Econometrica, Vol. 85, No. 3 (May, 2017), 735–767

STRONG DUALITY FOR A MULTIPLE-GOOD MONOPOLIST

BY CONSTANTINOS DASKALAKIS, ALAN DECKELBAUM, AND CHRISTOS TZAMOS[1]

We characterize optimal mechanisms for the multiple-good monopoly problem and provide a framework to find them. We show that a mechanism is optimal if and only if a measure μ derived from the buyer's type distribution satisfies certain stochastic dominance conditions. This measure expresses the marginal change in the seller's revenue under marginal changes in the rent paid to subsets of buyer types. As a corollary, we characterize the optimality of grand-bundling mechanisms, strengthening several results in the literature, where only sufficient optimality conditions have been derived. As an application, we show that the optimal mechanism for n independent uniform items each supported on $[c, c+1]$ is a grand-bundling mechanism, as long as c is sufficiently large, extending Pavlov's result for two items Pavlov (2011). At the same time, our characterization also implies that, for all c and for all sufficiently large n, the optimal mechanism for n independent uniform items supported on $[c, c+1]$ *is not* a grand-bundling mechanism.

KEYWORDS: Revenue maximization, mechanism design, strong duality, grand bundling.

1. INTRODUCTION

WE STUDY THE PROBLEM OF REVENUE maximization for a multiple-good monopolist. Given n heterogeneous goods and a probability distribution f over $\mathbb{R}^n_{\geq 0}$, we wish to design a mechanism that optimizes the monopolist's expected revenue against an additive (linear) buyer whose values for the goods are distributed according to f.

The single-good version of this problem—namely, $n = 1$—is well-understood, going back to Riley and Samuelson (1981), Myerson (1981), Maskin and Riley (1984), Riley and Zeckhauser (1983), where it was shown that a take-it-or-leave-it offer of the good at some price is optimal, and the optimal price can be easily calculated from f.

For general n, it has been known that the optimal mechanism may exhibit much richer structure. Even when the item values are independent, the mechanism may benefit from selling bundles of items or even lotteries over bundles of items (Preston McAfee, McMillan, and Whinston (1989), Bakos and Brynjolfsson (1999), Thanassoulis (2004), Manelli and Vincent (2006)). Moreover, no general framework to approach this problem has been proposed in the literature, making it dauntingly difficult both to identify optimal solutions and to certify the optimality of those solutions. As a consequence, seemingly simple special cases (even $n = 2$) remain poorly understood, despite much research for a few decades. See, for example, Rochet and Stole (2003) for a comprehensive survey of work spanning our problem, as well as Manelli and Vincent (2007) and Figalli, Kim, and McCann (2011) for additional references.

We propose a novel framework for revenue maximization based on duality theory. We identify a minimization problem that is dual to revenue maximization and prove that the optimal values of these problems are always equal. Our framework allows us to identify optimal mechanisms in general settings, and certify their optimality by providing a complementary solution to the dual problem, namely, finding a solution to the dual whose

[1]Daskalakis was supported by a Sloan Foundation Fellowship, a Microsoft Research Faculty Fellowship, and NSF Awards CCF-0953960 (CAREER) and CCF-1101491. Deckelbaum was supported by Fannie and John Hertz Foundation Daniel Stroock Fellowship and NSF Award CCF-1101491. Tzamos was supported by NSF Award CCF-1101491 and a Simons Award for Graduate Students in TCS.

 DOI: 10.3982/ECTA12618

objective value equals the mechanism's revenue. Our framework is applicable to arbitrary settings of n and f, with mild assumptions such as differentiability. In particular, we strengthen prior work Manelli and Vincent (2006), Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014), which identified optimal mechanisms in special cases. We exhibit the practicality of our framework by solving several examples. Importantly, we can leverage our duality theorem to characterize optimal multi-item mechanisms. From a technical standpoint, we provide new analytical methodology for multidimensional mechanism design by providing extensions to Monge–Kantorovich duality for optimal transportation. We proceed to discuss our contributions in detail, providing a roadmap to the paper, and conclude this section with a discussion of related work.

Strong Duality

Our first main result (presented as Theorem 2) formulates a dual problem to the optimal mechanism design problem, and establishes strong duality between the two problems. That is, we show that the optimal values of the two optimization problems are identical. Our approach for developing this dual problem is outlined below.

We start by formulating optimal mechanism design as a maximization problem over convex, nondecreasing, and 1-Lipschitz continuous functions u, representing the utility of the buyer as a function of her type, as in Rochet (1987). The objective function of this maximization problem can be written as the expectation of u with respect to a signed measure μ over the type space of the buyer. Measure μ is easily derived from the buyer's type distribution f (see Equation (3)) and expresses the marginal change in the seller's revenue under marginal changes in the rent paid to subsets of buyer types. Our formulation is summarized in Theorem 1, while Section 2.2 illustrates our formulation in the basic setting of independent uniform items.

In Theorem 2, we formulate a dual in the form of an optimal transportation problem, and establish strong duality between the two problems. Roughly speaking, our dual formulation is given the signed measure μ (from Theorem 1) and solves the following minimization problem: (i) first, it is allowed to choose any measure μ' that stochastically dominates μ with respect to convex increasing functions; (ii) second, it is supposed to find a coupling of the positive part μ'_+ of μ' with its negative part μ'_-, that is, find a transportation from μ'_+ to μ'_-; (iii) if a unit of mass of μ'_+ at x is transported to a unit of mass of μ'_- at y, we are charged $\|x-y\|_1$. The goal is to minimize the cost of the coupling with respect to the decisions in (i) and (ii).

While our dual formulation takes a simple form, establishing strong duality is quite technical. At a high level, our proof follows the proof of Monge–Kantorovich duality in Villani (2008), making use of the Fenchel–Rockafellar duality theorem, but the technical aspects of the proof are different due to the convexity constraint on feasible utility functions. The proof is presented in the Supplemental Material (Daskalakis, Deckelbaum, and Tzamos (2017)), but it is not necessary to understand the other results in this paper. We note that our formulation from Theorem 1 defines a convex optimization problem. One would hope then that infinite-dimensional linear programming techniques (Luenberger (1968), Anderson and Nash (1987)) can be leveraged to establish the existence of a strong dual. We are not aware of such an approach, and expect that such formulations will fail to establish existence of interior points in the primal feasible set, which is necessary for strong duality.

As already emphasized earlier, our identification of a strong dual implies that the optimal mechanism admits a certificate of optimality, in the form of a dual witness, for all settings of n and f. Hence, our duality framework can play the role of first-order conditions certifying the optimality of single-dimensional mechanisms. Where optimality of single-dimensional mechanisms can be certified by checking virtual welfare maximization, optimality of multidimensional mechanisms is always certifiable by providing dual solutions whose value matches the revenue of the mechanism, and such dual solutions take a simple form: they are transportation maps between measures.

Using our framework, we can provide shorter proofs of optimality of known mechanisms. As an illustrating example, we show in Section 5.1 how to use our framework to establish the optimality of the mechanism for two i.i.d. uniform [0, 1] items proposed by Manelli and Vincent (2006). Then in Section 5.2, we provide a simple illustration of the power of our framework, obtaining the optimal mechanism for two independent uniform [4, 16] and uniform [4, 7] items, a setting where the results of Manelli and Vincent (2006), Pavlov (2011), Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014) fail to apply. The optimal mechanism has the somewhat unusual structure shown in the diagram in Section 5, where types in Z are allocated nothing (and pay nothing), types in W are allocated the grand bundle (at price 12), while types in Y are allocated item 2 and get item 1 with probability 50% (at price 8).

Characterization of Optimal Mechanisms

Substantial effort in the literature has been devoted to studying optimality of mechanisms with a simple structure such as pricing mechanisms; see, for example, Manelli and Vincent (2006) and Daskalakis, Deckelbaum, and Tzamos (2013) for sufficient conditions under which mechanisms that only price the grand bundle of all items are optimal. Our second main result (presented as Theorem 3) obtains *necessary and sufficient* conditions characterizing the optimality of arbitrary mechanisms with a finite menu size. We proceed to describe our characterization result in more detail.

Suppose that we are given a feasible mechanism $\mathcal{M}$ whose set of possible allocations is finite. We can then partition the type set into finitely many subsets (called regions) $\mathcal{R}_1, \dots, \mathcal{R}_k$ of types who enjoy the same price and allocation. The question is this: *for what type distributions is $\mathcal{M}$ optimal?* Theorem 3 answers this question with a sharp characterization result: $\mathcal{M}$ is optimal *if and only if* the measure μ (derived from the type distribution as described above) satisfies k stochastic dominance conditions, one per region in the afore-defined partition. The type of stochastic dominance that μ restricted to region $\mathcal{R}_i$ ought to satisfy depends on the allocation to types from $\mathcal{R}_i$, namely, which set of items are allocated with probability 1, 0, or non-0/1.

Theorem 3 is important in that it reduces checking the optimality of mechanisms to checking standard stochastic dominance conditions between measures derived from the type distribution f, which is a concrete and easier task than arguing optimality against all possible mechanisms.

Theorem 3 is a corollary of our strong duality framework (Theorem 2), but requires a sequence of technical results. One direction of our characterization result requires turning the stochastic dominance conditions into dual solutions that can be plugged into Theorem 2 to establish the optimality of a given mechanism. The other direction requires showing that a dual solution certifying the optimality of a given mechanism also implies that the stochastic dominance conditions of Theorem 3 must hold.

A particularly simple special case of our characterization result pertains to the optimality of the grand-bundling mechanism. See Theorem 4. We show that the mechanism offering the grand bundle at price p is optimal *if and only if* measure μ satisfies a pair of stochastic dominance conditions. In particular, if Z are the types who cannot afford the grand bundle and W the types who can, then offering the grand bundle for p is optimal if and only if the following conditions hold:

- $\mu_-|_Z$, the negative part of μ restricted to Z, stochastically dominates $\mu_+|_Z$, the positive part of μ restricted to Z, with respect to all convex increasing functions;
- $\mu_+|_W$ stochastically dominates $\mu_-|_W$ with respect to all concave increasing functions.

Already our characterization of grand-bundling optimality settles a long line of research which only obtained sufficient conditions for the optimality of grand-bundling.

In turn, we illustrate the power of our characterization of grand-bundling optimality with Theorems 5 and 6, two results that are interesting in their own right. Theorem 5 generalizes the corresponding result of Pavlov (2011) from two to an arbitrary number of items. We show that, for any number of items n, there exists a large enough c such that the optimal mechanism for n i.i.d. uniform $[c, c+1]$ items is a grand-bundling mechanism. While maybe an intuitive claim, we do not see a direct way of proving it. Instead, we utilize Theorem 4 and construct intricate couplings establishing the stochastic dominance conditions required by the theorem. In view of Theorem 5, our companion theorem, Theorem 6, seems even more surprising. We show that in the same setting of n i.i.d. uniform $[c, c+1]$ items, for any fixed c it holds that, for all sufficiently large n, the optimal mechanism *is not* (!) a grand-bundling mechanism. See Section 6 for the proofs of these results.

Related Work

There is a rich literature on multi-item mechanism design pertaining to the multiple-good monopoly problem that we consider here. We refer the reader to the surveys Rochet and Stole (2003), Manelli and Vincent (2007), Figalli, Kim, and McCann (2011) for a detailed description, focusing on the work closest to ours.

Much work has focused on obtaining sufficient conditions for optimality of mechanisms. Hart and Nisan (2014), Menicucci, Hurkens, and Jeon (2015), and Haghpanah and Hartline (2015) provided sufficient conditions for the grand-bundling mechanism to be optimal. Manelli and Vincent (2006) provided conditions for the optimality of more complex deterministic mechanisms and, similarly, Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014) provided sufficient conditions for the optimality of general (possibly randomized) mechanisms. Finally, Haghpanah and Hartline (2015) provided an approach for reverse engineering sufficient conditions for a simple mechanism to be optimal. These works on sufficient conditions apply to limited settings of n and f. They typically proceed by relaxing some of the truthfulness constraints and are therefore only applicable when the relaxed constraints are not binding at the optimum.

In addition to sufficient conditions, a lot of work has focused on characterizing properties of optimal mechanisms. Armstrong (1996) have shown that optimal mechanisms always exclude a fraction of buyer types of low value from the mechanism. Thanassoulis (2004), Briest, Chawla, Kleinberg, and Weinberg (2010), and Hart and Nisan (2013) showed that randomization is necessary for optimal revenue extraction. In turn, Manelli and Vincent (2007) have shown that there exist type distributions for which optimal mechanisms are arbitrarily complex. Hart and Reny provided an interesting example where a product type distribution over two items stochastically dominates another, yet the optimal revenue from the weaker distribution is higher (Hart and Reny (2015)). Finally, some literature (Armstrong (1999), Hart and Nisan (2014), Babaioff, Immorlica, Lucier, and Weinberg (2014), Li and Yao (2013), Cai and Huang

(2013)) has focused on the revenue guarantees of simple mechanisms, for example, bundling all items together or selling them separately.

Rochet and Choné (1998) studied a closely related setting, providing a characterization of the optimal mechanism for the multiple-good monopoly problem where the monopolist has a (strictly) convex cost for producing copies of the goods. With strictly convex production costs, optimal mechanism design becomes a strictly concave maximization problem, which allows the use of first-order conditions to characterize optimal mechanisms. Our problem can be viewed as having a production cost that is 0 for selling at most one unit of each good and infinity otherwise. While still convex, our production function is not strictly convex and is discontinuous, making first-order conditions less useful for characterizing optimal mechanisms. This motivates the use of duality theory in our setting. From a technical standpoint, optimal mechanism design necessitates the development of new tools in optimal transport theory Villani (2008), extending Monge–Kantorovich duality to accommodate convexity constraints in the dual of the transportation problem. In our setting, the dual of the transportation problem corresponds to the mechanism design problem and these constraints correspond to the requirement that the utility function of the buyer be convex, which is intimately related to the truthfulness of the mechanism Rochet (1987). In turn, accommodating the convexity constraints in the mechanism design problem requires the introduction of mean-preserving spreads of measures in its transportation dual, resembling the "multidimensional sweeping" of Rochet and Choné.

Ultimately, our work relies on and develops further a fundamental connection of optimal transportation to designing optimal mechanisms. See Ekeland's notes on Optimal Transportation (Ekeland (2010)) for more connections to mechanism design.

