomes $\int_{0}^{\alpha(\hat{\theta}_{i},\theta_{-i})} c'_{i}(y) dy - \int_{0}^{\alpha(\theta)} c'_{i}(y) dy \ge \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v'_{i}(y,\hat{\theta}_{i}) dy$, which holds if and only if $\int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} c'_{i}(y) dy \ge \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_{i},\theta_{-i})} v'_{i}(y,\hat{\theta}_{i}) dy$, which is implied by the first inequality and Assumption 1.

Case 2. Suppose $\int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y < \int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} c'_i(y) \, \mathrm{d}y$ and $\int_0^{\alpha(\theta)} v'_i(y,\theta_i) \, \mathrm{d}y \ge \int_0^{\alpha(\theta)} c'_i(y) \, \mathrm{d}y$. Then (2) becomes $\int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} c'_i(y) \, \mathrm{d}y - \int_0^{\alpha(\theta)} v'_i(y,\theta_i) \, \mathrm{d}y \ge \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y$ which is implied by the first inequality and the definition of $\hat{\theta}_i$.

Case 3. Suppose $\int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y \ge \int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} c'_i(y) \, \mathrm{d}y \text{ and } \int_0^{\alpha(\theta)} v'_i(y,\theta_i) \, \mathrm{d}y < \int_0^{\alpha(\theta)} c'_i(y) \, \mathrm{d}y.$ Then (2) becomes $\int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y - \int_0^{\alpha(\theta)} c'_i(y) \, \mathrm{d}y \ge \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y$, which holds if and only if $\int_0^{\alpha(\theta)} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y \ge \int_0^{\alpha(\theta)} c'_i(y) \, \mathrm{d}y$, which is implied by the first inequality and Assumption 1.

Case 4. Suppose $\int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y \ge \int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} c'_i(y) \, \mathrm{d}y \text{ and } \int_0^{\alpha(\theta)} v'_i(y,\theta_i) \, \mathrm{d}y \ge \int_0^{\alpha(\theta)} c'_i(y) \, \mathrm{d}y.$ Then (2) becomes $\int_0^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y - \int_0^{\alpha(\theta)} v'_i(y,\theta_i) \, \mathrm{d}y \ge \int_{\alpha(\theta)}^{\alpha(\hat{\theta}_i,\theta_{-i})} v'_i(y,\hat{\theta}_i) \, \mathrm{d}y,$ which holds by the definition of $\hat{\theta}_i$.

Hence, $0 \in \Theta_i$ is a minimizer and for all θ_{-i} ,

$$g_{i}(\theta_{-i}) = \max\left\{\int_{0}^{\alpha(0,\theta_{-i})} v_{i}'(y,0) \, \mathrm{d}y, \int_{0}^{\alpha(0,\theta_{-i})} c_{i}'(y) \, \mathrm{d}y\right\} + \frac{1}{\lambda_{i}} \int_{0}^{\alpha(0,\theta_{-i})} \sum_{j\neq i} \lambda_{j} v_{j}'(y,\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}'(y) \, \mathrm{d}y$$
$$= \frac{1}{\lambda_{i}} \Big(\sum_{j\neq i} \lambda_{j} v_{j}(\alpha(0,\theta_{-i}),\theta_{j}) - \sum_{k\in I} \lambda_{k} c_{k}(\alpha(0,\theta_{-i}))\Big) + c_{i}(\alpha(0,\theta_{-i})).$$

Proposition 5. Consider any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$. Suppose that for any *i* and θ_{-i} , max $A^*(0, \theta_{-i})$ exists and there exists $\theta_i \in \Theta_i$ such that

$$v_i'(y,\theta_i) = \begin{cases} \nu_i'(y) & \text{if } y \le \max A^*(0,\theta_{-i}) \\ 0 & \text{if } y > \max A^*(0,\theta_{-i}) \end{cases},$$
(1)

where $\nu'_i(y) \geq \max_{j \neq i} v'_j(y, \theta_j)$ and $\nu'_i(y) > 0$ for all $y.^{50}$ For any utilitarian and

⁵⁰This is a weak richness condition on the domain of values which says that each individual may value marginal units of the public good as much as any other individual and that this marginal value may drop to zero at any point—namely, at max $A^*(0, \theta_{-i})$.

monotone decision rule α , there is no strategy-proof and cost-sharing universal participation mechanism that is no farther from ex-post budget-balance than a cost-sharing pivotal mechanism for every θ . The same holds replacing cost-sharing universal participation with the fair pricing principle per unit. Under Assumption 1, the same holds replacing the fair pricing principle per unit with the fair pricing principle.

Proof. Let

$$t_{i}(\theta) = \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\alpha(0, \theta_{-i}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\alpha(0, \theta_{-i})) \Big) - \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\alpha(\theta), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\alpha(\theta)) \Big)$$
$$= \frac{1}{\lambda_{i}} \int_{0}^{\alpha(0, \theta_{-i})} \sum_{j \neq i} \lambda_{j} v_{j}'(y, \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}'(y) \, \mathrm{d}y - \frac{1}{\lambda_{i}} \int_{0}^{\alpha(\theta)} \sum_{j \neq i} \lambda_{j} v_{j}'(y, \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}'(y) \, \mathrm{d}y$$

be *i*'s generalized pivotal payment and $\tau_i(\theta) = t_i(\theta) + c_i(\alpha(0, \theta_{-i}))$ be *i*'s CSP transfer.

Part I. First, we would like to show that for any θ_{-i} , there exists θ_i such that a CSP mechanism does not run a strict budget surplus—i.e., $\sum_{k \in I} \tau_k(\theta) \leq c(\alpha(\theta))$.

For any θ , $y^* \in A^*(\theta)$ if and only if for any $y' \leq y^* \leq y''$,

$$\int_{y'}^{y^*} \sum_{k \in I} \lambda_k v'_k(y, \theta_k) - \lambda_k c'_k(y) \, \mathrm{d}y \ge 0 \ge \int_{y^*}^{y''} \sum_{k \in I} \lambda_k v'_k(y, \theta_k) - \lambda_k c'_k(y) \, \mathrm{d}y.$$

That is, y^* is utilitarian if and only if moving from any $y' \leq y^*$ to y^* weakly increases total welfare and moving from y^* to any $y'' \geq y^*$ weakly decreases total welfare.

Fix any θ_{-i} . Let θ_i satisfy (1). Then for any $y' < \max A^*(0, \theta_{-i}) < y''$,

$$\int_{y'}^{\max A^*(0,\theta_{-i})} \sum_{l \neq i} \lambda_l v_l'(y,\theta_l) - \sum_{k \in I} \lambda_k c_k'(y) \, \mathrm{d}y \ge 0 > \int_{\max A^*(0,\theta_{-i})}^{y''} \sum_{l \neq i} \lambda_l v_l'(y,\theta_l) - \sum_{k \in I} \lambda_k c_k'(y) \, \mathrm{d}y,$$
(3)

where the strict inequality follows since we are considering max $A^*(0, \theta_{-i})$.

