
Volume 146 Number 2 March 2011 ISSN 0022-0531
ELSEVIER
JOURNAL OF
Economic
Theory
Editors: Christian Hellwig and Karl Shell
Associate Editors
Marco Battaglini
Dirk Bergemann
Alberto Bisin
Yeon-Koo Che
In-Koo Cho
Olivier Compte
Philip H. Dybvig
Larry G. Epstein
Yossi Feinberg
Christian Ghiglino
Jacob Goeree
Piero Gottardi
Jean-Michel Grandmont
Hugo Hopenhayn
Johannes Horner
Vijay Krishna
Felix Kubler
Ricardo Lagos
Massimo Marinacci
Thomas Mariotti
David Martimort
Kiminori Matsuyama
Richard McLean
Giuseppe Moscarini
Efe Ok
Alessandro Pavan
James Peck
Lasse Heje Pedersen
Fabrizio Perri
Benjamin Polak
Stephen A. Ross
Hamid Sabourian
William H. Sandholm
Manuel S. Santos
Chris Shannon
Marciano Siniscalchi
Lones Smith
William L. Thomson
Juuso Välimäki
Randall Wright
Administrative Editors: Edward McDevitt and Susan Witherow Schulze
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
Journal of Economic Theory 146 (2011) 454–478
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

When is multidimensional screening a convex program? ☆

Alessio Figalli [a], Young-Heon Kim [b], Robert J. McCann [c,*]

[a] *Department of Mathematics, The University of Texas at Austin, Austin, TX 78712, USA*
[b] *Department of Mathematics, University of British Columbia, Vancouver, BC, V6T 1Z2 Canada*
[c] *Department of Mathematics, University of Toronto, Toronto, Ontario, M5S 2E4 Canada*

Received 21 May 2010; final version received 24 October 2010; accepted 25 November 2010

Available online 6 January 2011

Abstract

A principal wishes to transact business with a multidimensional distribution of agents whose preferences are known only in the aggregate. Assuming a twist (= generalized Spence–Mirrlees single-crossing) hypothesis, quasi-linear utilities, and that agents can choose only pure strategies, we identify a structural condition on the value $b(x, y)$ of product type y to agent type x — and on the principal's costs $c(y)$ — which is necessary and sufficient for reducing the profit maximization problem faced by the principal to a convex program. This is a key step toward making the principal's problem theoretically and computationally tractable; in particular, it allows us to derive uniqueness and stability of the principal's optimal strategy — and similarly of the strategy maximizing the expected welfare of the agents when the principal's profitability is constrained. We call this condition non-negative cross-curvature: it is also (i) necessary and sufficient to guarantee convexity of the set of b-convex functions, (ii) invariant under reparametrization of agent and/or

☆ The initial impetus for this study emerged from discussions with Ivar Ekeland. RJM is pleased to express his gratitude to Ekeland for introducing him to the principal–agent problem in 1996, and for anticipating already at that time that it ought to be tackled using techniques from the mathematical theory of optimal transportation. We are grateful to Giuseppe Buttazzo and Guillaume Carlier also, for stimulating discussions, and to Shibing Chen for a careful reading of the manuscript. The authors are pleased to thank the Institute for Pure and Applied Mathematics at UCLA and the Institut Fourier at Grenoble for their generous hospitality during various stages of this work. RJM's research was supported in part by NSERC grant 217006-08 and NSF grant DMS-0354729. YHK was partially supported by NSF grant DMS-0635607 through the membership at Institute for Advanced Study, Princeton NJ and NSERC discovery grant 371642-09. AF was partially supported by NSF grant DMS-0969962. Any opinions, findings and conclusions or recommendations expressed in this material are those of authors and do not reflect the views of the Natural Sciences and Engineering Research Council of Canada (NSERC) or of the US National Science Foundation (NSF).

* Corresponding author. Fax: +1 416 978 4107.
E-mail addresses: figalli@math.utexas.edu (A. Figalli), yhkim@math.ubc.ca (Y.-H. Kim), mccann@math.toronto.edu (R.J. McCann).

0022-0531/$ – see front matter
doi:10.1016/j.jet.2010.11.006

product types by diffeomorphisms, and (iii) a strengthening of Ma, Trudinger and Wang's necessary and sufficient condition (A3w) for continuity of the correspondence between an exogenously prescribed distribution of agents and of products. We derive the persistence of economic effects such as the desirability for a monopoly to establish prices so high they effectively exclude a positive fraction of its potential customers, in nearly the full range of non-negatively cross-curved models.