2. REVENUE MAXIMIZATION AS OPTIMIZATION PROGRAM

2.1. *Setting up the Optimization Program*

Our goal is to find the revenue-optimal mechanism $\mathcal{M}$ for selling n goods to a single additive buyer. An additive buyer has a *type* x specifying his value for each good. The type x is an element of a *type space* $X = \prod_{i=1}^{n}[x_i^{\text{low}}, x_i^{\text{high}}]$, where $x_i^{\text{low}}, x_i^{\text{high}}$ are nonnegative real numbers. While the buyer knows his type with certainty, the mechanism designer only knows the probability distribution over X from which x is drawn. We assume that the distribution has a density $f : X \to \mathbb{R}$ that is continuous and differentiable with bounded derivatives.

Without loss of generality, by the revelation principle, we consider direct mechanisms. A (direct) mechanism consists of two functions: (i) an *allocation function* $\mathcal{P} : X \to [0, 1]^n$ specifying the probabilities, for each possible type declaration of the buyer, that the buyer will be allocated each good, and (ii) a *price function* $\mathcal{T} : X \to \mathbb{R}$ specifying, for each declared type of the buyer, the price that he is charged. When an additive buyer of type x declares himself to be of type $x' \in X$, he receives net expected utility $x \cdot \mathcal{P}(x') - \mathcal{T}(x')$.

We restrict our attention to mechanisms that are *incentive compatible*, meaning that the buyer must have adequate incentives to reveal his values for the items truthfully, and *individually rational*, meaning that the buyer has an incentive to participate in the mechanism.

DEFINITION 1: Mechanism $\mathcal{M} = (\mathcal{P}, \mathcal{T})$ over type space X is *incentive compatible (IC)* if and only if $x \cdot \mathcal{P}(x) - \mathcal{T}(x) \geq x \cdot \mathcal{P}(x') - \mathcal{T}(x')$ for all $x, x' \in X$.

DEFINITION 2: Mechanism $\mathcal{M} = (\mathcal{P}, \mathcal{T})$ over type space X is *individually rational (IR)* if and only if $x \cdot \mathcal{P}(x) - \mathcal{T}(x) \geq 0$ for all $x \in X$.

When a buyer truthfully reports his type to a mechanism $\mathcal{M} = (\mathcal{P}, \mathcal{T})$ (over type space X), we denote by $u : X \to \mathbb{R}$ the function that maps the buyer's valuation to the utility he receives by $\mathcal{M}$. It follows by the definitions of $\mathcal{P}$ and $\mathcal{T}$ that $u(x) = x \cdot \mathcal{P}(x) - \mathcal{T}(x)$. It is well-known (see Rochet (1987), Rochet and Choné (1998), and Manelli and Vincent (2006)) that an IC and IR mechanism has a convex, nonnegative, nondecreasing, and 1-Lipschitz utility function with respect to the ℓ_1 norm and that any utility function satisfying these properties is the utility function of an IC and IR mechanism with $\mathcal{P}(x) = \nabla u(x)$ and $\mathcal{T}(x) = \mathcal{P}(x) \cdot x - u(x)$.[2]

We clarify that a function u is 1-Lipschitz with respect to the ℓ_1 norm if $u(x) - u(y) \leq \|x - y\|_1$ for all $x, y \in X$. This is essentially equivalent to all partial derivatives having magnitude at most 1 in each dimension.

We will formulate the mechanism design problem as an optimization problem over feasible utility functions u. We first define the notation:

- $\mathcal{U}(X)$ is the set of all continuous, nondecreasing, and convex functions $u : X \to \mathbb{R}$.
- $\mathcal{L}_1(X)$ is the set of all 1-Lipschitz with respect to the ℓ_1 norm functions $u : X \to \mathbb{R}$.

In this notation, a mechanism $\mathcal{M}$ is IC and IR if and only if its utility function u satisfies $u \geq 0$ and $u \in \mathcal{U}(X) \cap \mathcal{L}_1(X)$. It follows that the optimal mechanism design problem can be viewed as an optimization problem:

$$\sup_{\substack{u \in \mathcal{U}(X) \cap \mathcal{L}_1(X) \\ u \geq 0}} \int_X \left[\nabla u(x) \cdot x - u(x)\right] f(x)\, dx.$$

Notice that for any utility u defining an IC and IR mechanism, the function $\tilde{u}(x) = u(x) - u(x^{\text{low}})$ also defines a valid IC and IR mechanism since $\tilde{u} \in \mathcal{U}(X) \cap \mathcal{L}_1(X)$ and $\tilde{u} \geq 0$. Moreover, $\tilde{u}$ achieves at least as much revenue as u, and thus it suffices in the above program to look only at feasible u with $u(x^{\text{low}}) = 0$.

We claim that we can therefore remove the constraint $u \geq 0$ and equivalently focus on solving

$$\sup_{u \in \mathcal{U}(X) \cap \mathcal{L}_1(X)} \int_X \left[\nabla u(x) \cdot x - \left(u(x) - u\left(x^{\text{low}}\right)\right)\right] f(x)\, dx. \tag{1}$$

Indeed, this objective function agrees with the prior one whenever $u(x^{\text{low}}) = 0$. Furthermore, for any $u \in \mathcal{U}(X) \cap \mathcal{L}_1(X)$, the function $\tilde{u}(x) = u(x) - u(x^{\text{low}})$ is nonnegative and achieves the same objective value. Applying the divergence theorem as in Manelli and Vincent (2006), we may rewrite the expression for expected revenue in (1) as follows:

$$\begin{aligned} &\int_X \left[\nabla u(x) \cdot x - \left(u(x) - u\left(x^{\text{low}}\right)\right)\right] f(x)\, dx \\ &\quad = \int_{\partial X} u(x) f(x)(x \cdot \hat{n})\, dx - \int_X u(x)\left(\nabla f(x) \cdot x + (n+1) f(x)\right) dx + u\left(x^{\text{low}}\right), \end{aligned} \tag{2}$$

[2]On the measure-0 set on which ∇u is not defined, we can use an analogous expression for $\mathcal{P}$ by choosing appropriate values of ∇u from the subgradient of u.

where $\hat{n}$ denotes the outer unit normal field to the boundary ∂X. To simplify notation, we make the following definition.

DEFINITION 3—Transformed Measure: The *transformed measure* of f is the (signed) measure μ (supported within X) given by the property that

$$\mu(A) \triangleq \int_{\partial X} \mathbb{I}_A(x) f(x)(x\cdot\hat{n})\,dx - \int_X \mathbb{I}_A(x)\big(\nabla f(x)\cdot x + (n+1)f(x)\big)\,dx + \mathbb{I}_A\big(x^{\mathrm{low}}\big) \quad (3)$$

for all measurable sets A.[3]

Interpretation of Transformed Measure

Given (2) and (3), the revenue of the seller in Formulation (1) can be written as $\int_X u\,d\mu$, which is a linear functional of u with respect to the measure μ. Hence, we will maintain the following intuition of what measure μ represents:

> "Measure μ quantifies the marginal change in revenue with respect to marginal changes in the rent paid to subsets of buyer types."

Moreover, our measure satisfies that $\mu(X) = \int_X 1\,d\mu = 0$. Indeed, if we substitute $u(x) = 1$ to the left-hand side of (2), we have that

$$\int_X \big[\nabla u(x)\cdot x - \big(u(x) - u\big(x^{\mathrm{low}}\big)\big)\big] f(x)\,dx = 0.$$

Furthermore, we have $|\mu|(X) < \infty$, since f, ∇f, and X are bounded.

Summarizing the above derivation, we obtain the following theorem.

THEOREM 1—Multi-Item Monopoly Problem: *The problem of determining the optimal IC and IR mechanism for a single additive buyer whose values for n goods are distributed according to the joint distribution* $f : X \to \mathbb{R}_{\geq 0}$ *is equivalent to solving the optimization problem*

$$\sup_{u\in\mathcal{U}(X)\cap\mathcal{L}_1(X)} \int_X u\,d\mu, \quad (4)$$

where μ *is the transformed measure of* f *given in* (3).

2.2. *Example*

Consider n independently distributed items, where the value of each item i is drawn uniformly from the bounded interval $[a_i, b_i]$ with $0 \leq a_i < b_i < \infty$. The support of the joint distribution is the set $X = \prod_i [a_i, b_i]$.

For notational convenience, define $v \triangleq \prod_i (b_i - a_i)$, the volume of X. The joint distribution of the items is given by the constant density function f taking value $1/v$ throughout X. The transformed measure μ of f is given by the relation

$$\mu(A) = \mathbb{I}_A(a_1,\dots,a_n) + \frac{1}{v}\int_{\partial X} \mathbb{I}_A(x)(x\cdot\hat{n})\,dx - \frac{n+1}{v}\int_X \mathbb{I}_A(x)\,dx$$

[3]It follows from boundedness of f's partial derivatives that μ is a Radon measure. Throughout this paper, all "measures" we use will be Radon measures.

for all measurable sets A. Therefore, by Theorem 1, the optimal revenue is equal to $\sup_{u\in\mathcal{U}(X)\cap\mathcal{L}_1(X)}\int_X u\,d\mu$, where μ is the sum of:

- a point mass of $+1$ at the point $(a_1,\dots,a_n)$.
- a mass of $-(n+1)$ distributed uniformly throughout the region X.
- a mass of $+\frac{b_i}{b_i-a_i}$ distributed uniformly on each surface $\{x\in\partial X : x_i=b_i\}$.
- a mass of $-\frac{a_i}{b_i-a_i}$ distributed uniformly on each surface $\{x\in\partial X : x_i=a_i\}$.

3. THE STRONG MECHANISM DESIGN DUALITY THEOREM

Thus far, we have compactly formulated the problem facing the multi-item monopolist as an optimization problem with respect to the buyer's utility function; see Formulation (4). Unfortunately, the problem is infinite-dimensional and cannot be solved directly. Moreover, the problem is not strictly convex so we cannot characterize its optimum using first-order conditions. This is an important point of departure in comparison with the work of Rochet and Choné (1998), where the strict convexity of the cost function allowed first-order conditions to drive the characterization. For more discussion, see Section 1.

In the absence of strict convexity, our approach is to use duality theory. We are seeking to identify a minimization problem, called "the dual problem," and which is linked to Formulation (4), henceforth called "the primal problem," as follows:

1. We want that the value of any solution to the dual problem is larger than the revenue achieved by any solution to the primal problem. If a minimization problem satisfies this property, it is called a "weak dual problem."

2. Additionally, we want that the optimum of the dual problem matches the optimum of the primal problem. A minimization problem satisfying this property is called a "strong dual problem." It is clear that a strong dual problem is also a weak dual problem. This type of strong dual problem is what we will identify in Theorem 2 of this section.

The importance of identifying a strong dual problem is the following. Given a candidate optimal mechanism, we are guaranteed that a solution to the dual problem with a matching objective value exists if and only if the candidate mechanism is indeed optimal. Therefore, solutions to the dual problem constitute "certificates of optimality" for solutions to the primal, and strong duality guarantees that such dual certificates are always possible to find for optimal solutions to the primal. Accordingly, we will be seeking solutions to our dual problem from Theorem 2 to obtain "certificates, or witnesses, of optimality" for candidate optimal mechanisms. By this we mean that we will be seeking solutions to the dual that prove (via duality theory) that a candidate optimal mechanism is indeed optimal. Moreover, these dual solutions take the form of optimal transportation maps between submeasures induced by measure μ of Definition 3. This tight connection between optimal mechanisms (primal solutions) and optimal transportation maps (dual solutions) drives our characterization of optimal mechanisms in Theorem 3, as well as the concrete examples we work out in Sections 5, 6.1, and 8. Moreover, by "reverse-engineering the duality theorem," we provide a framework for identifying optimal mechanisms in Section 7.

Recent work has applied duality theory to identify optimal mechanisms in the same setting as ours Manelli and Vincent (2006), Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014), albeit this work is restricted in that they only provided weak dual problems. These approaches remove constraints related to truthfulness from the primal formulation, and identify weak dual formulations to such relaxed primal formulations. As such, they provide no guarantee that they can identify dual certificates of optimality for optimal mechanisms. Indeed, while these techniques suffice in certain settings (namely, when the constraints removed from the primal happen not to be binding at the optimum), there are simple examples where they fail to apply. Section 5.2 provides such a two-item example with uniformly distributed values. In contrast to prior work, we achieve strong duality for the (unrelaxed) primal formulation and our approach is always guaranteed to work.

In this section, we show how to pin down the right dual formulation for the problem and prove strong duality. The proof of the result requires many analytical tools from measure theory. We give a rough sketch of the proof in this section and postpone the more technical details to the Supplemental Material.

3.1. *Measure-Theoretic Preliminaries*

We start with some useful measure-theoretic notation:

- $\Gamma(X)$ and $\Gamma_+(X)$ denote the sets of signed and unsigned (Radon) measures on X.
- Given an unsigned measure $\gamma \in \Gamma_+(X \times X)$, we denote by γ_1, γ_2 the two marginals of γ, that is, $\gamma_1(A) = \gamma(A \times X)$ and $\gamma_2(A) = \gamma(X \times A)$ for all measurable sets $A \subseteq X$.
- For a (signed) measure μ and a measurable $A \subseteq X$, we define the *restriction of* μ *to* A, denoted $\mu|_A$, by the property $\mu|_A(S) = \mu(A \cap S)$ for all measurable S.
- For a signed measure μ, we will denote by μ_+, μ_- the positive and negative parts of μ, respectively. That is, $\mu = \mu_+ - \mu_-$, where μ_+ and μ_- provide mass to disjoint subsets of X.

We will also be needing certain stochastic dominance properties, namely, first- and second-order stochastic dominance as well as the notion of convex dominance.

DEFINITION 4: We say that α *first-order* (respectively *second-order*) dominates β for $\alpha, \beta \in \Gamma(X)$, denoted $\alpha \succeq_1 \beta$ (respectively $\alpha \succeq_2 \beta$), if, for all nondecreasing continuous (respectively nondecreasing concave) functions $u : X \to \mathbb{R}$, $\int u\, d\alpha \geq \int u\, d\beta$.