By (3), for any $y' < \max A^*(0, \theta_{-i}) < y''$,

$$\int_{y'}^{\max A^*(0,\theta_{-i})} \sum_{k \in I} \lambda_k v'_k(y,\theta_k) - \lambda_k c'_k(y) \, \mathrm{d}y > 0 > \int_{\max A^*(0,\theta_{-i})}^{y''} \sum_{k \in I} \lambda_k v'_k(y,\theta_k) - \lambda_k c'_k(y) \, \mathrm{d}y,$$

where the first strict inequality follows since $\nu_i(y) > 0$ for all $y \leq A^*(0, \theta_{-i})$. Hence, $\alpha(\theta) = \max A^*(0, \theta_{-i})$. Since $\alpha(0, \theta_{-i}), \alpha(\theta) \in A^*(0, \theta_{-i})$,

$$\int_0^{\alpha(0,\theta_{-i})} \sum_{l\neq i} \lambda_l v_l'(y,\theta_l) - \sum_{k\in I} \lambda_k c_k'(y) \, \mathrm{d}y = \int_0^{\alpha(\theta)} \sum_{l\neq i} \lambda_l v_l'(y,\theta_l) - \sum_{k\in I} \lambda_k c_k'(y) \, \mathrm{d}y,$$

so *i*'s generalized pivotal payment is zero and $\tau_i(\theta) = c_i(\alpha(0, \theta_{-i})) \leq c_i(\alpha(\theta))$ by monotonicity.

For any $j \neq i$, by (3), for any $y' < \max A^*(0, \theta_{-i}) < y''$,

$$\int_{y'}^{\max A^*(0,\theta_{-i})} \sum_{l\neq j} \lambda_l v_l'(y,\theta_l) - \sum_{k\in I} \lambda_k c_k'(y) \,\mathrm{d}y \ge 0 > \int_{\max A^*(0,\theta_{-i})}^{y''} \sum_{l\neq j} \lambda_l v_l'(y,\theta_l) - \sum_{k\in I} \lambda_k c_k'(y) \,\mathrm{d}y$$

by definition of $v'_i(y,\theta_i)$. Hence, $\max A^*(0,\theta_{-i}) \in A^*(0,\theta_{-j})$. Since $\alpha(0,\theta_{-j}), \alpha(\theta) \in A^*(0,\theta_{-j})$,

$$\int_0^{\alpha(0,\theta_{-j})} \sum_{l\neq j} \lambda_l v_l'(y,\theta_l) - \sum_{k\in I} \lambda_k c_k'(y) \, \mathrm{d}y = \int_0^{\alpha(\theta)} \sum_{l\neq j} \lambda_l v_l'(y,\theta_l) - \sum_{k\in I} \lambda_k c_k'(y) \, \mathrm{d}y,$$

so j's generalized pivotal payment is zero and $\tau_j(\theta) = c_j(\alpha(0, \theta_{-j})) \leq c_j(\alpha(\theta))$ by monotonicity.

Hence, $\sum_{k \in I} \tau_k(\theta) \le c(\alpha(\theta))$ as desired.

Part II. Given any continuous environment $\mathcal{E}_C \in \mathbb{E}_C$, a mechanism f is strategy-proof and utilitarian if and only if it is a generalized Groves mechanism by Proposition 1. We may write i's generalized Groves transfer as the sum of her generalized pivotal payment and a term that depends only on her opponents' reports, $t_i(\theta) + h_i(\theta_{-i})$ for some $h_i : \Theta_{-i} \to \mathbb{R}$. For any i and θ_{-i} , if θ_i satisfies (1), then setting $h_i(\theta_{-i}) < c_i(\alpha(0, \theta_{-i}))$ runs a strictly larger budget-deficit than a CSP by Part I, and setting $h_i(\theta_{-i}) > c_i(\alpha(0, \theta_{-i}))$ violates CS-UP by Theorem 2 and violates the FPP per unit (and the FPP under Assumption 1) by Theorem 3.

We now proceed to prove Proposition 3, Proposition 4, and Theorem 4.

Consider any finite environment $\mathcal{E}_F \in \mathbb{E}_F$ and any utilitarian decision rule $\alpha : \Theta \to Y$. For any $\hat{Y} \subseteq Y = \{0, 1, \dots, \bar{y}\}$, let $A^*(\theta; \hat{Y}) = \arg \max_{y \in \hat{Y}} \sum_{i \in I} \lambda_i v_i(y, \theta_i) - \lambda_i c_i(y)$,

$$\hat{\alpha}(\theta; \hat{Y}) = \begin{cases} \alpha(\theta) & \text{if } \alpha(\theta) \in \hat{Y} \\ \max A^*(\theta; \hat{Y}) & \text{otherwise} \end{cases}$$

,

and

$$t_{i}(\theta; \hat{Y}) = \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\hat{\alpha}(0, \theta_{-i}; \hat{Y}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\hat{\alpha}(0, \theta_{-i}; \hat{Y})) \Big) \\ - \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\hat{\alpha}(\theta; \hat{Y}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\hat{\alpha}(\theta; \hat{Y})) \Big).$$

 $A^*(\theta; \hat{Y})$ is the set of utilitarian alternatives within \hat{Y} . $\hat{\alpha}(\theta; \hat{Y})$ is the decision rule that selects according to α when $\alpha(\theta)$ is in \hat{Y} and selects the largest utilitarian alternative in \hat{Y} otherwise. $t_i(\theta; \hat{Y})$ is the generalized pivotal transfer rule associated with $\hat{\alpha}(\theta; \hat{Y})$. For any $y \leq \bar{y}$, let

$$t_i^S(\theta; \{0, 1, \dots, y\}) = t_i(\theta; \{0, 1\}) + t_i(\theta; \{\hat{\alpha}(\theta; \{0, 1\}), 2\}) + t_i(\theta; \{\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{0, 1\}), 2\}), 3\}) + \dots + t_i(\theta; \{\hat{\alpha}(\theta; \dots), y\})$$

be the total transfer associated with running a sequence of binary pivotal mechanisms, where the selected alternative from $\{0,1\}$ is run against 2, the selected alternative from that mechanism is run against 3, and so on. Let $t_i(\theta) = t_i(\theta; Y)$ and $t_i^S(\theta) =$ $t_i^S(\theta; Y)$ be the generalized pivotal transfer and the sequential generalized pivotal transfer for *i*, respectively.

Lemma 1. Given any finite environment $\mathcal{E}_F \in \mathbb{E}_F$, $t_i^S(\theta) \ge t_i(\theta)$ for all θ .

Proof. First notice that

$$t_i^S(\theta) = t_i(\theta; \{0, 1\}) + t_i(\theta; \{\hat{\alpha}(\theta; \{0, 1\}), 2\}) + t_i(\theta; \{\hat{\alpha}(\theta; \{0, 1, 2\}), 3\}) + \ldots + t_i(\theta; \{\hat{\alpha}(\theta; \{0, 1, \ldots, \bar{y} - 1\}), \bar{y}\})$$

We now proceed by induction.