JEL classification: C61; D82; D42; H21; J42

Keywords: Principal–agent; Asymmetric information; Monopoly; Nonlinear pricing; Price discrimination; Multidimensional signalling; Screening; Social welfare maximization under budget constraint; Optimal taxation; Incentive compatibility; Mechanism design; Exclusion; Bunching; Robustness; Private imperfect incomplete information; Optimal transportation; Ma–Trudinger–Wang; Cross-curvature

1. Introduction

The principal–agent paradigm provides a microeconomic framework for modeling non-competitive decision problems which must be made in the face of informational asymmetry. Such problems range from monopolist nonlinear pricing [26,36,39,2] and product line design ("customer screening") [31], to optimal taxation [24], labour market signalling and contract theory [35,27], regulation of monopolies [4] including public utilities [28], and mechanism design [22,25]. A typical example would be the problem faced by a monopolist who wants to market automobiles $y \in Y$ to a population of potential buyers ("agents") $x \in X$. Knowing the value $b(x, y)$ of car y to buyer x, the relative frequency $d\mu(x)$ of different buyer types in the population, and the cost $c(y)$ she incurs in manufacturing car type y, the principal needs to decide which products (or product bundles) to manufacture and how much to charge for each of them, so as to maximize her profits.

In the simplest models there are only a finite number of product possibilities (e.g. with air conditioning, or without) and a finite number of buyer types (e.g. rich, middle-class, and poor); or possibly a one-dimensional continuum of product possibilities (parameterized, say, by quality) and of agent types (parameterized, say, by income) [24,35,26,4]. Of course, real cars depend on more than one parameter — fuel efficiency, comfort, options, reliability, styling, handling and safety, to name a few — as do car shoppers, who vary in wealth, income, age, commuting needs, family size, personal disposition, etc. Thus realistic modeling requires multidimensional type spaces $X \subset \mathbf{R}^m$ and $Y \subset \mathbf{R}^n$, as in [27,22,32,5,10]. Although such models can often be reduced to optimization problems in the calculus of variations [8,5], in the absence of convexity they remain dauntingly difficult to analyze. Convexity — whether manifest or hidden — rules out critical points other than global minima, and is often the key to locating and characterizing optimal strategies either numerically or theoretically. The purpose of the present article is to determine when convexity is present, assuming the dimensions $m = n$ of the agent and product type spaces coincide.

An archetypal model was addressed by Wilson [39], Armstrong [2], and Rochet and Choné [31]. A particular example from the last of these studies makes the simplifying hypotheses $X = Y = [0, \infty[^n$, $c(y) = |y|^2/2$, and $b(x, y) = \langle x, y\rangle$. By assuming this *bilinearity* of the buyers' valuations, Rochet and Choné were able to show that the principal's problem can be reduced to a quadratic minimization over the set of non-negative convex functions — itself a convex set. Although the convexity constraint makes this variational problem non-standard, for buyers distributed uniformly throughout the unit square in $\mathbf{R}^2$, they exploited a combination of theoretical and computational analysis to show a number of results of economic interest. Their most striking conclusion was that the profit motive alone leads the principal to discriminate between three different types of buyers: (i) low-end customers whom she will not market cars to, because — as Armstrong had already discovered — making cars affordable to this segment of the market would cost her too much of her mid-range and high-end profits; (ii) mid-range customers, whom she will encourage to choose from a one-parameter family of affordably-priced compromise vehicles; (iii) high-end customers, whom she will use both available dimensions of her product space to market expensive vehicles individually tailored to suit each customer's desires. Whether or not such *bunching* phenomena are robust is an unanswered question of considerable interest which — due to their specificity to particular valuation functions — the techniques of the foregoing authors remain unable to address. The possibility of non-robustness was highlighted in [5]; below we go further to suggest which specific perturbations of the valuation function $b(x, y)$ are most likely to yield robust results. On the other hand, our conclusions confirm Armstrong's assertion that what he called *the desirability of exclusion* is a very general phenomenon in the models we study (Theorem 4.8). This exclusion however, is less generic when the dimensions of the type and allocation spaces differ [10], or when their strict convexity fails [32]: see Deneckere and Severinov for a discussion of the case $(m, n) = (2, 1)$; see also [3].[1]

For general valuations $b(x, y)$, the principal's problem can be reformulated as a minimization problem over the space of b-convex functions (Definition 3.1), according to Carlier [8]. Such functions generally form a compact but non-convex set, which prevented Carlier from deducing much more than the existence of an optimal strategy for the principal — a result which can also be obtained using the method of Monteiro and Page [25] (for related developments see Basov [5] or Rochet and Stole [32]). Our present purpose is to identify conditions on the agent valuations which guarantee convexity of this feasible set (Theorem 3.2). In the setting we choose, the conditions we find will actually be necessary as well as sufficient for convexity; this necessity imparts a significance to these conditions even if they appear unexpected or unfamiliar. If, in addition, the principal's manufacturing cost $c(y)$ is b^*-convex, for $b^*(y, x) := b(x, y)$, the principal's problem becomes a convex program which renders it much more amenable to standard theoretical and computational techniques [11]. Although the resulting problem retains the complexities of the Wilson, Armstrong, and Rochet and Choné's models, we are able to deduce new results which remained inaccessible until now, such as conditions guaranteeing uniqueness (Theorem 4.6) and stability (Corollary 4.7) of the principal's optimal strategy. The same considerations and results apply also to the problem of maximizing the total welfare of the agents under the constraint that it remain possible for the principal to operate without sustaining a loss (Remark 5.1).