Similarly, for vector random variables A and B with values in X, we say that $A \succeq_1 B$ (respectively $A \succeq_2 B$) if $\mathbb{E}[u(A)] \geq \mathbb{E}[u(B)]$ for all nondecreasing continuous (respectively nondecreasing concave) functions $u : X \to \mathbb{R}$.

DEFINITION 5: We say that α *convexly dominates* β for $\alpha, \beta \in \Gamma(X)$, denoted $\alpha \succeq_{\text{cvx}} \beta$, if, for all (nondecreasing, convex) functions $u \in \mathcal{U}(X)$, $\int u\, d\alpha \geq \int u\, d\beta$.

Similarly, for vector random variables A and B with values in X, we say that $A \succeq_{\text{cvx}} B$ if $\mathbb{E}[u(A)] \geq \mathbb{E}[u(B)]$ for all $u \in \mathcal{U}(X)$.

Interpretation of Convex Dominance

For intuition, a measure $\alpha \succeq_{\text{cvx}} \beta$ if we can transform β to α by doing the following two operations:

1. sending (positive) mass to coordinate-wise larger points: this makes the integral $\int u\, d\beta$ larger since u is nondecreasing.
2. spreading (positive) mass so that the mean is preserved: this makes the integral $\int u\, d\beta$ larger since u is convex.

The existence of a valid transformation using the above operations is equivalent to convex dominance. This follows by Strassen's theorem presented in the Supplemental Material.

3.2. *Mechanism Design Duality*

The main result of this paper is that the mechanism design problem, formulated as a maximization problem in Theorem 1, has a strong dual problem, as follows:

THEOREM 2—Strong Duality Theorem: *Let $\mu \in \Gamma(X)$ be the transformed measure of the probability density f according to Definition 3. Then*

$$\sup_{u \in \mathcal{U}(X) \cap \mathcal{L}_1(X)} \int_X u \, d\mu = \inf_{\substack{\gamma \in \Gamma_+(X \times X) \\ \gamma_1 - \gamma_2 \succeq_{\text{cvx}} \mu}} \int_{X \times X} \|x - y\|_1 \, d\gamma(x, y) \tag{5}$$

and both the supremum and infimum are achieved. Moreover, the infimum is achieved for some γ^* such that $\gamma_1^*(X) = \gamma_2^*(X) = \mu_+(X)$, $\gamma_1^* \succeq_{\text{cvx}} \mu_+$, and $\gamma_2^* \preceq_{\text{cvx}} \mu_-$.

Interpretation of the Strong Dual Problem

The dual problem of minimizing $\int \|x - y\|_1 \, d\gamma$ is an optimization problem that can be intuitively thought of as a two-step process:

Step 1: Transform μ into a new measure μ' with $\mu'(X) = 0$ such that $\mu' \succeq_{\text{cvx}} \mu$. This step is similar to sweeping as defined in Rochet and Choné (1998) where they transformed the original measure by mean-preserving spreads. However, here we are also allowed to perform positive mass transfers to coordinate-wise larger points.

Step 2: Find a joint measure $\gamma \in \Gamma_+(X \times X)$ with $\gamma_1 = \mu'_+$, $\gamma_2 = \mu'_-$ such that $\int \|x - y\|_1 \, d\gamma(x, y)$ is minimized. This is an optimal mass transportation problem where the cost of transporting a unit of mass from a point x to a point y is the ℓ_1 distance $\|x - y\|_1$, and we are asked for the cheapest method of transforming the positive part of μ' into the negative part of μ'. Transportation problems of this form have been studied in the mathematical literature. See Villani (2008).

Overall, our goal in the dual problem is to match the positive part of μ to the negative part of μ at a minimum cost where some operations come for free, namely, we can choose any $\mu' \succeq_{\text{cvx}} \mu$ that is convenient to us, foreseeing that transporting μ'_+ to μ'_- comes at a cost equal to the total ℓ_1 distance that mass travels.

We remark that establishing that the right-hand side of (5) is a weak dual for the left-hand side is easy. Proving strong duality is significantly more challenging, and relies on nontrivial analytical tools such as the Fenchel–Rockafellar duality theorem. We postpone that proof to the Supplemental Material, and proceed to show weak duality.

LEMMA 1—Weak Duality: *Let $\mu \in \Gamma(X)$. Then*

$$\sup_{u \in \mathcal{U}(X) \cap \mathcal{L}_1(X)} \int_X u \, d\mu \le \inf_{\substack{\gamma \in \Gamma_+(X \times X) \\ \gamma_1 - \gamma_2 \succeq_{\text{cvx}} \mu}} \int_{X \times X} \|x - y\|_1 \, d\gamma.$$

PROOF: For any feasible u for the left-hand side and feasible γ for the right-hand side, we have

$$\int_X u \, d\mu \le \int_X u \, d(\gamma_1 - \gamma_2) = \int_{X \times X} \big(u(x) - u(y)\big) \, d\gamma(x, y) \le \int_{X \times X} \|x - y\|_1 \, d\gamma(x, y),$$

where the first inequality follows from $\gamma_1 - \gamma_2 \succeq_{\text{cvx}} \mu$ and the second inequality follows from the 1-Lipschitz condition on u. *Q.E.D.*

From the proof of Lemma 1, we note the following "complementary slackness" conditions that a pair of optimal primal and dual solutions must satisfy.

COROLLARY 1: *Let u^* and γ^* be feasible for their respective problems above. Then $\int u^* \, d\mu = \int \|x - y\|_1 \, d\gamma^*$ if and only if both of the following conditions hold*:

1. $\int u^* \, d(\gamma_1^* - \gamma_2^*) = \int u^* \, d\mu$.
2. $u^*(x) - u^*(y) = \|x - y\|_1$, $\gamma^*(x, y)$-*almost surely*.

PROOF: The inequalities in the proof of Lemma 1 are tight precisely when both conditions hold. *Q.E.D.*

Interpretation of the Complementary Slackness Conditions

REMARK 1: It is useful to geometrically interpret Corollary 1:

CONDITION 1: We view $\gamma_1^* - \gamma_2^*$ (denote this by μ') as a "shuffled" μ. Stemming from the $\mu' \succeq_{\text{cvx}} \mu$ constraint, the shuffling of μ into μ' is obtained via any sequence of the following operations: (1) Picking a positive point mass δ_x from μ_+ and sending it from point x to some other point $y \geq x$ (coordinate-wise). The constraint $\int u^* \, d\mu' = \int u^* \, d\mu$ requires that $u^*(x) = u^*(y)$. Recall that u^* is nondecreasing, so $u^*(z) = u^*(x)$ for all $z \in \prod_j [x_j, y_j]$. Thus, if y is strictly larger than x in coordinate i, then $(\nabla u^*)_i = 0$ at all points z "in between" x and y. The other operation we are allowed, called a "mean-preserving spread," is (2) picking a positive point mass δ_x from μ_+, splitting the point mass into several pieces, and sending these pieces to multiple points while preserving the center of mass. The constraint $\int u^* \, d\mu' = \int u^* \, d\mu$ requires that u^* varies *linearly* between x and all points z that received a piece.

CONDITION 2: The second condition is more straightforward than the first. We view γ^* as a "transport" map between its component measures γ_1^* and γ_2^*. The condition states that if γ^* transports from location x to location y, then $u^*(x) = u^*(y) + \|x - y\|_1$. If, for some coordinate i, $x_i < y_i$, then $\|z - y\|_1 < \|x - y\|_1$ for z with $z_j = \max(x_j, y_j)$. This leads to a contradiction since $u^*(x) - u^*(y) \leq u^*(z) - u^*(y) \leq \|z - y\|_1 < \|x - y\|_1$. Therefore, it must be the case that (1) x is component-wise greater than or equal to y and (2) if $x_i > y_i$ in coordinate i, then $(\nabla u^*)_i = 1$ at all points "in between" x and y. That is, the mechanism allocates item i with probability 1 to all those types.

By Lemma 1 and Corollary 1, if we can find a "tight pair" of u^* and γ^*, then they are optimal for their respective problems. This is useful since constructing a γ that satisfies the conditions of Corollary 1 serves as a certificate of optimality for a mechanism. Theorem 2 shows that this approach always works: for any optimal u^*, there always exists a γ^* satisfying the conditions of Corollary 1.

REMARK 2: It is useful to discuss what in our dual formulation in the RHS of (5) makes it a strong dual, comparing to the previous work Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014). If we were to tighten the $\gamma_1 - \gamma_2 \succeq_{\text{cvx}} \mu$ constraint in our dual formulation to a first-order stochastic dominance constraint, we essentially recover the duality framework of Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014). Tightening the dual constraint maintains the weak duality but creates a gap between the optimal primal and dual values. In particular, the dual problem resulting from tightening this constraint becomes a strong dual problem for a relaxed version of the mechanism design problem in which the convexity constraint on u is dropped.

4. SINGLE-ITEM APPLICATIONS AND INTERPRETATION

Before considering multi-item settings, it is instructive to study the application of our strong duality theorem to single-item settings. We seek to relate the task of minimizing the transportation cost in the dual problem from Theorem 2 to the structure of Myerson's solution Myerson (1981).

Consider the task of selling a single item to a buyer whose value z for the item is distributed according to a twice-differentiable regular distribution F supported on $[\underline{z}, \bar{z}]$.[4] Since $n = 1$, if we were to apply our duality framework to this setting, we would choose μ according to (3) as follows:

$$\begin{aligned}\mu(A) &= \mathbb{I}_A(\underline{z}) \cdot \big(1 - f(\underline{z}) \cdot \underline{z}\big) + \mathbb{I}_A(\bar{z}) \cdot f(\bar{z}) \cdot \bar{z} - \int_{\underline{z}}^{\bar{z}} \mathbb{I}_A(z)\big(f'(z) \cdot z + 2f(z)\big)\,dz \\ &= \mathbb{I}_A(\underline{z}) \cdot \big(1 - f(\underline{z}) \cdot \underline{z}\big) + \mathbb{I}_A(\bar{z}) \cdot f(\bar{z}) \cdot \bar{z} - \int_{\underline{z}}^{\bar{z}} \mathbb{I}_A(z)\left(\left(z - \frac{1 - F(z)}{f(z)}\right) f(z)\right)' dz.\end{aligned}$$

We can interpret the transportation problem of Theorem 2, defined in terms of μ, as:

- The sub-population of buyers having the right-most type, $\bar{z}$, in the support of the distribution have an excess supply of $f(\bar{z}) \cdot \bar{z}$.
- The sub-population of buyers with the left-most type, $\underline{z}$, in the support have an excess supply of $1 - f(\underline{z}) \cdot \underline{z}$.
- Finally, the sub-population of buyers at each other type, z, have a demand of

$$\left(\left(z - \frac{1 - F(z)}{f(z)}\right) f(z)\right)' dz.$$

One way to satisfy the above supply/demand requirements is to have every infinitesimal buyer of type z push mass of $z - \frac{1-F(z)}{f(z)}$ to its left. Since the fraction of buyers at z is $f(z)$, the total amount of mass staying with them is then $((z - \frac{1-F(z)}{f(z)})f(z))'\,dz$ as required. Notice, in particular, that buyers with positive virtual types will push mass to their left, while buyers with negative virtual types will push mass to their right.

The afore-described transportation map is feasible for our transportation problem as it satisfies all demand/supply constraints. We also claim that this solution is optimal. To see this, consider the mechanism that allocates the item to all buyers with nonnegative virtual type at a fixed price p^*. The resulting utility function is of the form $\max\{z - p^*, 0\}$. We claim that this utility function satisfies the complementary slackness conditions of Remark 1 with respect to the transportation map identified above. Indeed, when $z > p^*$, u is linear with $u'(z) = 1$ and mass is sent to the left—which is allowed by Part 2 of the remark, while, when $z < p^*$, u is 0 with $u'(z) = 0$ and mass is sent to the right—allowed by Part 1(1) of the remark.

In conclusion, when F is regular, the virtual values dictate exactly how to optimally solve the optimal transportation problem of Theorem 2. Each infinitesimal buyer of type z will push mass that equals its virtual value to its left. In particular, the optimal transportation does not need to use mean-preserving spreads. Moreover, measure μ can be

[4] We remind the reader that a differentiable distribution F is *regular* when its Myerson virtual value function $\phi(z) = z - \frac{1-F(z)}{f(z)}$ is increasing in its support, where f is the distribution density function.

interpreted as the "negative marginal normalized virtual value," as it assigns measure $-((z - \frac{1-F(z)}{f(z)})f(z))' \, dz$ to the interval $[z, z + dz]$, when $z \neq \underline{z}, \bar{z}$.

When F is not regular, the afore-described transportation map is not optimal due to the non-monotonicity of the virtual values. In this case, we need to preprocess our measure μ via mean-preserving spreads, prior to the transport, and ironing dictates how to do these mean-preserving spreads. In other words, ironing dictates how to perform the sweeping of the type set prior to transport.

5. MULTI-ITEM APPLICATIONS OF DUALITY

We now give two examples of using Theorem 2 to prove optimality of mechanisms for selling two uniformly distributed independent items.

5.1. *Two Uniform* $[0, 1]$ *Items*

Using Theorem 2, we provide a short proof of optimality of the mechanism for two i.i.d. uniform $[0, 1]$ items proposed by Manelli and Vincent (2006), which we refer to as the MV-mechanism:

EXAMPLE 1: The optimal IC and IR mechanism for selling two items whose values are distributed uniformly and independently on the interval $[0, 1]$ is the following menu:

- buy any single item for a price of $\frac{2}{3}$; or
- buy both items for a price of $\frac{4-\sqrt{2}}{3}$.

Let Z be the set of types that receive no goods and pay 0 to the MV-mechanism. Also, let A, B be the set of types that receive only goods 1 and 2, respectively, and W be the set of types that receive both goods. The sets A, B, Z, W are illustrated in Figure 1 and separated by solid lines.

Let us now try to prove that the MV-mechanism is indeed optimal. As a first step, we need to compute the transformed measure μ of the uniform distribution on $[0, 1]^2$. We have already computed μ in Section 2.2. It has a point mass of $+1$ at $(0, 0)$, a mass of -3 distributed uniformly over $[0, 1]^2$, a mass of $+1$ distributed uniformly on the top boundary of $[0, 1]^2$, and a mass of $+1$ distributed uniformly on the right boundary. Notice that the total net mass is equal to 0 within each region Z, A, B, or W.