Step 1. We would like to show $t_i^S(\theta; \{0, 1\}) \ge t_i(\theta; \{0, 1\})$. By definition, $t_i^S(\theta; \{0, 1\}) = t_i(\theta; \{0, 1\})$.

Step 2. Suppose by induction that $t_i^S(\theta; \{0, \ldots, y\}) \ge t_i(\theta; \{0, \ldots, y\})$. We would like to show that $t_i^S(\theta; \{0, \ldots, y+1\}) \ge t_i(\theta; \{0, \ldots, y+1\})$.

Case 1. Suppose $\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\}) = y+1$. Then $y^* \equiv \hat{\alpha}(\theta; \{0, \dots, y+1\}) \leq y+1$, and by Proposition 2, $y^* \in A^*(0, \theta_{-i}; \{0, \dots, y+1\})$. Hence, $t_i(\theta; \{0, \dots, y+1\}) = 0$. The result follows since generalized pivotal payments are non-negative.

Case 2. Suppose $\hat{\alpha}(0, \theta_{-i}; \{0, ..., y + 1\}) < y + 1$. We have

$$t_{i}(\theta; \{0, \dots, y+1\}) = \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\})) \Big) \\ - \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\hat{\alpha}(\theta; \{0, \dots, y+1\}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\hat{\alpha}(\theta; \{0, \dots, y+1\})) \Big)$$

and

$$\begin{split} t_{i}^{S}(\theta; \{0, \dots, y+1\}) &= t_{i}^{S}(\theta; \{0, \dots, y\}) \\ &+ \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\})) \Big) \\ &- \frac{1}{\lambda_{i}} \Big(\sum_{j \neq i} \lambda_{j} v_{j}(\hat{\alpha}(\theta; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\}), \theta_{j}) - \sum_{k \in I} \lambda_{k} c_{k}(\hat{\alpha}(\theta; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\})) \Big). \end{split}$$

Hence,

$$\begin{split} t_i^S(\theta; \{0, \dots, y+1\}) &- t_i(\theta; \{0, \dots, y+1\}) = t_i^S(\theta; \{0, \dots, y\}) \\ &+ \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\})) \Big) \\ &- \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\})) \Big) \\ &\geq \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y\})) \Big) \\ &- \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y\})) \Big) \\ &+ \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\})) \Big) \\ &- \frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\})) \Big) \\ &= -\frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \{0, \dots, y+1\})) \Big) \\ &= -\frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\})) \Big) \\ &= -\frac{1}{\lambda_i} \Big(\sum_{j \neq i} \lambda_j v_j(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\}), \theta_j) - \sum_{k \in I} \lambda_k c_k(\hat{\alpha}(0, \theta_{-i}; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\})) \Big) \\ &> 0. \end{split}$$

where the first equality follows by $\hat{\alpha}(\theta; \{0, \dots, y+1\}) = \hat{\alpha}(\theta; \hat{\alpha}(\theta; \{0, \dots, y\}) \cup \{y+1\})$, the first inequality by the inductive hypothesis, and the final equality by the Case 2 assumption.

Proposition 3. Consider any finite environment $\mathcal{E}_F \in \mathbb{E}_F$ and suppose Assumption 2. For any $\theta \in \Theta$, the total generalized pivotal payment can be no less than zero and no more than $\frac{\lambda_H}{\lambda_L}c(\alpha(\theta))$, and the total CSP payment can be no less than zero and no more than $\frac{\lambda_H}{\lambda_L}c(\alpha(\theta)) + c(\bar{y})$. That is,

1.
$$0 \leq \sum_{i=1}^{n} t_i(\theta) \leq \frac{\lambda_H}{\lambda_L} c(\alpha(\theta))$$
 and

2. $0 \leq \sum_{i=1}^{n} \tau_i(\theta) \leq \frac{\lambda_H}{\lambda_L} c(\alpha(\theta)) + c(\bar{y}),$

where $t_i(\theta)$ is i's transfer in a generalized pivotal mechanism and $\tau_i(\theta) = t_i(\theta) + c_i(\alpha(0, \theta_{-i}))$ is i's transfer in a CSP mechanism. If the decision rule is monotone, the upper bound on the total CSP payment can be lowered to $\frac{\lambda_H + \lambda_L}{\lambda_L} c(\alpha(\theta))$.

Proof. **Part I.** Step 1. First, we show $0 \leq \sum_{i=1}^{n} t_i(\theta) \leq \frac{\lambda_H}{\lambda_L} c(\alpha(\theta))$ for any binary environment $\mathcal{E}_B \in \mathbb{E}_B$. Consider any $\theta, \hat{\theta} \in \Theta$ where $v_k(1, \hat{\theta}_k) < v_k(1, \theta_k)$ for some k and $v_j(1, \hat{\theta}_j) = v_j(1, \theta_j)$ for all $j \neq k$.

- 1. If $\sum_{i=1}^{n} \lambda_i v_i(1, \theta_i) < \sum_{i=1}^{n} \lambda_i c_i(1), \sum_{i=1}^{n} t_i(\theta) = \sum_{i=1}^{n} t_i(\hat{\theta}) = 0.$
- 2. If $\sum_{i=1}^{n} \lambda_i v_i(1, \theta_i) \ge \sum_{i=1}^{n} \lambda_i c_i(1)$ and $\sum_{i=1}^{n} \lambda_i v_i(1, \hat{\theta}_i) < \sum_{i=1}^{n} \lambda_i c_i(1), \sum_{i=1}^{n} t_i(\theta) \ge 0 = \sum_{i=1}^{n} t_i(\hat{\theta}).$
- 3. If $\sum_{i=1}^{n} \lambda_i v_i(1, \theta_i) \ge \sum_{i=1}^{n} \lambda_i c_i(1)$ and $\sum_{i=1}^{n} \lambda_i v_i(1, \hat{\theta}_i) \ge \sum_{i=1}^{n} \lambda_i c_i(1), \sum_{i=1}^{n} t_i(\hat{\theta}) \ge \sum_{i=1}^{n} t_i(\theta) \ge 0.$

To see (3), notice that k's generalized pivotal payment remains unchanged between θ and $\hat{\theta}$. Consider any $j \neq k$. If j is pivotal under θ_{-j} , then she remains pivotal under $\hat{\theta}_{-j}$ and her generalized pivotal payment strictly increases. If j is not pivotal under θ_{-j} and remains not pivotal under $\hat{\theta}_{-j}$, her generalized pivotal payment remains zero. If j is not pivotal under θ_{-j} , but is pivotal under $\hat{\theta}_{-j}$, her payment increases from zero to some positive amount.

Hence, to maximize $\sum_{i=1}^{n} t_i(\theta)$ with respect to θ , it must be that $\sum_{i=1}^{n} \lambda_i v_i(1, \theta_i) = \sum_{i=1}^{n} \lambda_i c_i(1)$ and the good is produced, i.e., every *i* is exactly pivotal. In this case, each *i*'s generalized pivotal payment is $v_i(1, \theta_i)$, and the total generalized pivotal payment is $\sum_{i=1}^{n} v_i(1, \theta_i)$.