2. Hypotheses: the basic framework

As in Ma, Trudinger and Wang's work concerning the smoothness of optimal mappings [21], let us assume the buyer valuations satisfy the following hypotheses. Let $\overline{X}$ denote the closure of any given set $X \subset \mathbf{R}^n$, and for each $(x_0, y_0) \in \overline{X} \times \overline{Y}$ assume:

[1] A different robustness result concerning exclusion was found by Barelli, Basov, Bugarin and King [3], who relax the convexity assumption on the space of agents while allowing a fairly wide class of valuations $b(x, y)$. No hypothesis analogous to our **(B3)** below appears in their work, though they relax our convexity hypothesis **(B2)** considerably and work under a different hypothesis than **(B1)**. We are grateful to an anonymous referee, for bringing this work to our attention after the present manuscript had been submitted.

(B0) $b \in C^4(\overline{X} \times \overline{Y})$, where $X \subset \mathbf{R}^n$ and $Y \subset \mathbf{R}^n$ are open and bounded;

(B1) (bi-twist) $\left.\begin{array}{l} x \in \overline{X} \longmapsto D_y b(x, y_0) \\ y \in \overline{Y} \longmapsto D_x b(x_0, y) \end{array}\right\}$ are diffeomorphisms onto their ranges;

(B2) (bi-convexity) $\left.\begin{array}{l} X_{y_0} := D_y b(X, y_0) \\ Y_{x_0} := D_x b(x_0, Y) \end{array}\right\}$ are convex subsets of $\mathbf{R}^n$.

Here the subscript x_0 serves as a reminder that Y_{x_0} denotes a subset of the cotangent space $T^*_{x_0} X = \mathbf{R}^n$ to X at x_0. Note **(B1)** is strengthened form of the multidimensional generalization of the Spence–Mirrlees single-crossing condition expressed in separate works from the 1990s by Rüschendorf, Gangbo, and Levin; see e.g. [1]. It turns out to imply that the marginal utility of buyer type x_0 determines the product he selects uniquely and smoothly (cf. (4.2)), and similarly that buyer type who selects product y_0 will be a well-defined smooth function of y_0 and the marginal cost of that product; **(B1)** is much less restrictive than the generalized single crossing condition proposed by McAfee and McMillan [22], since the iso-price curves in the latter context become hyperplanes, effectively reducing the problem to a single dimension. Hypothesis **(B2)** turns out to be necessary (but not sufficient) for the convexity of the principal's optimization problem and strategy space.[2] We also assume

(B3) (non-negative cross-curvature)

$$\frac{\partial^4}{\partial s^2 \partial t^2}\bigg|_{(s,t)=(0,0)} b\big(x(s), y(t)\big) \geqslant 0 \tag{2.1}$$

whenever either of the two curves $s \in [-1, 1] \longmapsto D_y b(x(s), y(0))$ and $t \in [-1, 1] \longmapsto D_x b(x(0), y(t))$ forms an affinely parameterized line segment (in $\overline{X}_{y(0)} \subset \mathbf{R}^n$, or in $\overline{Y}_{x(0)} \subset \mathbf{R}^n$, respectively).

If the inequality (2.1) becomes strict whenever $x'(0)$ and $y'(0)$ are non-vanishing, we say the valuation function b is *positively cross-curved*, and denote this by $\mathbf{(B3)_u}$.[3]

[2] Necessity of the convexity of Y_{x_0} for that of $\mathcal{V}^b_Y$ in Theorem 3.2 was pointed out to us by Brendan Pass in his response to this manuscript. In the context of Spence–Mirrlees and Rochet–Choné type valuations (Examples 3.3–3.4 below), the convexity of Y_{x_0} permits the space Y of product types to be interpreted as representing randomized (mixed) strategies.