To prove optimality of the MV-mechanism, we will construct an optimal γ^* for the dual program of Theorem 2 to match the positive mass μ_+ to the negative μ_-. Our γ^* will be decomposed into $\gamma^* = \gamma^Z + \gamma^A + \gamma^B + \gamma^W$ and to ensure that $\gamma_1^* - \gamma_2^* \succeq_{\text{cvx}} \mu$, we will show that

$$\gamma_1^Z - \gamma_2^Z \succeq_{\text{cvx}} \mu|_Z; \qquad \gamma_1^A - \gamma_2^A \succeq_{\text{cvx}} \mu|_A; \qquad \gamma_1^B - \gamma_2^B \succeq_{\text{cvx}} \mu|_B; \qquad \gamma_1^W - \gamma_2^W \succeq_{\text{cvx}} \mu|_W.$$

We will also show that the conditions of Corollary 1 hold for each of the measures γ^Z, γ^A, γ^B, and γ^W separately, namely, $\int u^* \, d(\gamma_1^S - \gamma_2^S) = \int_S u^* \, d\mu$ and $u^*(x) - u^*(y) = \|x - y\|_1$ hold γ^S-almost surely for $S = Z$, A, B, and W.

Construction of γ^Z: Since $\mu_+|_Z$ is a point mass at $(0, 0)$ and $\mu_-|_Z$ is distributed throughout a region which is coordinate-wise greater than $(0, 0)$, we notice that $\mu|_Z \preceq_{\text{cvx}} 0$. We set γ^Z to be the zero measure, and the relation $\gamma_1^Z - \gamma_2^Z = 0 \succeq_{\text{cvx}} \mu|_Z$, as well as the two necessary equalities from Corollary 1, are trivially satisfied.

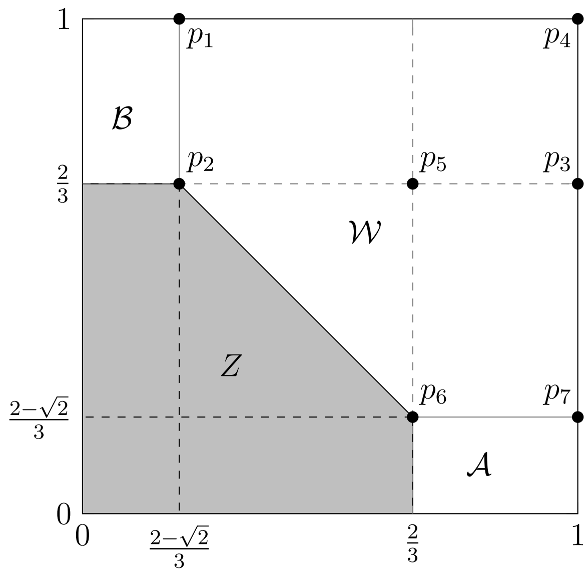

FIGURE 1.—The MV-mechanism for two i.i.d. uniform [0, 1] items.

Construction of γ^A *and* γ^B: In region A, $\mu_+|_A$ is distributed on the right boundary while $\mu_-|_A$ is distributed uniformly on the interior of A. We construct γ^A by transporting the positive mass $\mu_+|_A$ to the left to match the negative mass $\mu_-|_A$. Notice that this indeed matches completely the positive mass to the negative since $\mu(A) = 0$ and intuitively minimizes the ℓ_1 transportation distance. To see that the two necessary equalities from Corollary 1 are satisfied, notice that $\gamma_1^A = \mu_+|_A$, $\gamma_2^A = \mu_-|_A$ so the first equality holds. The second inequality holds as we are transporting mass only to the left and thus the measure γ^A is concentrated on pairs $(x, y) \in A \times A$ such that $1 = x_1 \geq y_1 \geq \frac{2}{3}$ and $x_2 = y_2$. Moreover, for all such pairs (x, y), we have that $u(x) - u(y) = (x_1 - \frac{2}{3}) - (y_1 - \frac{2}{3}) = x_1 - y_1 = \|x - y\|_1$. The construction of γ^B is similar.

Construction of γ^W: We construct an explicit matching that only matches leftwards and downwards without doing any prior mass shuffling. We match the positive mass on the segment $p_1 p_4$ to the negative mass on the rectangle $p_1 p_2 p_3 p_4$ by moving mass downwards. We match the positive mass of the segment $p_3 p_7$ to the negative mass on the rectangle $p_3 p_5 p_6 p_7$ by moving mass leftwards. Finally, we match the positive mass on the segment $p_3 p_4$ to the negative mass on the triangle $p_2 p_5 p_6$ by moving mass downwards and leftwards. Notice that all positive/negative mass in region W has been accounted for, all of $(\mu|_W)_+$ has been matched to all of $(\mu|_W)_-$, and all moves were down and to the left, establishing $u(x) - u(y) = (x_1 + x_2 - \frac{4-\sqrt{2}}{3}) - (y_1 + y_2 - \frac{4-\sqrt{2}}{3}) = x_1 + x_2 - y_1 - y_2 = \|x - y\|_1$.

5.2. *Two Uniform but not Identical Items*

We now present an example with two items whose values are distributed uniformly and independently on the intervals [4, 16] and [4, 7]. We note that the distributions are not identical, and thus the characterization of Pavlov (2011) does not apply. In addition, the relaxation-based duality framework of Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014) (see Remark 2) fails in this example: if we were

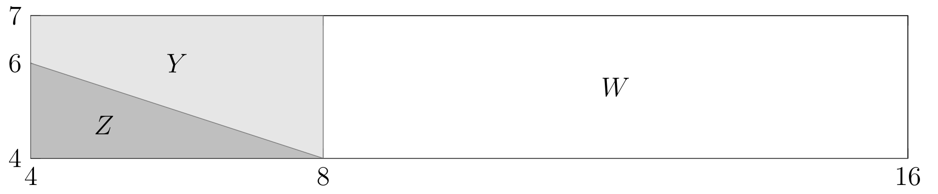

FIGURE 2.—Partition of $[4, 16] \times [4, 7]$ into different regions by the optimal mechanism.

to relax the constraint that the utility function u be convex, the "mechanism design program" would have a solution with greater revenue than is actually possible.

EXAMPLE 2: The optimal IC and IR mechanism for selling two items whose values are distributed uniformly and independently on the intervals $[4, 16]$ and $[4, 7]$ partitions the buyer types into 3 different regions as in Figure 2 and allocates the goods as follows:

- If the buyer's declared type is in region Z, he receives no goods and pays nothing.
- If the buyer's declared type is in region Y, he pays a price of 8 and receives the first good with probability 50% and the second good with probability 1.
- If the buyer's declared type is in region W, he gets both goods for a price of 12.

The proof of optimality of our proposed mechanism works by constructing a measure $\gamma = \gamma^Z + \gamma^Y + \gamma^W$ separately in each region. The constructions of γ^W and γ^Z are similar to the previous example. The construction of γ^Y, however, is a little more intricate as it requires an initial shuffling of the mass before computing the optimal way to transport the resulting mass. The proof is presented in the Supplemental Material.

5.3. *Discussion*

Our examples in this section serve to illustrate how to use our duality theorem to verify the optimality of our proposed mechanisms, without explaining how we identified these mechanisms. These mechanisms were in fact identified by "reverse-engineering" the duality theorem. The next two sections provide tools for performing this reverse-engineering. In particular, Section 6 provides a characterization of mechanism optimality in terms of stochastic dominance conditions satisfied in regions partitioning the type space. Alleviating the need to reverse-engineer the duality theorem, Section 7 prescribes a straightforward procedure for identifying optimal mechanisms. We use this procedure to solve several examples in Section 8.

6. CHARACTERIZING OPTIMAL FINITE-MENU MECHANISMS

To prove the optimality of our mechanisms in the examples of Section 5, we explicitly constructed a measure γ *separately* for each subset of types enjoying the same allocation in the optimal mechanism, establishing that the conditions of Corollary 1 are satisfied for each such subset of types separately. In this section, we show that decomposing the solution γ of the optimal transportation dual of Theorem 2 into "regions" of types enjoying the same allocation in the optimal solution u of the primal, and working on these regions separately to establish the complementary slackness conditions of Corollary 1, is guaranteed to work.

Even with this understanding of the structure of dual witnesses, it may still be nontrivial work to identify a witness certifying the optimality of a given mechanism. We thus develop a more usable framework for certifying the optimality of mechanisms, which does not involve finding dual witnesses at all. In particular, we show in Theorem 3 that a given mechanism $\mathcal{M}$ is optimal for some f *if and only if* appropriate stochastic dominance conditions are satisfied by the restriction of the transformed measure μ of Definition 3 to each region of types enjoying the same allocation under $\mathcal{M}$. We thus provide conditions that are both *necessary* and *sufficient* for a given mechanism $\mathcal{M}$ to be optimal, a characterization result.

To describe our characterization, we define the intuitive notion a "menu" that a certain mechanism offers.

DEFINITION 6: The *menu of a mechanism* $\mathcal{M} = (\mathcal{P}, \mathcal{T})$ is the set

$$\text{Menu}_{\mathcal{M}} = \{(p, t) : \exists x \in X, (p, t) = (\mathcal{P}(x), \mathcal{T}(x))\}.$$

Clearly, an IC mechanism allocates to every type x the option in the menu that maximizes that type's utility. Figure 3 shows an example of a menu and the corresponding partition of the type set into subsets of types that prefer each option in the menu.

The revenue of a mechanism with a finite menu-size comes from choices in the menu that are bought with strictly positive probability. The menu might contain options that

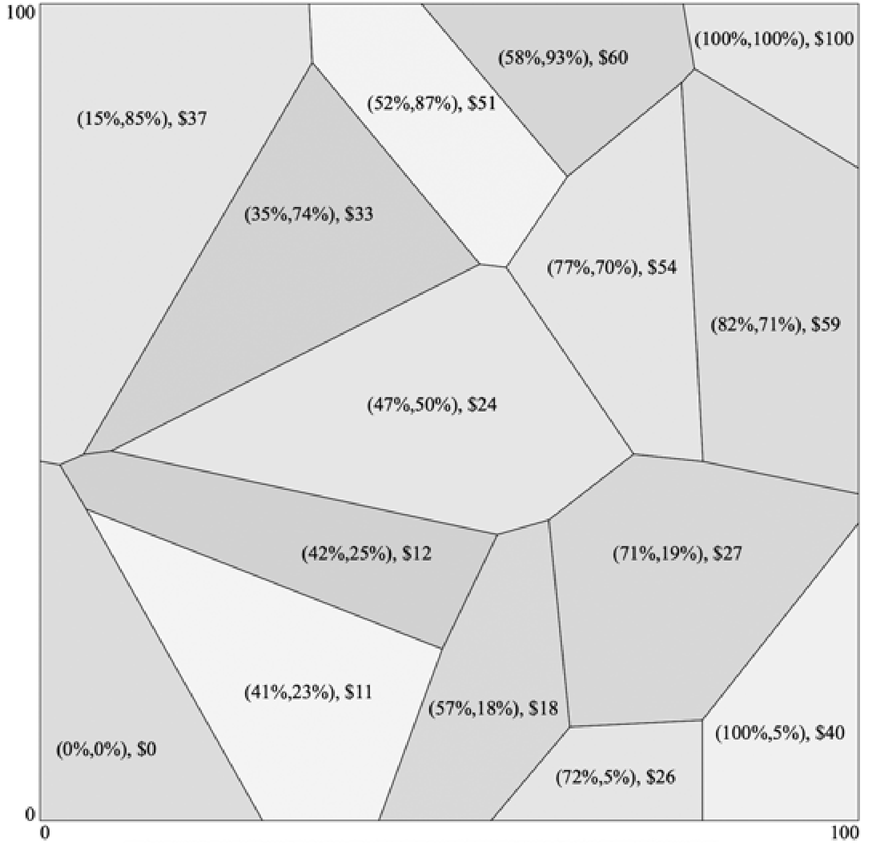

FIGURE 3.—Partition of the type set $X = [0, 100]^2$ induced by some menu of lotteries.

are only bought with probability 0, but we can get another mechanism that gives identical revenue by removing all those options. We call this the *essential form* of a mechanism.

DEFINITION 7: A mechanism $\mathcal{M}$ is in *essential form* if, for all options $(p, t) \in \text{Menu}_{\mathcal{M}}$, $\Pr_f[\{x \in X : (p, t) = (\mathcal{P}(x), \mathcal{T}(x))\}] > 0$.

We will now show our main result of this section under the assumption that the menu-size is finite. We expect that our tools can be used to extend the results to the case of infinite menu-size with a more careful analysis. We stress that the point of our result is not to provide sufficient conditions to certify optimality of mechanisms, as in Manelli and Vincent (2006), Daskalakis, Deckelbaum, and Tzamos (2013), Giannakopoulos and Koutsoupias (2014), but to provide necessary and sufficient conditions. In particular, we show that verifying optimality is *equivalent* to checking a collection of measure-theoretic inequalities, and this applies to arbitrary mechanisms with a finite menu-size. The proof of our result is intricate, requiring several technical lemmas, so it is postponed to the Supplemental Material. The most crucial component of the proof establishes that the optimal dual solution γ in Theorem 2 never convexly shuffles mass across regions of types that enjoy different allocations. (I.e., to obtain $\mu' = \gamma_1 - \gamma_2$ from μ, we never need to move mass across different regions.) Similarly, we argue that the optimal γ never transports mass across regions.

Before formally stating our result, it is helpful to provide some intuition behind it. Consider a region R corresponding to a menu choice $(\vec{p}, t)$ of an optimal mechanism $\mathcal{M}$. As we have already discussed, we can establish that the dual witness γ, which witnesses the optimality of $\mathcal{M}$, does not transport mass between regions and, likewise, the associated "convex shuffling" transforming μ to $\mu' = \gamma_1 - \gamma_2$ does not shuffle across regions. Given this, our complementary slackness conditions of Corollary 1 imply then that $\mu_+|_R$ can be transformed to $\mu_-|_R$ using the following (intra-region R) operations:

- spreading positive mass within R so that the mean is preserved;
- sending (positive) mass from a point $x \in R$ to a coordinate-wise larger point $y \in R$ if, for all coordinates where $y_i > x_i$, we have that the corresponding probability of the menu choice satisfies $p_i = 0$;
- sending (positive) mass from a point $x \in R$ to a coordinate-wise smaller point $y \in R$ if, for all coordinates where $y_i < x_i$, we have that $p_i = 1$.