We now show that $\sum_{i=1}^{n} v_i(1, \theta_i)$ is bounded from above by $\frac{\lambda_H}{\lambda_L}c$ under Assumption 2. For ease of notation, let $v_i \equiv v_i(1, \theta_i)$ and $c_i \equiv c_i(1)$. We seek

$$\max_{(\lambda_i, v_i, c_i)_{i \in I}} \sum_{i=1}^n v_i \quad \text{subject to} \quad \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \lambda_i c_i.$$

The solution involves placing all the value on individuals with the lowest λ_i (without loss of generality, a single individual k) and all the cost share on individuals with the highest λ_i (without loss of generality, a single individual l), so that $c_j = 0$ for all $j \neq l$, $c_l = c$, $v_j = 0$ for all $j \neq k$, and $\lambda_L v_k = \lambda_H c$. Hence,

$$\sum_{i=1}^{n} v_i \le \frac{\lambda_H}{\lambda_L} c.$$

Since the total generalized pivotal payment is zero when the good is not produced, the Step 1 result follows.

Step 2. We now show this holds for any finite environment $\mathcal{E}_F \in \mathbb{E}_F$. To show that the total generalized pivotal payment is non-negative, note that individual generalized pivotal payments are themselves non-negative. To show that the total generalized pivotal payment

is no more than $\frac{\lambda_H}{\lambda_L} c(\alpha(\theta))$, note that

$$\begin{split} \sum_{i=1}^{n} t_{i}(\theta) &\leq \sum_{i=1}^{n} t_{i}^{S}(\theta) \\ &= \sum_{i=1}^{n} t_{i}(\theta; \{0,1\}) + t_{i}(\theta; \{\hat{\alpha}(\theta; \{0,1\}), 2\}) \\ &+ t_{i}(\theta; \{\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{0,1\}), 2\}), 3\}) + \ldots + t_{i}(\theta; \{\hat{\alpha}(\theta; \ldots), \bar{y}\}) \\ &\leq \frac{\lambda_{H}}{\lambda_{L}} \left(c(\hat{\alpha}(\theta; \{0,1\})) - c(0) \right) \\ &+ \frac{\lambda_{H}}{\lambda_{L}} \left(c(\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{0,1\}), 2\})) - c(\hat{\alpha}(\theta; \{0,1\}))) \right) \\ &+ \frac{\lambda_{H}}{\lambda_{L}} \left(c(\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{0,1\}), 2\}), 3\})) - c(\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{0,1\}), 2\})) \right) \\ &+ \ldots + \frac{\lambda_{H}}{\lambda_{L}} \left(c(\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \ldots), \bar{y}\})) - c(\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \ldots), \bar{y} - 1\}))) \right) \\ &= \frac{\lambda_{H}}{\lambda_{L}} \left(c(\hat{\alpha}(\theta; \{\hat{\alpha}(\theta; \ldots), \bar{y}\})) - c(0) \right) \\ &= \frac{\lambda_{H}}{\lambda_{L}} c(\alpha(\theta)), \end{split}$$

where the first inequality follows by Lemma 1 and the second inequality follows by Step 1.

Part II. The fair share payment $c_i(\alpha(0, \theta_{-i}))$ is always non-negative. Hence, the total CSP payment is no less than zero.

The total generalized pivotal payment $\sum_{i=1}^{n} t_i(\theta)$ is no more than $\frac{\lambda_H}{\lambda_L} c(\alpha(\theta))$ by Part I. The total fair share payment $\sum_{i=1}^{n} c_i(\alpha(0, \theta_{-i}))$ can be no more than $c(\bar{y})$, since $\sum_{i=1}^{n} c_i(\alpha(0, \theta_{-i})) \leq \sum_{i=1}^{n} c_i(\bar{y}) = c(\bar{y})$, and with a monotone decision rule can be no more than $c(\alpha(\theta))$, since $\sum_{i=1}^{n} c_i(\alpha(0, \theta_{-i})) \leq \sum_{i=1}^{n} c_i(\alpha(0, \theta_{-i})) \leq \sum_{i=1}^{n} c_i(\alpha(\theta)) = c(\alpha(\theta))$. Hence, the total CSP payment can be no more than $\frac{\lambda_H + \lambda_L}{\lambda_L} c(\alpha(\theta)) + c(\bar{y})$ and with a monotone decision rule, no more than $\frac{\lambda_H + \lambda_L}{\lambda_L} c(\alpha(\theta))$.

Let $\mathbb{E}_B \subset \mathbb{E}_F$ denote the set of binary public good environments, i.e., the set of finite public good environments with $\bar{y} = 1$.

Lemma 2. Consider any sequence of i.i.d. random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}(\theta_i(1)^2) < \infty$, and consider any induced sequence of binary public good environments $\mathcal{E}_B^n \in \mathbb{E}_B^{\mathbb{N}}$ satisfying Assumption 2. The probability that there is at least one pivotal player in \mathcal{E}_B^n goes to zero as n goes to infinity.

Proof. Let $\theta_i \equiv \theta_i(1)$, $c_n \equiv c^n(1)$, and $c_i^n(z^n) \equiv c_i^n(1, z^n)$. θ_i is an i.i.d. sequence of random variables representing each *i*'s willingness to pay for the public good, c_n is an arbitrary

sequence of real numbers representing the total cost of the public good in an environment with population size n, and $(c_i^n(z^n))_{i=1}^n$ is a (non-i.i.d.) sequence of random vectors representing the cost share for each individual i in an environment with population size n given the profile of observable characteristics z^n . Let $\mu \equiv \mathbb{E}(\theta_i)$ be the mean willingness to pay for the good, $\sigma^2 \equiv \operatorname{Var}(\theta_i)$, and $\theta_n^* \equiv \max_{i \in \{1,...,n\}} \theta_i$. Then we have that

$$\mathbb{P}(k \text{ is pivotal in } \mathcal{E}_B^n) \le \mathbb{P}(\sum_{j \ne k, j \le n} \lambda_j(z_j) \theta_j \le \sum_{i=1}^n \lambda_i(z_i) c_i^n(z^n) \text{ and } \sum_{i=1}^n \lambda_i(z_i) \theta_i \ge \sum_{i=1}^n \lambda_i(z_i) c_i^n(z^n))$$