Our characterization result involves stochastic dominance conditions that are slightly more general than the standard notions of first, second, and convex dominance. We need the following definition, which extends the notion of convex dominance.

DEFINITION 8: We say that a function $u : X \to \mathbb{R}$ is $\vec{v}$-monotone for a vector $\vec{v} \in \{-1, 0, +1\}^n$ if it is nondecreasing in all coordinates i for which $v_i = 1$ and non-increasing in all coordinates i for which $v_i = -1$.

A measure α *convexly dominates* a measure β *with respect to* a vector $\vec{v} \in \{-1, 0, +1\}^n$, denoted $\alpha \succeq_{\text{cvx}(\vec{v})} \beta$, if, for all convex $\vec{v}$-monotone functions $u \in \mathcal{U}(X)$,

$$\int u \, d\alpha \geq \int u \, d\beta.$$

Similarly, for vector random variables A and B with values in X, we say that $A \succeq_{\text{cvx}(\vec{v})} B$ if $\mathbb{E}[u(A)] \geq \mathbb{E}[u(B)]$ for all convex $\vec{v}$-monotone functions $u \in \mathcal{U}(X)$.

The definition of convex dominance presented earlier coincides with convex dominance with respect to the vector $\vec{1}$. Moreover, convex dominance with respect to the vector $-\vec{1}$ is related to second-order stochastic dominance as follows:

$$\alpha \succeq_{\mathrm{cvx}(-\vec{1})} \beta \quad \Leftrightarrow \quad \beta \succeq_2 \alpha.$$

Measures satisfying the dominance condition of Definition 8 must have equal mass.

PROPOSITION 1: *Fix two measures $\alpha, \beta \in \Gamma(X)$ and a vector $v \in \{-1, 0, 1\}^n$. If it holds that $\alpha \succeq_{\mathrm{cvx}(\vec{v})} \beta$, then $\alpha(X) = \beta(X)$.*

We are now ready to describe our main characterization theorem. Our characterization, stated below as Theorem 3 and proven in the Supplemental Material, is given in terms of the conditions of Definition 9.

DEFINITION 9—Optimal Menu Conditions: A mechanism $\mathcal{M}$ *satisfies the optimal menu conditions with respect to* μ if, for all menu choices $(p, t) \in \mathrm{Menu}_{\mathcal{M}}$, we have

$$\mu_+|_R \preceq_{\mathrm{cvx}(\vec{v})} \mu_-|_R,$$

where $R = \{x \in X : (\mathcal{P}(x), \mathcal{T}(x)) = (p, t)\}$ is the subset of types that receive (p, t) and $\vec{v}$ is the vector whose ith coordinate v_i takes value 1 if $p_i = 0$, value -1 if $p_i = 1$, or value 0 if $p_i \in (0, 1)$.

THEOREM 3—Optimal Menu Theorem: *Let μ be the transformed measure of a probability density f as per Definition 3. Then a mechanism $\mathcal{M}$ with finite menu-size is an optimal IC and IR mechanism for a single additive buyer whose values for n goods are distributed according to the joint distribution f* if and only if *its essential form satisfies the optimal menu conditions with respect to μ.*

Interpretation of the Optimal Menu Conditions

A simple interpretation of the optimal menu conditions that Theorem 3 claims are necessary and sufficient for the optimality of mechanisms is this. Take some region R of the type set X corresponding to the types that are allocated a specific menu choice (p, t) by optimal mechanism $\mathcal{M}$. Let us consider the revenue $\int_R u^* \, d\mu$ extracted by $\mathcal{M}$ from the types in region R. Is it possible to extract more revenue from these types? We claim that the optimal menu condition for region R guarantees that no mechanism can possibly extract more from the types in region R. Indeed, consider any utility function u induced by some other mechanism. The revenue extracted by this other mechanism in region R is $\int_R u \, d\mu = \int_R u^* \, d\mu + \int_R (u - u^*) \, d\mu \leq \int_R u^* \, d\mu$. That $\int_R (u - u^*) \, d\mu \leq 0$ follows directly from the optimal menu condition for region R. Indeed, since $u^*(x) = p \cdot x - t$ in region R, it follows that, whatever choice of u we make, $u - u^*$ is a convex $\vec{v}$-monotone function in region R, where $\vec{v}$ is the vector defined by p as per Definition 9. Our condition in region R reads $\mu|_R \preceq_{\mathrm{cvx}(\vec{v})} 0$, hence $\int_R (u - u^*) \, d\mu \leq 0$. Our line of argument implies the sufficiency of the optimal menu conditions, as they imply that, for each region separately, no mechanism can beat the revenue extracted by $\mathcal{M}$. The more surprising part (and harder to prove) is that the conditions are also necessary, implying that optimal mechanisms are locally optimal for every region R of types that they allocate the same menu choice to.

A particularly simple special case of our characterization result pertains to the optimality of the grand-bundling mechanism. Theorem 3 implies that the mechanism that offers the grand bundle at price p is optimal *if and only if* the transformed measure μ satisfies a pair of stochastic dominance conditions. In particular, we obtain the following theorem:

THEOREM 4—Grand-Bundling Optimality: *For a single additive buyer whose values for n goods are distributed according to the joint distribution f, the mechanism that only offers the bundle of all items at price p is optimal if and only if the transformed measure μ of f satisfies $\mu|_{\mathcal{W}} \succeq_2 0 \succeq_{\text{cvx}} \mu|_{\mathcal{Z}}$, where $\mathcal{W}$ is the subset of types that can afford the grand bundle at price p, and $\mathcal{Z}$ the subset of types who cannot.*

Next, we explore implications of our characterization of grand-bundling optimality.

6.1. *Example Applications of Grand-Bundling Optimality*

We now present an example application of our characterization result to determine the optimality of mechanisms that make a take-it-or-leave-it offer of the grand bundle of all items at some price. Our result applies to a setting with arbitrarily many items, which is relatively rare in the literature. More specifically, we consider a setting with n i.i.d. goods whose values are uniformly distributed on $[c, c+1]$. It is easy to see that the ratio of the revenue achievable by grand bundling to the social welfare goes to 1 when either n or c goes to infinity.[5] This implies that grand bundling is optimal or close to optimal for large values of n and c. Indeed, the following theorem shows that, for every n, grand bundling is the optimal mechanism for large values of c.

THEOREM 5: *For any integer $n > 0$, there exists a c_0 such that, for all $c \geq c_0$, the optimal mechanism for selling n i.i.d. goods whose values are uniform on $[c, c+1]$ is a take-it-or-leave-it offer for the grand bundle.*

REMARK 3: Pavlov (2011) proved the above result for two items, and explicitly solved for $c_0 \approx 0.077$. In our proof, for simplicity of analysis, we do not attempt to exactly compute c_0 as a function of n.

Our proof of Theorem 5 uses the following lemma, which enables us to appropriately match regions on the surface of a hypercube. The proofs of this lemma and of Theorem 5 appear in the Supplemental Material.

LEMMA 2: *For $n \geq 2$ and $\rho > 1$, define the $(n-1)$-dimensional subsets of* $[0,1]^n$:

$$A = \left\{ x : 1 = x_1 \geq x_2 \geq \cdots \geq x_n \text{ and } x_n \leq 1 - \left(\frac{\rho - 1}{\rho}\right)^{1/(n-1)} \right\},$$

$$B = \{ y : y_1 \geq \cdots \geq y_n = 0 \}.$$

There exists a continuous bijective map $\varphi : A \to B$ such that

- *For all $x \in A$, x is component-wise greater than or equal to $\varphi(x)$.*

[5]This follows by setting a price for the grand bundle equal to $(c + \frac{1}{2})n - \sqrt{n} \log cn$ and noting that a straightforward application of Hoeffding's inequality gives that the bundle is accepted with probability close to 1.

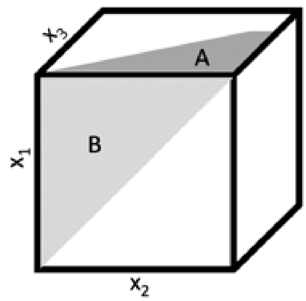

FIGURE 4.—The regions of Lemma 2 for the case $n = 3$.

- *For subsets $S \subseteq A$ which are measurable under the $(n-1)$-dimensional surface Lebesgue measure $v(\cdot)$, it holds that $\rho \cdot v(S) = v(\varphi(S))$.*
- *For all $\varepsilon > 0$, if $\varphi_1(x) \le \varepsilon$, then $x_n \ge 1 - (\frac{\varepsilon^{n-1}+\rho-1}{\rho})^{1/(n-1)}$.*

The main difficulty in proving Theorem 5 is verifying the necessary stochastic dominance relations above the grand-bundling hyperplane. Our proof appropriately partitions this part of the hypercube into $2(n! + 1)$ regions and uses Lemma 2 to show a desired stochastic dominance relation holds for an appropriate pairing of regions. An example of such a pairing for the case of three items is illustrated in Figure 4. The proof of Theorem 5 is in the Supplemental Material.

We now consider what happens when n becomes large while c remains fixed. In this case, in contrast to the previous result, we show using our strong duality theorem that grand bundling is *never* the optimal mechanism for sufficiently large values of n.

THEOREM 6: *For any $c \ge 0$, there exists an integer n_0 such that, for all $n \ge n_0$, the optimal mechanism for selling n i.i.d. goods whose values are uniform on $[c, c+1]$ is* not *a take-it-or-leave-it offer for the grand bundle.*

PROOF: Given c, let n be large enough so that

$$\frac{n+1}{n!} + \frac{nc}{(n-1)!} < 1.$$

To prove the theorem, we will assume that an optimal grand-bundling price p exists and reach a contradiction.

As shown in Section 2.2, under the transformed measure μ the hypercube has mass $-(n+1)$ in the interior, $+1$ on the origin, $c+1$ on every positive surface $x_i = c+1$, and $-c$ on every negative surface $x_i = c$.

According to Theorem 3, for grand bundling at price p to be optimal, it must hold that $\mu|_{Z_p} \preceq_{\text{cvx}} 0$ for the region $Z_p = \{x : \|x\|_1 \le p\}$. If $p > nc + 1$, this could not happen, since for the function $\mathbb{1}_{x_1=c+1}(x)$ (which is increasing and convex in $[c, c+1]^n$), we have that $\int_{Z_p} \mathbb{1}_{x_1=c+1}\, d\mu = \mu(Z_p \cap \{x_1 = c+1\}) = \mu_+(Z_p \cap \{x_1 = c+1\}) > 0$, which violates the $\mu|_{Z_p} \preceq_{\text{cvx}} 0$ condition.

To complete the proof, we now consider the case that $p \le nc + 1$ and will derive a contradiction. For the necessary condition $\mu|_{Z_p} \preceq_{\text{cvx}} 0$ to hold, it must be that $\mu(Z_p) = 0$. Since $p \le nc + 1$, none of the positive outer surfaces of the cube have nontrivial intersection with Z_p, so all the positive mass in Z_p is located at the origin. Therefore, $\mu_+(Z_p) = 1$, which means that $\mu_-(Z_p) = 1$ as well. Moreover, since $p \le nc + 1 \Rightarrow Z_p \subseteq Z_{nc+1}$, we also have that $\mu_-(Z_{nc+1}) \ge \mu_-(Z_p) = 1$.

To reach a contradiction, we will show that $\mu_-(Z_{nc+1}) < 1$. We observe that we can compute $\mu_-(Z_{nc+1})$ directly by summing the n-dimensional volume of the negative interior with the $(n-1)$-dimensional volumes of each of the n negative surfaces enclosed in Z_{nc+1}.[6] The first is equal to

$$(n+1) \times \mathrm{Vol}\big[\big\{x \in (c, c+1)^n : \|x\|_1 \le nc+1\big\}\big]$$
$$= (n+1) \times \mathrm{Vol}\big[\big\{x \in (0, 1)^n : \|x\|_1 \le 1\big\}\big] = \frac{(n+1)}{n!},$$

while the latter is equal to

$$n \times c \times \mathrm{Vol}\big[\big\{x \in (c, c+1)^{n-1} : \|x\|_1 + c \le nc+1\big\}\big]$$
$$= n \times c \times \mathrm{Vol}\big[\big\{x \in (0, 1)^{n-1} : \|x\|_1 \le 1\big\}\big] = \frac{nc}{(n-1)!}.$$

Therefore, we get that $1 \le \mu_-(Z_{nc+1}) = \frac{(n+1)}{n!} + \frac{nc}{(n-1)!}$, which is a contradiction since we chose n to be sufficiently large to make this quantity less than 1. *Q.E.D.*

7. CONSTRUCTING OPTIMAL MECHANISMS

7.1. *Preliminaries*

The results of the previous section characterize optimal mechanisms and give us the tools to check if a mechanism is optimal. In this section, we show how to use the optimal menu conditions we developed to identify candidate mechanisms. In particular, Theorem 3 implies that (in the finite-menu case) to find an optimal mechanism, we need to identify a set of choices for the menu such that, for every region R that corresponds to a menu outcome, it holds that $\mu_+|_R \preceq_{\mathrm{cvx}(\vec{v})} \mu_-|_R$ for the appropriate vector $\vec{v}$. This implies that $\mu_+(R) = \mu_-(R)$, so at the very least the total positive and the total negative mass in each region need to be equal. This property immediately helps us exclude a large class of mechanisms and guides us to identify potential candidates. We note that in this section we will develop techniques which apply not just to finite-menu mechanisms but to mechanisms with infinite menus as well.

We will restrict ourselves to a particularly useful class of mechanisms defined completely by the set of types that are excluded from the mechanism, that is, they receive no items and pay nothing. We call this set of types the *exclusion set* of a mechanism. The exclusion set gives rise to a mechanism where the utility of a buyer is equal to the ℓ_1 distance between the buyer's type and the closest point in the exclusion set. All known instances of optimal mechanisms for independently distributed items fall under this category. We proceed to define these concepts formally.

[6]The geometric intuition of this step of the argument is that, for large enough n, the fraction of the n-dimensional hypercube $[0, 1]^n$ which lies below the diagonal $\|x\| = 1$ goes to zero, and similarly the fraction of $(n-1)$-dimensional surface area on the boundaries which lies below the diagonal also goes to zero as n gets large.

DEFINITION 10—Exclusion Set: Let $X = \prod_{i=1}^{n}[x_i^{\text{low}}, x_i^{\text{high}}]$. An *exclusion set Z of X* is a convex, compact, and decreasing[7] subset of X with nonempty interior.

DEFINITION 11—Mechanism of an Exclusion Set: Every exclusion set Z of X induces a mechanism whose utility function $u_Z : X \to \mathbb{R}$ is defined by

$$u_Z(x) = \min_{z \in Z} \|z - x\|_1.$$

Note that, since the exclusion set Z is closed, for any $x \in X$ there exists a $z \in Z$ such that $u_Z(x) = \|z - x\|_1$. Moreover, we show below that any such utility function u_Z satisfies the constraints of the mechanism design problem. That is, the mechanism corresponding to u_Z is IC and IR. The proof of the following claim is straightforward casework and appears in the Supplemental Material.

CLAIM 1: *Let Z be an exclusion set of X. Then u_Z is nonnegative, nondecreasing, convex, and has Lipschitz constant (with respect to the ℓ_1 norm) at most* 1. *In particular, u_Z is the utility function of an incentive compatible and individually rational mechanism.*

7.2. *Constructing Optimal Mechanisms for Two Items*

To provide sufficient conditions for u_Z to be optimal for the case of two items, we define the concept of a canonical partition. A canonical partition divides X into regions such that the mechanism's allocation function within each region has a similar form. Roughly, the canonical partition separates X based on which direction (either "down," "left," or "diagonally") one must travel to reach the closest point in Z. While the definition is involved, the geometric picture of Figure 5 is straightforward.

DEFINITION 12—Critical Price, Critical Point, Outer Boundary Functions: Let Z be an exclusion set of X. Denote by P the maximum value $P = \max\{x + y : (x, y) \in Z\}$; we call

[7]A decreasing subset $Z \subset X$ satisfies the property that, for all $a, b \in X$ such that a is component-wise less than or equal to b, if $b \in Z$, then $a \in Z$ as well.

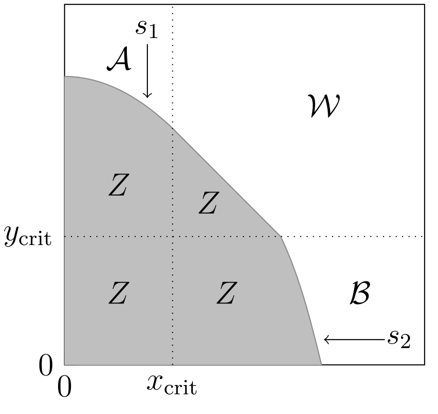

FIGURE 5.—The canonical partition.

P the *critical price*. We now define the *critical point* $(x_{\text{crit}}, y_{\text{crit}})$, such that

$$x_{\text{crit}} = \min\{x : (x, P - x) \in Z\} \quad \text{and} \quad y_{\text{crit}} = \min\{y : (P - y, y) \in Z\}.$$

We define the *outer boundary functions of* Z to be the functions s_1, s_2 given by

$$s_1(x) = \max\{y : (x, y) \in Z\} \quad \text{and} \quad s_2(y) = \max\{x : (x, y) \in Z\},$$

with domain $[0, x_{\text{crit}}]$ and $[0, y_{\text{crit}}]$, respectively.

DEFINITION 13—Canonical Partition: Let Z be an exclusion set of X with critical point $(x_{\text{crit}}, y_{\text{crit}})$ as in Definition 12. We define *the canonical partition of* X *induced by* Z to be the partition of X into $Z \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{W}$, where

$$\mathcal{A} = \{(x, y) \in X : x < x_{\text{crit}}\} \setminus Z; \qquad \mathcal{B} = \{(x, y) \in X : y < y_{\text{crit}}\} \setminus Z;$$
$$\mathcal{W} = X \setminus (Z \cup \mathcal{A} \cup \mathcal{B}),$$

as shown in Figure 5.

Note that the outer boundary functions s_1, s_2 of an exclusion set Z are concave and thus are differentiable almost everywhere on $[0, c_1]$ and have non-increasing derivatives.

We now restate the utility function u_Z of a mechanism with exclusion set Z in terms of a canonical partition.

CLAIM 2: *Let* Z *be an exclusion set of* X *with outer boundary functions* s_1, s_2 *and critical price* P, *and let* $Z \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{W}$ *be its canonical partition. Then for all* $(v_1, v_2) \in X$, *the utility function* u_Z *of the mechanism with exclusion set* Z *is given by*

$$u_Z(v_1, v_2) = \begin{cases} 0, & \text{if } (v_1, v_2) \in Z, \\ v_2 - s_1(v_1), & \text{if } (v_1, v_2) \in \mathcal{A}, \\ v_1 - s_2(v_2), & \text{if } (v_1, v_2) \in \mathcal{B}, \\ v_1 + v_2 - P, & \text{if } (v_1, v_2) \in \mathcal{W}. \end{cases}$$

PROOF: The proof is fairly straightforward casework. We prove one of the cases here, and the remaining cases are similar.

Pick any $v = (v_1, v_2) \in \mathcal{A}$. We will show that the closest $z \in Z$ is the point $z^* = (v_1, s_1(v_1))$. Pick $z' = (z'_1, z'_2) \in Z$ such that $u_Z(v) = \|v - z'\|_1$. It must be the case that $z'_1 \leq v_1$, since otherwise (v_1, z'_2) would be in Z (as Z is decreasing) and strictly closer to v.

We now have that $\|v - z'\|_1 \geq \|v\|_1 - \|z'\|_1 \geq \|v\|_1 - \max_{x \in [0, v_1]}(x + s_1(x))$. Since the less restricted maximization problem, $\max_{x \in [0, x_{\text{crit}}]}(x + s_1(x))$, is maximized at x_{crit} and the function $(x + s_1(x))$ is concave, the maximum of the more constrained version is achieved at $x = v_1$. Thus, we have that $\|v - z'\|_1 \geq \|v\|_1 - v_1 - s_1(v_1) = v_2 - s_1(v_1) = \|v - z^*\|_1$. *Q.E.D.*

We now describe sufficient conditions under which u_Z is optimal.

DEFINITION 14—Well-Formed Canonical Partition: Let $Z \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{W}$ be a canonical partition of X induced by exclusion set Z and let μ be a signed Radon measure on X such that $\mu(X) = 0$. We say that the canonical partition is *well-formed with respect to* μ if the following conditions are satisfied:

1. $\mu|_Z \preceq_{\text{cvx}} 0$ and $\mu|_{\mathcal{W}} \succeq_2 0$, and
2. for all $v \in X$ and all $\varepsilon > 0$:
- $\mu|_{\mathcal{A}}([v_1, v_1 + \varepsilon] \times [v_2, \infty)) \geq 0$, with equality whenever $v_2 = 0$,
- $\mu|_{\mathcal{B}}([v_1, \infty) \times [v_2, v_2 + \varepsilon]) \geq 0$, with equality whenever $v_1 = 0$.

We point out the similarities between a well-formed canonical partition and the sufficient conditions for menu optimality of Theorem 3. Condition 1 gives exactly the stochastic dominance conditions that need to hold in regions Z and $\mathcal{W}$. We interpret Condition 2 as saying that $\mu|_{\mathcal{A}}$ (resp. $\mu|_{\mathcal{B}}$) allows for the positive mass in any vertical (resp. horizontal) "strip" to be matched to the negative mass in the strip by only transporting "downwards" (resp. "leftwards"). These conditions guarantee (single-dimensional) first-order dominance of the measures along each strip, which is a stronger requirement than the convex dominance conditions of Theorem 3. In practice, when μ is given by a density function, we verify these conditions by analyzing the integral of the density function along appropriate vertical or horizontal lines. Even though Theorem 3 applies only for mechanisms with finite menus, we prove in Theorem 7 that a mechanism induced by an exclusion set is optimal for a two-item instance if the canonical partition of its exclusion set is well-formed. Refer back to Figure 5 to visualize such a mechanism.

THEOREM 7: *Let μ be the transformed measure of a probability density function f. If there exists an exclusion set Z inducing a canonical partition $Z \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{W}$ of X that is well-formed with respect to μ, then the optimal IC and IR mechanism for a single additive buyer whose values for two goods are distributed according to the joint distribution f is the mechanism induced by exclusion set Z. In particular, the mechanism uses the following allocation and price for a buyer with reported type $(x, y) \in X$:*
- *if $(x, y) \in Z$, the buyer receives no goods and is charged 0;*
- *if $(x, y) \in \mathcal{A}$, the buyer receives item 1 with probability $-s_1'(x)$, item 2 with probability 1, and is charged $s_1(x) - xs_1'(x)$;*
- *if $(x, y) \in \mathcal{B}$, the buyer receives item 2 with probability $-s_2'(y)$, item 1 with probability 1, and is charged $s_2(y) - ys_2'(y)$;*
- *if $(x, y) \in \mathcal{W}$, the buyer receives both goods with probability 1 and is charged P;*

where s_1, s_2 are the boundary functions and P is the critical price as in Definition 12.

PROOF: We will show that u_Z maximizes $\sup_{u \in \mathcal{U}(X) \cap \mathcal{L}_1(X)} \int_X u \, d\mu$. By Corollary 1, it suffices to provide a $\gamma \in \Gamma_+(X \times X)$ such that $\gamma_1 - \gamma_2 \succeq_{\text{cvx}} \mu$, $\int u_Z \, d(\gamma_1 - \gamma_2) = \int u_Z \, d\mu$, and $u_Z(x) - u_Z(y) = \|x - y\|_1$ holds γ-almost surely. The γ we construct will never transport mass between regions. That is, $\gamma = \gamma_Z + \gamma_{\mathcal{W}} + \gamma_{\mathcal{A}} + \gamma_{\mathcal{B}}$, where[8]

- $\gamma_Z = 0$. We notice that $(\gamma_Z)_1 - (\gamma_Z)_2 = 0 \succeq_{\text{cvx}} \mu|_Z$.
- $\gamma_{\mathcal{W}}$ is constructed such that $(\gamma_{\mathcal{W}})_1 - (\gamma_{\mathcal{W}})_2 \succeq_{\text{cvx}} \mu|_{\mathcal{W}}$ and the component-wise inequality $x \geq y$ holds $\gamma_{\mathcal{W}}(x, y)$ almost surely.[9] As in our proof of Theorem 3, the existence of such a $\gamma_{\mathcal{W}}$ is guaranteed by Strassen's theorem for second-order dominance (presented in the Supplemental Material).
- $\gamma_{\mathcal{A}} \in \Gamma_+(\mathcal{A} \times \mathcal{A})$ will be constructed to have respective marginals $\mu_+|_{\mathcal{A}}$ and $\mu_-|_{\mathcal{A}}$, and so that, $\gamma_{\mathcal{A}}(x, y)$ almost surely, it holds that $x_1 = y_1$ and $x_2 \geq y_2$. Thus, $(\gamma_{\mathcal{A}})_1 - (\gamma_{\mathcal{A}})_2 = \mu|_{\mathcal{A}}$, and $\gamma_{\mathcal{A}}$ sends positive mass "downwards."[10] We claim that such a map can indeed be

[8]We chose this notation for simplicity, where $\gamma_Z \in \Gamma_+(Z \times Z)$, $\gamma_{\mathcal{W}} \in \Gamma_+(\mathcal{W} \times \mathcal{W})$, and so on.

[9]As in Example 2 and as discussed in Remark 1, we aim for $\gamma_{\mathcal{W}}$ to transport "downwards and leftwards" since both items are allocated with probability 1 in $\mathcal{W}$.

[10]Once again, the intuition for this construction follows Remark 1.

constructed, by noticing that Property 2 of Definition 14 guarantees that, restricted to any vertical strip inside $\mathcal{A}$, μ_+ first-order stochastically dominates μ_-.[11] Hence, Strassen's theorem for first-order dominance guarantees that, restricted to that strip, μ_+ can be coupled with μ_- so that, with probability 1, mass is only moved downwards.

Measure $\gamma_{\mathcal{A}}$ satisfies $x_1 = y_1$, $\gamma_{\mathcal{A}}(x, y)$ almost surely, and hence also

$$u_Z(x) - u_Z(y) = \big(x_2 - s(x_1)\big) - \big(y_2 - s(y_1)\big) = x_2 - y_2 = \|x - y\|_1.$$

- $\gamma_{\mathcal{B}} \in \Gamma_+(\mathcal{B} \times \mathcal{B})$ is constructed analogously to $\gamma_{\mathcal{A}}$, except sending mass "leftwards." That is, $\gamma_{\mathcal{B}}(x, y)$ almost-surely, the relationships $x_1 \geq y_1$ and $x_2 = y_2$ hold.

It follows by our construction that $\gamma = \gamma_Z + \gamma_{\mathcal{W}} + \gamma_{\mathcal{A}} + \gamma_{\mathcal{B}}$ satisfies all necessary properties to certify optimality of u_Z. *Q.E.D.*

8. APPLYING THEOREM 7 TO FIND OPTIMAL MECHANISMS

In this section, we provide example applications of Theorem 7. A technical difficulty is verifying the stochastic dominance relation $\mu|_{\mathcal{W}} \succeq_2 0$ required to apply the theorem. In our examples, we will have the stronger condition $\mu|_{\mathcal{W}} \succeq_1 0$, which is easier to verify, yet still imposes technical difficulties. In Section 8.1, we present a useful tool, Lemma 3, for verifying first-order stochastic dominance. In Section 8.2, we then provide example applications of Theorem 7 and Lemma 3 to solve for optimal mechanisms.