and

 $\mathbb{P}(\exists \text{ pivotal player in } \mathcal{E}_B^n)$

$$\begin{split} &\leq \mathbb{P}\Big(\exists k:\sum_{j\neq k,\,j\leq n}\lambda_j(z_j)\theta_j\leq \sum_{i=1}^n\lambda_i(z_i)c_i^n(z^n) \text{ and } \sum_{i=1}^n\lambda_i(z_i)\theta_i\geq \sum_{i=1}^n\lambda_i(z_i)c_i^n(z^n)\Big)\\ &=\mathbb{P}\Big(\exists k:0\leq \sum_{i=1}^n\lambda_i(z_i)c_i^n(z^n)-\sum_{j\neq k,\,j\leq n}\lambda_j(z_j)\theta_j \text{ and } \lambda_k(z_k)\theta_k\geq \sum_{i=1}^n\lambda_i(z_i)c_i^n(z^n)-\sum_{j\neq k,\,j\leq n}\lambda_j(z_j)\theta_j\Big)\\ &\leq \mathbb{P}\Big(\exists k:\lambda_k(z_k)\theta_k\geq \Big|\sum_{j\neq k,\,j\leq n}\lambda_j(z_j)\theta_j-\sum_{i=1}^n\lambda_i(z_i)c_i^n(z^n)\Big|\Big)\\ &\leq \mathbb{P}\Big(\exists k:2\lambda_k(z_k)\theta_k\geq \Big|\sum_{i=1}^n\lambda_i(z_i)\theta_i-\sum_{i=1}^n\lambda_i(z_i)c_i^n(z^n)\Big|\Big)\\ &\leq \mathbb{P}\Big(\exists k:2\lambda_H\theta_k\geq \lambda_L\Big|\sum_{i=1}^n\theta_i-\sum_{i=1}^nc_i^n(z^n)\Big|\Big)\\ &=\mathbb{P}\Big(\theta_n^*\geq \frac{\lambda_L}{2\lambda_H}\Big|\sum_{i=1}^n\theta_i-c_n\Big|\Big),\end{split}$$

where the third inequality follows by the triangle inequality and the fourth by Assumption 2.

We would now like to show that $\mathbb{P}\left(\theta_n^* \geq \frac{\lambda_L}{2\lambda_H} \left| \sum_{i=1}^n \theta_i - c_n \right| \right) \to 0$ as $n \to \infty$. For any $\varepsilon > 0$,

$$\begin{split} \mathbb{P}\Big(\theta_n^* \geq \frac{\lambda_L}{2\lambda_H} \Big| \sum_{i=1}^n \theta_i - c_n \Big| \Big) \\ &= \mathbb{P}\Big(\frac{\theta_n^*}{\sigma\sqrt{n}} \geq \frac{\lambda_L}{2\lambda_H} \Big| \frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} - \frac{c_n - \mu n}{\sigma\sqrt{n}} \Big| \Big) \\ &= \mathbb{P}\Big(\frac{\theta_n^*}{\sigma\sqrt{n}} \geq \frac{\lambda_L}{2\lambda_H} \Big| \frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} - \frac{c_n - \mu n}{\sigma\sqrt{n}} \Big|, \frac{\lambda_L}{2\lambda_H} \Big| \frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} - \frac{c_n - \mu n}{\sigma\sqrt{n}} \Big| > \frac{\varepsilon\lambda_L}{2\lambda_H} \Big) \\ &+ \mathbb{P}\Big(\frac{\theta_n^*}{\sigma\sqrt{n}} \geq \frac{\lambda_L}{2\lambda_H} \Big| \frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} - \frac{c_n - \mu n}{\sigma\sqrt{n}} \Big|, \frac{\lambda_L}{2\lambda_H} \Big| \frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} - \frac{c_n - \mu n}{\sigma\sqrt{n}} \Big| \le \frac{\varepsilon\lambda_L}{2\lambda_H} \Big) \\ &\leq \mathbb{P}\Big(\frac{\theta_n^*}{\sigma\sqrt{n}} \geq \frac{\varepsilon\lambda_L}{2\lambda_H}\Big) + \mathbb{P}\Big(\Big| \frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} - \frac{c_n - \mu n}{\sigma\sqrt{n}} \Big| \le \varepsilon\Big) \\ &= \mathbb{P}\Big(\frac{\theta_n^*}{\sigma\sqrt{n}} \geq \frac{\varepsilon\lambda_L}{2\lambda_H}\Big) + \mathbb{P}\Big(\frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} \le \frac{c_n - \mu n}{\sigma\sqrt{n}} + \varepsilon\Big) - \mathbb{P}\Big(\frac{\sum_{i=1}^n (\theta_i - \mu)}{\sigma\sqrt{n}} < \frac{c_n - \mu n}{\sigma\sqrt{n}} - \varepsilon\Big) \\ &\equiv \mathbb{P}\Big(\frac{\theta_n^*}{\sigma\sqrt{n}} \geq \frac{\varepsilon\lambda_L}{2\lambda_H}\Big) + F_n\Big(\frac{c_n - \mu n}{\sigma\sqrt{n}} + \varepsilon\Big) - F_n\Big(\frac{c_n - \mu n}{\sigma\sqrt{n}} - \varepsilon\Big), \end{split}$$

where F_n is the cdf of $\frac{\sum_{i=1}^{n}(\theta_i-\mu)}{\sigma\sqrt{n}}$. $\frac{\theta_n^*}{\sigma\sqrt{n}} \stackrel{p}{\to} 0$ by Rob (1982, pp. 211-212, proof of Lemma 1 and 2), so $\mathbb{P}(\frac{\theta_n^*}{\sigma\sqrt{n}} > \frac{\varepsilon\lambda_L}{2\lambda_H}) \to 0$ as $n \to \infty$. We would now like to show that $F_n(\frac{c_n-\mu n}{\sigma\sqrt{n}}+\varepsilon) - F_n(\frac{c_n-\mu n}{\sigma\sqrt{n}}-\varepsilon) \to 0$ as $n \to \infty$.

 $\frac{\sum_{i=1}^{n}(\theta_{i}-\mu)}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$ by the central limit theorem. By van der Vaart (1998, p. 12, Lemma 2.11), if a sequence of random variables X_{n} with cdf G_{n} converges in distribution to a random variable X with cdf G and G is continuous, then $G_{n} \to G$ uniformly. In particular, for any $\delta > 0$ there exists $N \in \mathbb{N}$ such that for all n > N and $x \in \mathbb{R}$, $|G_{n}(x) - G(x)| < \delta$.

Let Φ denote the cdf of a standard normal, which is continuous. Then $F_n \to \Phi$ uniformly and for any $x \in \mathbb{R}$ and n > N,

$$\left| \left(F_n(x+\varepsilon) - F_n(x-\varepsilon) \right) - \left(\Phi(x+\varepsilon) - \Phi(x-\varepsilon) \right) \right|$$

= $\left| \left(F_n(x+\varepsilon) - \Phi(x+\varepsilon) \right) + \left(\Phi(x-\varepsilon) - F_n(x-\varepsilon) \right) \right|$
 $\leq |F_n(x+\varepsilon) - \Phi(x+\varepsilon)| + |F_n(x-\delta) - \Phi(x-\delta)|$
 $< 2\delta.$

Note that $\max_x \Phi(x+\varepsilon) - \Phi(x-\varepsilon) = \Phi(\varepsilon) - \Phi(-\varepsilon)$. Hence, for any $\varepsilon, \delta > 0$, there exists N such that for all n > N,

$$0 \le F_n \left(\frac{c_n - \mu n}{\sigma \sqrt{n}} + \varepsilon\right) - F_n \left(\frac{c_n - \mu n}{\sigma \sqrt{n}} - \varepsilon\right) \le \Phi(\varepsilon) - \Phi(-\varepsilon) + 2\delta,$$

and, since $\Phi(\varepsilon) - \Phi(-\varepsilon) \to 0$ as $\varepsilon \to 0$, the result follows.