8.1. *Verifying First-Order Stochastic Dominance*

A useful tool for verifying first-order dominance between measures is the following.[12]

LEMMA 3: *Let $\mathcal{C} = [p_1, q_1) \times [p_2, q_2)$, where q_1 and q_2 are possibly infinite, and let R be a decreasing nonempty subset of $\mathcal{C}$. Consider two measures $\kappa, \lambda \in \Gamma_+(\mathcal{C})$ with bounded integrable density functions $g, h : \mathcal{C} \to \mathbb{R}_{\geq 0}$, respectively, that satisfy the conditions*:

- *$g(x, y) = h(x, y) = 0$ for all $(x, y) \in R$.*
- *$\int_{\mathcal{C}} g(x, y)\, dx\, dy = \int_{\mathcal{C}} h(x, y)\, dx\, dy$.*
- *For any basis vector $e_i \in \{e_1 \equiv (1, 0), e_2 \equiv (0, 1)\}$ and any point $z \in R$:*

$$\int_0^{q_i - z_i} g(z + \tau e_i) - h(z + \tau e_i)\, d\tau \leq 0.$$

- *There exist nonnegative functions $\alpha : [p_1, q_1) \to \mathbb{R}_{\geq 0}$ and $\beta : [p_2, q_2) \to \mathbb{R}_{\geq 0}$, and an increasing function $\eta : \mathcal{C} \to \mathbb{R}$ such that, for all $(x, y) \in \mathcal{C} \setminus R$,*

$$g(x, y) - h(x, y) = \alpha(x) \cdot \beta(y) \cdot \eta(x, y).$$

Then $\kappa \succeq_1 \lambda$.

Lemma 3 provides a sufficient condition for a measure to stochastically dominate another in the first order. Its proof is given in the Supplemental Material and is an application of a claim which states that an equivalent condition for first-order stochastic

[11]Indeed, as $\varepsilon \to 0$, Property 2 states exactly the one-dimensional equivalent condition for first-order stochastic dominance in terms of cumulative density functions.

[12]The lemma also appeared as Theorem 7.4 of Daskalakis, Deckelbaum, and Tzamos (2013) without a proof. We provide a detailed proof in the Supplemental Material.

dominance is that one measure has more mass than the other on all sets that are unions of *finitely many* "increasing boxes." When the conditions of Lemma 3 are satisfied, we can induct on the number of boxes by removing one box at a time. We note that Lemma 3 is applicable even to distributions with unbounded support.

Interpreting the Conditions of Lemma 3

Lemma 3 is applicable whenever two density functions, g and h, are nonzero on some set $\mathcal{C} \setminus R$, where R is a decreasing subset of some two-dimensional box $\mathcal{C}$. This setting is motivated by Figure 5 and Theorem 7. Recall that, in order to apply Theorem 7, we need to check a second-order stochastic dominance condition in region $\mathcal{W}$, namely, $\mu|_{\mathcal{W}} \succeq_2 0$.

While Theorem 7 demands checking a second-order stochastic dominance condition, an easier and sufficient goal is to check first-order stochastic dominance, namely, $\mu|_{\mathcal{W}} \succeq_1 0$. To do this, we can readily use Lemma 3, by taking $\mathcal{C} = [x_{\text{crit}}, \infty) \times [y_{\text{crit}}, \infty)$, $R = \mathcal{C} \cap Z$, and g, h the densities corresponding to measures $\mu_+|_{\mathcal{W}}$ and $\mu_-|_{\mathcal{W}}$. The way region $\mathcal{W}$ is defined in Theorem 7 guarantees that the two measures have equal mass, so the first two conditions of the lemma will be satisfied automatically. For the third condition, we need to verify that, if we integrate $g - h$ along either a vertical or a horizontal line outwards starting from any point in R, the result is *non-positive*. The last condition of Lemma 3 requires that the density function of the measure $\mu|_{\mathcal{W}}$, that is, $g - h$, have an appropriate form. If the values of the buyer for the two items are independently distributed according to distributions with densities f_1 and f_2, then the density of measure μ in the interior according to Equation (3) can be written as $-f_1(x)f_2(y)(\frac{f_1'(x)x}{f_1(x)} + \frac{f_2'(y)y}{f_2(y)} + 3)$. The last condition of the lemma is thus satisfied if the functions $\frac{f_1'(x)x}{f_1(x)}$ and $\frac{f_2'(y)y}{f_2(y)}$ are decreasing, a condition that is easy to verify.

8.2. *Examples*

We apply Theorem 7 to obtain optimal mechanisms in several two-item settings. In Section 8.2.1, we consider two independent items distributed according to beta distributions. We find the optimal mechanism, showing that it actually offers an uncountably infinite menu of lotteries. We conclude with Section 8.2.2, where we discuss extensions of Theorem 7 to distributions with infinite support, providing the optimal mechanism for two arbitrary independent exponential items, as well as the optimal mechanism for an instance with two independent power-law items.

8.2.1. *An Optimal Mechanism With Infinite Menu-Size: Two Beta Items*

In this section, we will use Theorem 7 to calculate the optimal mechanism for two items distributed according to Beta distributions. In doing so, we illustrate a general approach for finding closed-form descriptions of optimal mechanisms via the following steps: (i) definition of the sets S_{top} and S_{right}, (ii) computation of a critical price p^*, (iii) definition of a canonical partition in terms of (i) and (ii), and (iv) application of Theorem 7. Our approach succeeds in pinning down optimal mechanisms in all examples considered in Sections 8.2.1–8.2.2, and we expect it to be broadly applicable. Finally, it is noteworthy that the optimal mechanism for the setting studied in this section offers the buyer a menu of uncountably many lotteries to choose from. Using our approach, we can nevertheless compute and succinctly describe the optimal mechanism. We also note in Remark 4 that our identified mechanism is essentially unique, hence the uncountability of the menu is inevitable.

Consider two items whose values are distributed independently according to the distributions $\text{Beta}(a_1, b_1)$ and $\text{Beta}(a_2, b_2)$, respectively. That is, the distributions are given by the following two density functions on $[0, 1]$:

$$f_1(x) = \frac{1}{B(a_1, b_1)} x^{a_1-1}(1-x)^{b_1-1}; \qquad f_2(y) = \frac{1}{B(a_2, b_2)} y^{a_2-1}(1-y)^{b_2-1}.$$

To find the optimal mechanism for our example setting, we first compute the measure μ induced by f. Notice that

$$\begin{aligned}
&-\nabla f(x, y) \cdot (x, y) - 3f(x, y) \\
&= -xf_2(y)\frac{\partial f_1(x)}{\partial x} - yf_1(x)\frac{\partial f_2(y)}{\partial y} - 3f_1(x)f_2(y) \\
&= -(a_1 - 1)f_1(x)f_2(y) + (b_1 - 1)\frac{x}{1-x}f_1(x)f_2(y) \\
&\quad - (a_b - 1)f_1(x)f_2(y) + (b_2 - 1)\frac{y}{1-y}f_1(x)f_2(y) - 3f_1(x)f_2(y) \\
&= f_1(x)f_2(y)\left(\frac{b_1 - 1}{1-x} + \frac{b_2 - 1}{1-y} + (1 - a_1 - b_1 - a_2 - b_2)\right),
\end{aligned}$$

where the last equality used the identity $\frac{x}{1-x} = \frac{1}{1-x} - 1$. We also observe that $f_1(x)x = 0$ whenever $x = 0$ or $x = 1$ (as long as $b_1 > 1$), and an analogous property holds for y. Thus, the transformed measure μ is composed of:

- a point mass of $+1$ at the origin; and
- mass distributed on $[0, 1]^2$ according to the density function

$$f_1(x)f_2(y)\left(\frac{b_1 - 1}{1-x} + \frac{b_2 - 1}{1-y} + (1 - a_1 - b_1 - a_2 - b_2)\right).$$

Note that in the case $b_i = 1$, our analysis still holds, except there is also positive mass on the boundary $x_i = 1$.

Deriving the Optimal Mechanism for a Concrete Setting of Parameters. We now analyze a concrete example of two independent Beta distributed items where $a_1 = a_2 = 1$ and $b_1 = b_2 = 2$. That is, we consider two items whose values are distributed independently according to the following two density functions on $[0, 1]$:

$$f_1(x) = 2(1-x); \qquad f_2(y) = 2(1-y).$$

As discussed above, the transformed measure μ comprises:

- a point mass of $+1$ at the origin; and
- mass distributed on $[0, 1]^2$ according to the density function

$$f_1(x)f_2(y)\left(\frac{1}{1-x} + \frac{1}{1-y} - 5\right).$$

Note that the density of μ is positive on $\mathcal{P} = \{(x, y) \in (0, 1)^2 : \frac{1}{1-x} + \frac{1}{1-y} > 5\} \cup \{\vec{0}\}$ and non-positive on $\mathcal{N} = \{(x, y) \in [0, 1)^2 \setminus \{\vec{0}\} : \frac{1}{1-x} + \frac{1}{1-y} \leq 5\}$, and that $\mathcal{N} \cup \{\vec{0}\}$ is a decreasing set.

Step (i). We first attempt to identify candidate functions for s_1 and s_2 that will lead to a well-formed canonical partition. We do this by defining two sets $S_{\text{top}}, S_{\text{right}} \subset [0,1)^2$. We require that $(x,y) \in S_{\text{top}}$ iff $\int_y^1 \mu(x,t)\,dt = 0$. That is, starting from any point $z \in S_{\text{top}}$ and integrating the density of μ "upwards" from $t = y$ to $t = 1$ yields zero. Since $\mathcal{N} \cup \{\vec{0}\}$ is a decreasing set, it follows that $S_{\text{top}} \subset \mathcal{N}$ and that integrating μ upwards starting from any point above S_{top} yields a positive integral. Similarly, we say that $(x,y) \in S_{\text{right}}$ iff $\int_x^1 \mu(t,y)\,dt = 0$, noting that $S_{\text{right}} \subset \mathcal{N}$. S_{top} and S_{right} are shown in Figure 6.

We analytically compute that $(x,y) \in S_{\text{top}}$ if and only if $y = \frac{2-3x}{4-5x}$. Similarly, $(x,y) \in S_{\text{right}}$ if and only if $x = \frac{2-3y}{4-5y}$.

In particular, for any $x \leq 2/3$, there exists a y such that $(x,y) \in S_{\text{top}}$, and there does not exist such a y if $x > 2/3$. Furthermore, it is easy to verify by computing the second derivative of $\frac{\partial^2}{\partial x^2}\frac{2-3x}{4-5x} = -\frac{20}{(4-5x)^3} < 0$ that the region below S_{top} and the region below S_{right} are strictly convex.

Step (ii). We now need to calculate the critical point and the critical price. To do this, we set the critical price $p^* \approx 0.5535$ as the intercept of the 45° line in Figure 6, which causes $\mu(Z) = 0$ for the set $Z \subset [0,1]^2$ lying below S_{top}, S_{right}, and the 45° line. We can also compute the critical point $(x_{\text{crit}}, y_{\text{crit}}) \approx (0.0618, 0.0618)$ by finding the intersection of the critical price line with the sets S_{top} and S_{bottom}. Moreover, by the definition of the sets S_{top} and S_{bottom}, we know that the candidate boundary functions are $s_1(x) = \frac{2-3x}{4-5x}$ and $s_2(y) = \frac{2-3y}{4-5y}$, with domain $[0, x_{\text{crit}})$ and $[0, y_{\text{crit}})$, respectively.

Step (iii). We can now compute the canonical partition and decompose $[0,1]^2$ into the following regions:

$$\mathcal{A} = \{(x,y) : x \in [0, x_{\text{crit}}) \text{ and } y \in [s_1(x), 1]\};$$
$$\mathcal{B} = \{(x,y) : y \in [0, y_{\text{crit}}) \text{ and } x \in [s_2(y), 1]\};$$

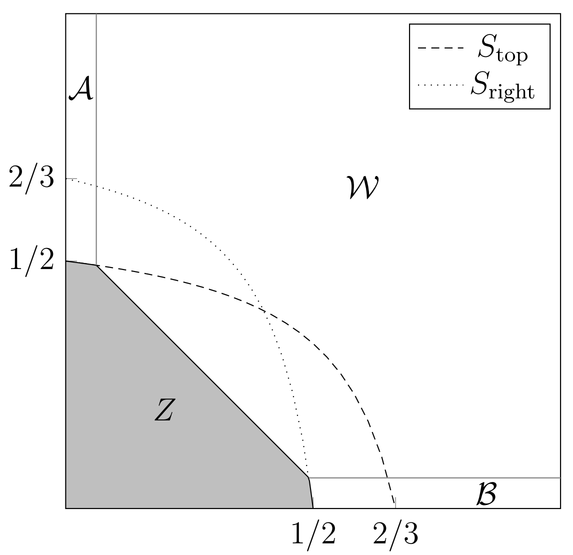

FIGURE 6.—The well-formed canonical partition for $f_1(x) = 2(1-x)$ and $f_2(y) = 2(1-y)$.

$$\mathcal{W} = \{(x, y) \in [x_{\text{crit}}, 1] \times [y_{\text{crit}}, 1] : x + y \geq p^*\};$$
$$Z = [0, 1]^2 \setminus (\mathcal{W} \cup \mathcal{A} \cup \mathcal{B}),$$

as illustrated in Figure 6.

Step (iv). We claim that the canonical partition $Z \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{W}$ is well-formed with respect to μ. Condition 2 is satisfied by construction of S_{top} and S_{right} and the corresponding discussion in Step (i). To check for Condition 1, note that given the definition of p^*, it holds that for all regions $R = Z, \mathcal{A}, \mathcal{B},$ and $\mathcal{W}$, we have $\mu(R) = 0$. Recall that $S_{\text{top}}, S_{\text{right}} \subset \mathcal{N}$ and, since $\mathcal{N} \cup \{\vec{0}\}$ is a decreasing set, μ has negative density along these curves and all points below either curve, other than at the origin. Hence, $\mu_-|_Z \succeq_1 \mu_+|_Z$, which implies that $\mu|_Z \preceq_{\text{cvx}} 0$. Hence, the only nontrivial condition of Definition 14 that we need to verify is $\mu|_{\mathcal{W}} \succeq_2 0$. In fact, we can apply Lemma 3 to conclude the stronger dominance relation $\mu|_{\mathcal{W}} \succeq_1 0$. See the Supplemental Material. Having verified all conditions of Definition 14, we apply Theorem 7 to conclude the following.

EXAMPLE 3: The optimal mechanism for selling two independent items whose values are distributed according to $f_1(x) = 2(1 - x)$ and $f_2(y) = 2(1 - y)$ has the following outcome for a buyer of type (x, y):

- If $(x, y) \in Z$, the buyer receives no goods and is charged 0.
- If $(x, y) \in \mathcal{A}$, the buyer receives item 1 with probability $-s_1'(x) = \frac{2}{(4-5x)^2}$, item 2 with probability 1, and is charged $s_1(x) - xs_1'(x) = \frac{2-3x}{4-5x} + \frac{2x}{(4-5x)^2}$.
- If $(x, y) \in \mathcal{B}$, the buyer receives item 2 with probability $-s_2'(y) = \frac{2}{(4-5y)^2}$, item 1 with probability 1, and is charged $s_2(y) - ys_2'(y) = \frac{2-3y}{4-5y} + \frac{2y}{(4-5y)^2}$.
- If $(x, y) \in \mathcal{W}$, the buyer receives both items and is charged $p^* \approx 0.5535$.

REMARK 4: Note that the mechanism identified in Example 3 offers an uncountably large menu of lotteries. One could wonder whether there exists a different optimal mechanism offering a finite menu. Using our duality theorem, we can easily argue that the utility function induced by every optimal mechanism equals the utility function $u(x)$ induced by our mechanism in Example 3. Hence, up to the choice of subgradients at the measure-zero set of types where $\nabla u(x)$ is discontinuous, the allocations offered by any optimal mechanism must agree with those of our mechanism in Example 3. Therefore, every optimal mechanism must offer an uncountably large menu. The proof of uniqueness is given in the Supplemental Material.

Summary of Beta Distributions. Example 3 shows that the optimal mechanism for two Beta distributed items offers a continuum of lotteries, thereby having infinite menu-size complexity (Hart and Nisan (2013)). Still, using our techniques, we can obtain a succinct and easily computable description of the mechanism.

Working similarly to Example 3, we can obtain the optimal mechanism for broader settings of parameters. Figure 7 illustrates the optimal mechanism for two items distributed according to Beta distributions with different parameters. The reader can experiment with different settings of parameters at http://christos.me/betas.

8.2.2. *Distributions of Unbounded Support: Exponential and Power-Law*

So far, this paper has focused on type distributions with bounded support. In this section, we note that Theorem 1, Lemma 1, and Theorem 7 can be easily modified to accommodate settings with unbounded type spaces, as long as the type distribution decays

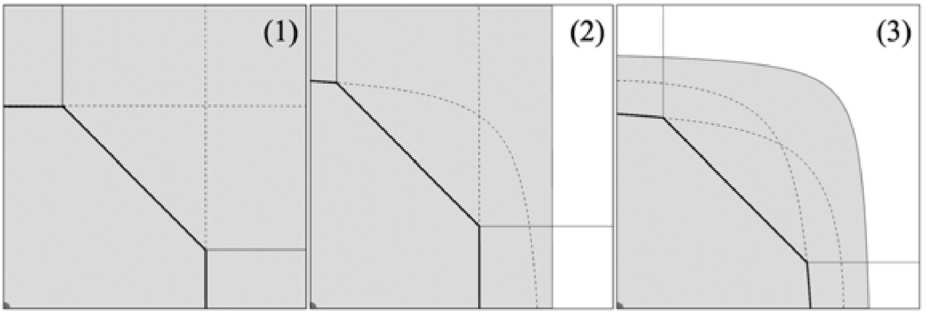

FIGURE 7.—Canonical partitions for different cases of Beta distributions. The shaded region is where the measure μ becomes negative. (Note that when the second parameter b_i of the Beta distribution of some item i equals 1, μ has positive mass on the outer boundary $x_i = 1$.) (1) Beta$(1, 1)$ and Beta$(1, 1)$, (2) Beta$(2, 2)$ and Beta$(1, 1)$, (3) Beta$(2, 2)$ and Beta$(2, 2)$.

sufficiently rapidly towards infinity. On the other hand, we do not know extensions of our strong duality theorem (Theorem 2), and the optimal menu conditions (Theorem 3) for unbounded type distributions, due to technical issues.

In the Supplemental Material, we provide a short discussion of the modifications required to obtain an analog of Theorem 7 for unbounded distributions that are sufficiently fast-decaying, and present below two example settings that can be analyzed using the modified characterization theorem. Both examples are taken from Daskalakis, Deckelbaum, and Tzamos (2013).

In Example 4, the optimal mechanism for selling two power-law items is a grand-bundling mechanism. The canonical partition induced by the exclusion set of the grand-bundling mechanism is degenerate (regions $\mathcal{A}$ and $\mathcal{B}$ are empty), and establishing the optimality of the mechanism amounts to establishing that the first-order stochastic dominance condition for the induced measure μ holds in region $\mathcal{W}$.

EXAMPLE 4: The optimal IC and IR mechanism for selling two items whose values are distributed independently according to the probability densities $f_1(x) = 5/(1+x)^6$ and $f_2(y) = 6/(1+y)^7$, respectively, is a take-it-or-leave-it offer of the bundle of the two goods for price $p^* \approx 0.35725$.

Example 5 provides a complete solution for the optimal mechanism for two items distributed according to independent exponential distributions. In this case, the canonical partition, illustrated in Figure 8, induced by the exclusion set of the mechanism is missing region $\mathcal{A}$, and possibly region $\mathcal{B}$ (if $\lambda_1 = \lambda_2$).

EXAMPLE 5: For all $\lambda_1 \geq \lambda_2 > 0$, the optimal IC and IR mechanism for selling two items whose values are distributed independently according to exponential distributions f_1 and f_2 with respective parameters λ_1 and λ_2 offers the following menu:

1. receive nothing, and pay 0;
2. receive the first item with probability 1 and the second item with probability λ_2/λ_1, and pay $2/\lambda_1$; and
3. receive both items, and pay p^*;

where p^* is the unique $0 < p^* \leq 2/\lambda_2$ such that

$$\mu\big(\{(x, y) \in \mathbb{R}^2_{\geq 0} : x + y \leq p^* \text{ and } \lambda_1 x + \lambda_2 y \leq 2\}\big) = 0,$$

where μ is the transformed measure of the joint distribution.

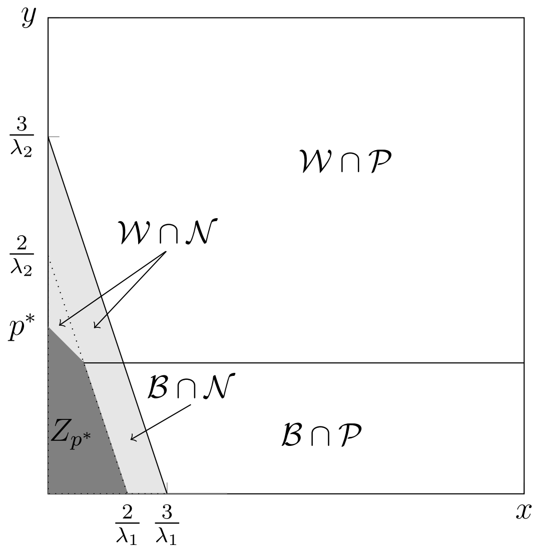

FIGURE 8.—The canonical partition of $\mathbb{R}^n_{\geq 0}$ for the proof of Example 5. In this diagram, $p^* > 2/\lambda_1$. If $p^* \leq 2/\lambda_1$, $\mathcal{B}$ is empty. The positive part μ_+ of μ is supported inside $\mathcal{P} \cap \{\vec{0}\}$ while the negative part μ_- is supported within $Z_{p^*} \cup \mathcal{N}$.

9. CONCLUSIONS

We provided a duality-based framework for revenue maximization in a multiple-good monopoly. Our framework shows that every optimal mechanism has a certificate of optimality, taking the form of an optimal transportation map between measures. Using this framework, we characterized optimal mechanisms, showing that a mechanism is optimal if and only if certain stochastic dominance conditions are satisfied by a measure induced by the buyer's type distribution. This measure expresses the marginal change in the seller's revenue under marginal changes in the rent paid to subsets of buyer types.

We also provided several tools for checking the pertinent stochastic dominance conditions in two dimensions. These tools were useful in establishing the optimality of mechanisms in a multitude of two-item examples that we studied. While our characterization holds for an arbitrary number of items, verifying stochastic dominance in higher dimensions becomes significantly harder. An interesting future direction is to develop tools for checking stochastic dominance in higher dimensions. This will be useful for establishing optimality of mechanisms for three and more items.

Another important research direction is to obtain conditions for the type distribution under which the optimal mechanism has a simple closed-form description. For example, are there broad conditions implying that grand bundling is optimal or that the optimal mechanism takes the form of the mechanisms in Theorem 7?

Finally, a major open problem is to extend our results to multiple bidders. Even for the presumably simple setting of two bidders with independent and identical values for two items that are uniformly distributed in [0, 1], the revenue-optimal mechanism is unknown.

REFERENCES

ANDERSON, E. J., AND P. NASH (1987): *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Chichester: Wiley. [736]

ARMSTRONG, M. (1996): "Multiproduct Nonlinear Pricing," *Econometrica: Journal of the Econometric Society*, 64 (1), 51–75. [738]

——— (1999): "Price Discrimination by a Many-Product Firm," *Review of Economic Studies*, 66 (1), 151–168. [738]

BABAIOFF, M., N. IMMORLICA, B. LUCIER, AND S. M. WEINBERG (2014): "A Simple and Approximately Optimal Mechanism for an Additive Buyer," in *55th Annual IEEE Symposium on Foundations of Computer Science—FOCS*. Los Alamitos, CA: IEEE Computer Soc., 21–30. [738]

BAKOS, Y., AND E. BRYNJOLFSSON (1999): "Bundling Information Goods: Pricing, Profits, and Efficiency," *Management Science*, 45 (12), 1613–1630. [735]

BRIEST, P., S. CHAWLA, R. KLEINBERG, AND S. M. WEINBERG (2010): "Pricing Randomized Allocations," in *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*. Philadelphia, PA: SIAM, 585–597. [738]

CAI, Y., AND Z. HUANG (2013): "Simple and Nearly Optimal Multi-Item Auctions," in *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*. Philadelphia, PA: SIAM, 564–577. [738]

DASKALAKIS, C., A. DECKELBAUM, AND C. TZAMOS (2013): "Mechanism Design via Optimal Transport," in *14th ACM Conference on Electronic Commerce (EC)*. [736-738,742,745,748,751,759,764]

——— (2017): "Supplement to 'Strong Duality for a Multiple-Good Monopolist'," *Econometrica Supplemental Material*, 85, http://dx.doi.org/10.3982/ECTA12618. [736]

EKELAND, I. (2010): "Notes on Optimal Transportation," *Economic Theory*, 42 (2), 437–459. [739]

FIGALLI, A., Y.-H. KIM, AND R. J. MCCANN (2011): "When Is Multidimensional Screening a Convex Program?" *Journal of Economic Theory*, 146 (2), 454–478. [735,738]

GIANNAKOPOULOS, Y., AND E. KOUTSOUPIAS (2014): "Duality and Optimality of Auctions for Uniform Distributions," in *The 15th ACM Conference on Electronic Commerce (EC)*. [736-738,742,745,748,751]

HAGHPANAH, N., AND J. HARTLINE (2015): "Reverse Mechanism Design," in *The 16th ACM Conference on Economics and Computation (EC)*. [738]

HART, S., AND N. NISAN (2013): "The Menu-Size Complexity of Auctions," in *14th ACM Conference on Electronic Commerce (EC)*. [738,763]

——— (2014): "How Good Are Simple Mechanisms for Selling Multiple Goods?" [738]

HART, S., AND P. J. RENY (2015): "Maximal Revenue With Multiple Goods: Nonmonotonicity and Other Observations," *Theoretical Economics*, 10 (3), 893–922. [738]

LI, X., AND A. C.-C. YAO (2013): "On Revenue Maximization for Selling Multiple Independently Distributed Items," *Proceedings of the National Academy of Sciences*, 110 (28), 11232–11237. [738]

LUENBERGER, D. G. (1968): *Optimization by Vector Space Methods*. Wiley. [736]

MANELLI, A. M., AND D. R. VINCENT (2006): "Bundling as an Optimal Selling Mechanism for a Multiple-Good Monopolist," *Journal of Economic Theory*, 127 (1), 1–35. [735-738,740,742,747,751]

——— (2007): "Multidimensional Mechanism Design: Revenue Maximization and the Multiple-Good Monopoly," *Journal of Economic Theory*, 137 (1), 153–185. [735,738]

MASKIN, E., AND J. RILEY (1984): "Monopoly With Incomplete Information," *RAND Journal of Economics*, 15 (2), 171–196. [735]

MENICUCCI, D., S. HURKENS, AND D.-S. JEON (2015): "On the Optimality of Pure Bundling for a Monopolist," *Journal of Mathematical Economics*, 60, 33–42. [738]

MYERSON, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6 (1), 58–73. [735, 746]

PAVLOV, G. (2011): "Optimal Mechanism for Selling Two Goods," *B.E. Journal of Theoretical Economics*, 11 (1), 1–35. [735,737,738,748,753]

PRESTON MCAFEE, R., J. MCMILLAN, AND M. D. WHINSTON (1989): "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *Quarterly Journal of Economics*, 104 (2), 371–383. [735]

RILEY, J., AND R. ZECKHAUSER (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *The Quarterly Journal of Economics*, 98 (2), 267–289. [735]

RILEY, J. G., AND W. F. SAMUELSON (1981): "Optimal Auctions," *American Economic Review*, 71 (3), 381–392. [735]

ROCHET, J.-C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context," *Journal of Mathematical Economics*, 16 (2), 191–200. [736,739,740]

ROCHET, J.-C., AND P. CHONÉ (1998): "Ironing, Sweeping, and Multidimensional Screening," *Econometrica*, 66 (4), 783–826. [739,740,742,744]

ROCHET, J.-C., AND L. A. STOLE (2003): "The Economics of Multidimensional Screening," *Econometric Society Monographs*, 35, 150–197. [735,738]

THANASSOULIS, J. (2004): "Haggling Over Substitutes," *Journal of Economic Theory*, 117 (2), 217–245. [735, 738]

VILLANI, C. (2008): *Optimal Transport: Old and New*, Vol. 338. Berlin: Springer. [736,739,744]

EECS, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, U.S.A.; costis@csail.mit.edu,
Math, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, U.S.A.; deckel@alum.mit.edu,
and
EECS, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, U.S.A.; tzamos@mit.edu.

Co-editor Dirk Bergemann handled this manuscript.

Manuscript received 2 July, 2014; final version accepted 20 October, 2016; available online 15 November, 2016.