Proposition 4. Consider any sequence of i.i.d. random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}((\theta_i(y_1) - \theta_i(y_0))^2) < \infty$ for all $y_0, y_1 \in Y$, and consider any induced sequence of finite environments $\mathcal{E}_F^n \in \mathbb{E}_F^{\mathbb{N}}$ satisfying Assumption 2. The probability that there is at least one pivotal player in \mathcal{E}_F^n goes to zero as n goes to infinity.

Proof. For any $y_0, y_1 \in Y$ with $y_0 < y_1$, let $\hat{Y} = \{y_0, y_1\}$ and $\mathcal{E}_B^n(\hat{Y}) \in \mathbb{E}_B^n$ be the environment $\mathcal{E}_F^n \in \mathbb{E}_F^n$ with \hat{Y} substituted for Y. For any $i \in I$, let $\hat{\theta}_i \equiv \theta_i(y_1) - \theta_i(y_0)$ and $\hat{c}^n \equiv c^n(y_1) - c^n(y_0)$. Note that $(\hat{\theta}_i)_{i \in \mathbb{N}}$ is i.i.d. and $\mathbb{E}(\hat{\theta}_i^2) < \infty$, so by Lemma 2, $\mathbb{P}(\exists \text{ pivotal player in } \mathcal{E}_B^n(\hat{Y})) \to 0 \text{ as } n \to \infty$.

Let $\hat{\mathbb{Y}} = \{\{y_0, y_1\} : y_0, y_1 \in Y \text{ and } y_0 < y_1\}$. Note that $|\hat{\mathbb{Y}}| = K(K+1)/2$. Then

$$\mathbb{P}\left(\exists \text{ pivotal player in } \mathcal{E}_{F}^{n}\right) \leq \mathbb{P}\left(\exists \hat{Y} \in \hat{\mathbb{Y}} : \exists \text{ pivotal player in } \mathcal{E}_{B}^{n}(\hat{Y})\right)$$
$$\leq \sum_{\hat{Y} \in \hat{\mathbb{Y}}} \mathbb{P}\left(\exists \text{ pivotal player in } \mathcal{E}_{B}^{n}(\hat{Y})\right)$$
$$\to 0 \quad \text{as} \quad n \to \infty.$$

Theorem 4. Consider any sequence of *i.i.d.* random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}((\theta_i(y_1) - \theta_i(y_0))^2) < \infty$ for all $y_0, y_1 \in Y$, and consider any induced sequence of finite environments $\mathcal{E}_F^n \in \mathbb{E}_F^{\mathbb{N}}$ satisfying Assumption 2. Any sequence of cost-sharing pivotal mechanisms (α^n, τ^n) is asymptotically ex-post budget-balanced.

Proof. **Part I.** The probability that no agent is pivotal goes to one by Proposition 4, and the CSP is EPBB when no agent is pivotal.

Part II. Let $S^n(\theta^n, z^n) \equiv \sum_{i=1}^n \tau_i^n(\theta^n, z^n) - c^n(\alpha^n(\theta^n, z^n))$ be the budget surplus of the CSP mechanism (α^n, τ^n) . Let Π^n be the event in which there exists a pivotal player in \mathcal{E}_F^n .

Case 1. Suppose $\frac{c^n(\bar{y})}{n} \not\to \infty$ as $n \to \infty$. Then

$$\begin{split} \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right|) &= \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \Pi^{n}) \cdot \mathbb{P}(\Pi^{n}) + \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \neg \Pi^{n}) \cdot \mathbb{P}(\neg \Pi^{n}) \\ &\leq \frac{\lambda_{H}}{\lambda_{L}} \cdot \frac{c^{n}(\bar{y})}{n} \cdot \mathbb{P}(\Pi^{n}) \\ &\to 0 \quad \text{as} \quad n \to \infty, \end{split}$$

where the inequality follows by Proposition 3 (the distance from EPBB in a CSP is bounded by $\frac{\lambda_H}{\lambda_L}c(\bar{y})$ since $\frac{\lambda_H}{\lambda_L}c(\alpha(\theta)) + c(\bar{y}) - c(\alpha(\theta)) = \frac{\lambda_H - \lambda_L}{\lambda_L}c(\alpha(\theta)) + c(\bar{y}) \leq \frac{\lambda_H - \lambda_L}{\lambda_L}c(\bar{y}) + c(\bar{y}) = \frac{\lambda_H}{\lambda_L}c(\bar{y})$) and the last line follows by Proposition 4.

Case 2. Suppose $\frac{c^n(\bar{y})}{n} \to \infty$ as $n \to \infty$. Let y^* be the smallest y such that $\frac{c^n(y)}{n} \to \infty$ as $n \to \infty$. Since $c^n(y)$ is non-decreasing in y for all n, $\frac{c^n(y)}{n} \to \infty$ for any $y \ge y^*$.

If for all $y \geq \hat{y}$ the total welfare benefit is strictly smaller than the total welfare cost, $\sum_{i=1}^{n} \lambda_i(z_i)\theta_i(y) < \sum_{i=1}^{n} \lambda_i(z_i)c_i^n(y, z^n)$, then the total fair share payment $\sum_{i=1}^{n} c_i^n(\alpha(0, \theta_{-i}), z^n)$ can be no more than $c^n(\hat{y} - 1)$, since any $y \geq \hat{y}$ is never an efficient decision without i for any i. Hence, the total CSP payment can be no less than zero and no more than $\frac{\lambda_H}{\lambda_L}c^n(\alpha(\theta)) + c^n(\hat{y} - 1)$ by Proposition 3, and the distance from EPBB is bounded by $\frac{\lambda_H}{\lambda_L}c^n(\hat{y} - 1)$.

Let

$$\langle (y) \equiv \sum_{i=1}^{n} \lambda_i(z_i)\theta_i(y) < \sum_{i=1}^{n} \lambda_i(z_i)c_i^n(y, z^n) \quad \text{and} \quad \geq (y) \equiv \sum_{i=1}^{n} \lambda_i(z_i)\theta_i(y) \ge \sum_{i=1}^{n} \lambda_i(z_i)c_i^n(y, z^n)$$

be the event that the total welfare benefit is strictly smaller, or weakly larger, than the welfare cost of alternative y, respectively. Then

$$\begin{aligned} \frac{1}{n} \mathbb{E}(\left|S^n(\theta^n, z^n)\right|) &= \frac{1}{n} \mathbb{E}(\left|S^n(\theta^n, z^n)\right| \mid \forall y \ge y^*, \ <(y)) \mathbb{P}(\forall y \ge y^*, \ <(y)) \\ &+ \frac{1}{n} \mathbb{E}(\left|S^n(\theta^n, z^n)\right| \mid \exists y \ge y^*, \ \ge(y)) \mathbb{P}(\exists y \ge y^*, \ \ge(y)). \end{aligned}$$

We now show each term goes to zero. Let $\mu_y = \mathbb{E}(\theta_i(y))$ and $\sigma_y^2 = \operatorname{Var}(\theta_i(y))$. Then

$$\begin{split} &\frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \forall y \geq y^{*}, <(y)) \\ &= \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \Pi^{n} \text{ and } \forall y \geq y^{*}, <(y)) \cdot \mathbb{P}(\Pi^{n} \mid \forall y \geq y^{*}, <(y)) \\ &+ \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \neg \Pi^{n} \text{ and } \forall y \geq y^{*}, <(y)) \cdot \mathbb{P}(\neg \Pi^{n} \mid \forall y \geq y^{*}, <(y)) \\ &\leq \frac{\lambda_{H}}{\lambda_{L}} \cdot \frac{c^{n}(\hat{y} - 1)}{n} \cdot \frac{\mathbb{P}(\Pi^{n})}{\mathbb{P}(\forall y \geq y^{*}, <(y))} \\ &\to 0 \quad \text{as} \quad n \to \infty, \end{split}$$

where the inequality follows by Proposition 3 and the last line follows by Proposition 4 and

$$\begin{split} \mathbb{P}(\exists y \ge y^*, \ \ge(y)) &\leq \sum_{y=y^*}^{\bar{y}} \mathbb{P}\Big(\sum_{i=1}^n \lambda_i(z_i)\theta_i(y) \ge \sum_{i=1}^n \lambda_i(z_i)c_i^n(y, z^n)\Big) \\ &\leq \sum_{y=y^*}^{\bar{y}} \mathbb{P}\Big(\lambda_H \sum_{i=1}^n \theta_i(y) \ge \lambda_L \sum_{i=1}^n c_i^n(y, z^n)\Big) \\ &= \sum_{y=y^*}^{\bar{y}} \mathbb{P}\Big(\frac{1}{n} \sum_{i=1}^n \theta_i(y) \ge \frac{\lambda_L}{\lambda_H} \frac{c^n(y)}{n}\Big) \\ &\to 0 \quad \text{as} \quad n \to \infty, \end{split}$$

since for all $y \ge y^*$, $\frac{1}{n} \sum_{i=1}^n \theta_i(y) \xrightarrow{p} \mu_y$ by the law of large numbers and $\frac{c^n(y)}{n} \to \infty$ by assumption.

Finally,

$$\begin{split} &\frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \exists y \geq y^{*}, \geq (y))\mathbb{P}(\exists y \geq y^{*}, \geq (y)) \\ &= \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \geq (y^{*}), <(y^{*}+1), \ldots, <(\bar{y})) \\ &\times \mathbb{P}(\geq (y^{*}), <(y^{*}+1), \ldots, <(\bar{y})) \\ &+ \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \geq (y^{*}+1), <(y^{*}+2), \ldots, <(\bar{y})) \\ &\times \mathbb{P}(\geq (y^{*}+1), <(y^{*}+2), \ldots, <(\bar{y})) \\ &+ \ldots \\ &+ \frac{1}{n} \mathbb{E}(\left|S^{n}(\theta^{n}, z^{n})\right| \mid \geq (\bar{y})) \times \mathbb{P}(\geq (\bar{y})) \\ &\leq \frac{\lambda_{H}}{\lambda_{L}} \cdot \frac{c^{n}(y^{*})}{n} \cdot \mathbb{P}(\geq (y^{*})) + \frac{\lambda_{H}}{\lambda_{L}} \cdot \frac{c^{n}(y^{*}+1)}{n} \cdot \mathbb{P}(\geq (y^{*}+1)) + \ldots + \frac{\lambda_{H}}{\lambda_{L}} \cdot \frac{c^{n}(\bar{y})}{n} \cdot \mathbb{P}(\geq (\bar{y})). \end{split}$$

For any $y = y^*, \ldots, \bar{y}$, there exists $N \in \mathbb{N}$ such that for all n > N, $\frac{c^n(y)}{n} > \frac{\lambda_H}{\lambda_L} \mu_y$ and

$$\begin{aligned} \frac{c^n(y)}{n} \cdot \mathbb{P}(\geq(y)) &= \frac{c^n(y)}{n} \cdot \mathbb{P}\Big(\sum_{i=1}^n \lambda_i(z_i)\theta_i(y) \geq \sum_{i=1}^n \lambda_i(z_i)c_i^n(y, z^n)\Big) \\ &\leq \frac{c^n(y)}{n} \cdot \mathbb{P}\Big(\lambda_H \sum_{i=1}^n \theta_i(y) \geq \lambda_L \sum_{i=1}^n c_i^n(y, z^n)\Big) \\ &= \frac{c^n(y)}{n} \cdot \mathbb{P}\Big(\sum_{i=1}^n \theta_i(y) \geq \frac{\lambda_L}{\lambda_H}c^n(y)\Big) \\ &\leq \frac{c^n(y)}{n} \cdot \mathbb{P}\Big(\Big|\sum_{i=1}^n \theta_i(y) - n\mu_y\Big| \geq \Big|\frac{\lambda_L}{\lambda_H}c^n(y) - n\mu_y\Big|\Big) \\ &\leq \frac{c^n(y)}{n} \cdot \frac{n\sigma_y^2}{\Big(\frac{\lambda_L}{\lambda_H}c^n(y) - n\mu_y\Big)^2} \\ &= \frac{c^n(y)}{n} \cdot \frac{\sigma_y^2}{n\Big(\frac{\lambda_L}{\lambda_H}\frac{c^n(y)}{n} - \mu_y\Big)^2} \\ &\to 0 \quad \text{as} \quad n \to \infty, \end{aligned}$$

where the first inequality follows by Assumption 2, the second inequality by $\frac{c^n(y)}{n} > \frac{\lambda_H}{\lambda_L} \mu_y$, and the last inequality by Chebyshev's inequality.

Theorem 5. Consider any sequence of i.i.d. random vectors (θ_i, z_i) defined on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathbb{E}((\theta_i(y_1) - \theta_i(y_0))^2) < \infty$ for all $y_0, y_1 \in Y$, and consider any induced sequence of finite environments $\mathcal{E}_F^n \in \mathbb{E}_F^{\mathbb{N}}$ satisfying Assumption 2. Let α^n be a sequence of utilitarian and monotone decision rules and τ^n be a sequence of CSP transfer rules. A sequence of utilitarian and monotone mechanisms $(\alpha^n, \hat{\tau}^n)$ satisfies asymptotic ex-post budget-balance, strategy-proofness for each n, and cost-sharing universal participation for each n if and only if, for all i,

$$\hat{\tau}^n_i(\theta^n, z^n) = \tau^n_i(\theta^n, z^n) + h^n_i(\theta^n_{-i}, z^n),$$

for some sequence of functions $h_i^n: \Theta_0^{n-1} \times Z_0^n \to \mathbb{R}$ such that

- 1. $h_i^n(\theta_{-i}^n, z^n) \leq 0$ for all $\theta_{-i}^n \in \Theta_0^{n-1}$, $z^n \in Z_0^n$, *i*, and *n*,
- 2. $\lim_{n\to\infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \text{ for all } i) = 1, \text{ and }$
- 3. $\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(h_i^n(\theta_{-i}^n, z^n)) = 0.$

The same holds replacing cost-sharing universal participation with the fair pricing principle per unit. Under Assumption 1, the same holds replacing the fair pricing principle per unit with the fair pricing principle.

Proof. Part I. We would like to show that $(\alpha^n, \hat{\tau}^n)$ satisfies strategy-proofness and CS-UP (the FPP per unit, and the FPP under Assumption 1) for each n if and only if, for all i, $\hat{\tau}_i^n(\theta^n, z^n) = \tau_i^n(\theta^n, z^n) + h_i^n(\theta_{-i}^n, z^n)$ for some sequence of functions $h_i^n : \Theta_0^{n-1} \times Z_0^n \to \mathbb{R}$ such that (1). This follows immediately from Proposition 1 and Theorem 2 (Theorem 3). \Box

Part II. We would like to show that $(\alpha^n, \hat{\tau}^n)$ satisfies strategy-proofness, CS-UP (the FPP per unit, and the FPP under Assumption 1) for each n, and that the probability of EPBB goes to 0 as $n \to \infty$ if and only if, for all i, $\hat{\tau}_i^n(\theta^n, z^n) = \tau_i^n(\theta^n, z^n) + h_i^n(\theta_{-i}^n, z^n)$ for some sequence of functions $h_i^n : \Theta_0^{n-1} \times Z_0^n \to \mathbb{R}$ such that (1) and (2).

$$\mathbb{P}\Big(\sum_{i=1}^{n} \hat{\tau}_{i}^{n}(\theta^{n}, z^{n}) - c^{n}(\alpha^{n}(\theta^{n}, z^{n})) = 0\Big)$$

$$= \mathbb{P}\Big(\sum_{i=1}^{n} \hat{\tau}_{i}^{n}(\theta^{n}, z^{n}) - c^{n}(\alpha^{n}(\theta^{n}, z^{n}) = 0 \mid \neg \exists \text{ pivotal player in } \mathcal{E}_{F}^{n}\Big) \mathbb{P}(\neg \exists \text{ pivotal player in } \mathcal{E}_{F}^{n})$$

$$+ \mathbb{P}\Big(\sum_{i=1}^{n} \hat{\tau}_{i}^{n}(\theta^{n}, z^{n}) - c^{n}(\alpha^{n}(\theta^{n}, z^{n}) = 0 \mid \exists \text{ pivotal player in } \mathcal{E}_{F}^{n}\Big) \mathbb{P}(\exists \text{ pivotal player in } \mathcal{E}_{F}^{n})$$

$$= \mathbb{P}(\text{EDRP} \mid n_{2} \text{ on a pivotal}) \mathbb{P}(n_{2} \text{ on a pivotal}) + \mathbb{P}(\text{EDRP} \mid \text{ compare pivotal}) \mathbb{P}(\alpha \text{ proves privatel}) \mathbb{P}(\alpha \text{ proves proves privatel}) \mathbb{P}(\alpha \text{ proves proves privatel}) \mathbb{P}(\alpha \text{ proves p$$

 $\equiv \mathbb{P}(\text{EPBB} \mid \text{no one pivotal})\mathbb{P}(\text{no one pivotal}) + \mathbb{P}(\text{EPBB} \mid \text{someone pivotal})\mathbb{P}(\text{someone pivotal}).$ Hence,

$$\begin{split} \lim_{n \to \infty} \mathbb{P}(\text{EPBB}) &= \lim_{n \to \infty} \mathbb{P}(\text{EPBB} \mid \text{no one pivotal}) \lim_{n \to \infty} \mathbb{P}(\text{no one pivotal}) \\ &+ \lim_{n \to \infty} \mathbb{P}(\text{EPBB} \mid \text{someone pivotal}) \lim_{n \to \infty} \mathbb{P}(\text{someone pivotal}) \\ &= \lim_{n \to \infty} \mathbb{P}(\text{EPBB} \mid \text{no one pivotal}) \\ &= \lim_{n \to \infty} \mathbb{P}\Big(\sum_{i=1}^n h_i^n(\theta_{-i}^n, z^n) = 0 \mid \text{no one pivotal}\Big) \\ &= \lim_{n \to \infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \forall i \mid \text{no one pivotal}) \\ &= \lim_{n \to \infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \forall i), \end{split}$$

where the second line follows since $\lim_{n\to\infty} \mathbb{P}(\text{someone pivotal}) = 0$ (proof of Theorem 4 Part I), the third line follows since a CSP is EPBB when no one is pivotal, the fourth line follows from Part I, and the last line follows since

$$\begin{split} &\lim_{n \to \infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \ \forall i) \\ &= \lim_{n \to \infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \ \forall i \mid \text{no one pivotal}) \lim_{n \to \infty} \mathbb{P}(\text{no one pivotal}) \\ &+ \lim_{n \to \infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \ \forall i \mid \text{someone pivotal}) \lim_{n \to \infty} \mathbb{P}(\text{someone pivotal}) \\ &= \lim_{n \to \infty} \mathbb{P}(h_i^n(\theta_{-i}^n, z^n) = 0 \ \forall i \mid \text{no one pivotal}). \end{split}$$

Part III. We would like to show that $(\alpha^n, \hat{\tau}^n)$ is such that the expected distance from EPBB goes to 0 as $n \to \infty$ if and only if, for all i, $\hat{\tau}_i^n(\theta^n, z^n) = \tau_i^n(\theta^n, z^n) + \hat{h}_i^n(\theta^n, z^n)$ for some sequence of functions $\hat{h}_i^n : \Theta_0^n \times Z_0^n \to \mathbb{R}$ such that $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\hat{h}_i^n(\theta^n, z^n)) = 0$.

Any transfer rule can be represented by $\hat{\tau}_i^n(\theta^n, z^n) = \tau_i^n(\theta^n, z^n) + \hat{h}_i^n(\theta^n, z^n)$. Then

$$\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \Big(\Big| \sum_{i=1}^n \tau_i^n(\theta^n, z^n) + \hat{h}_i^n(\theta^n, z^n) - c^n(\alpha^n(\theta^n, z^n)) \Big| \Big) = 0$$
$$\iff \lim_{n \to \infty} \frac{1}{n} \mathbb{E} \Big(\Big| \sum_{i=1}^n \tau_i^n(\theta^n, z^n) - c^n(\alpha^n(\theta^n, z^n)) \Big| \Big) + \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E} \Big(\hat{h}_i^n(\theta^n, z^n) \Big) = 0$$

where the second line follows since $\lim_{x\to\infty} f(x) = 0 \iff \lim_{x\to\infty} |f(x)| = 0^{51}$ and the desired result follows by Theorem 4.

The result follows from the conjunction of Parts I, II, and III.

⁵¹Notice $|f(x) - 0| < \varepsilon \iff ||f(x)| - 0| < \varepsilon$.