[3] We will eventually see that condition **(B3)** can alternately be characterized as in Lemma 4.3 using Definition 4.1; the convexity asserted by that lemma may appear more intuitive and natural than **(B3)** from point of view of applications. Historically, non-negative cross-curvature arose as a strengthening of Trudinger and Wang's criterion (A3w) guaranteeing smoothness of optimal maps in the Monge–Kantorovich transportation problem [37]; unlike us, they require (2.1) only if, in addition,

$$\frac{\partial^2}{\partial s \partial t}\bigg|_{(s,t)=(0,0)} b\big(x(s), y(t)\big) = 0. \tag{2.2}$$

Necessity of Trudinger and Wang's condition for continuity was shown by Loeper [19], who (like Trudinger, and independently Kim and McCann [16]) also noted its covariance and some of its relations to the geometric notion of curvature. Their condition relaxes a hypothesis proposed with Ma [21], which required strict positivity of (2.1) when (2.2) holds. The strengthening considered here was first studied in a different but equivalent form by Kim and McCann, where both the original and the modified conditions were shown to correspond to pseudo-Riemannian sectional curvature conditions induced by buyer valuations on $X \times Y$, thus highlighting their invariance under reparametrization of either X or Y by diffeomorphism; see Lemma 4.5 of [16]. The same lemma shows it costs no generality to require both curves

3. Results concerning the principal–agent problem

A mathematical concept of central relevance to us is encoded in the following definition.

Definition 3.1 *(b-convex).* A function $u : \overline{X} \longmapsto \mathbf{R}$ is called b-convex if $u = (u^{b^*})^b$, where

$$v^b(x) = \sup_{y \in \overline{Y}} b(x, y) - v(y) \quad \text{and} \quad u^{b^*}(y) = \sup_{x \in \overline{X}} b(x, y) - u(x). \tag{3.1}$$

In other words, if u is its own second b-transform, i.e. a supremal convolution (or generalized Legendre transform) of some function $v : \overline{Y} \longmapsto \mathbf{R} \cup \{+\infty\}$ with b. The set of b-convex functions will be denoted by $\mathcal{V}^b_{\overline{Y}}$. Similarly, we define the set $\mathcal{U}^{b^*}_{\overline{X}}$ of b^*-convex functions to consist of those $v : \overline{Y} \longmapsto \mathbf{R}$ satisfying $v = (v^b)^{b^*}$.

Although some authors permit b-convex functions to take the value $+\infty$, our hypothesis **(B0)** ensures b-convex functions are Lipschitz continuous and thus that the suprema defining their b-transforms are finitely attained. Our first result is the following.

Theorem 3.2 *(b-convex functions form a convex set). Assuming $b : X \times Y \longmapsto \mathbf{R}$ satisfies* **(B0)**–**(B2)***, hypothesis* **(B3)** *becomes necessary and sufficient for the convexity of the set $\mathcal{V}^b_{\overline{Y}}$ of b-convex functions on $\overline{X}$.*

To understand the relevance of this theorem to economic theory, let us recall a mathematical formulation of the principal–agent problem based on [8] and [29,30]. In this context, each product $y \in \overline{Y}$ costs the principal $c(y)$ to manufacture, and she is free to market this product to the population $\overline{X}$ of agents at any lower semicontinuous price $v(y)$ that she chooses. She is aware that product y has value $b(x, y)$ to agent $x \in \overline{X}$, and that in response to any price menu $v(y)$ she proposes, each agent will compute his indirect utility by combining his valuation for product y with its price quasi-linearly

$$u(x) = v^b(x) := \max_{y \in \overline{Y}} b(x, y) - v(y), \tag{3.2}$$

and will choose to buy a product $y_{b,v}(x)$ which attains the maximum, meaning $u(x) = b(x, y_{b,v}(x)) - v(y_{b,v}(x))$. However, let us assume that there is a distinguished point $y_\emptyset \in \overline{Y}$ representing the null product (or outside option), which the principal is compelled to offer to all agents at zero profit,

$$v(y_\emptyset) = c(y_\emptyset), \tag{3.3}$$

$s \in [-1, 1] \longmapsto D_y b(x(s), y(0))$ and $t \in [-1, 1] \longmapsto D_x b(x(0), y(t))$ to be line segments for (2.1) to hold. Other variants and refinements of Ma, Trudinger, and Wang's condition have been proposed and investigated by Figalli and Rifford and Loeper and Villani for different purposes at about the same time; see e.g. [17].

Kim and McCann showed non-negative cross-curvature guarantees tensorizability of condition **(B3)**, which is useful for building examples of valuation functions which satisfy it [17]; in suitable coordinates, it guarantees convexity of each b-convex function, as they showed with Figalli [13]; see Proposition 4.4. Hereafter we show, in addition, that it is necessary and sufficient to guarantee convexity of the set $\mathcal{V}^b_{\overline{Y}}$ of b-convex functions. A variant on the sufficiency was observed simultaneously and independently from us in a different context by Sei (Lemma 1 of [34]), who was interested in the function $b(x, y) = -d^2_{S^n}(x, y)$, and used it to give a convex parametrization of a family of statistical densities he introduced on the round sphere $X = Y = S^n$.

either because both quantities vanish (representing the null transaction), or because, as in [6], there is a competing supplier or regulator from whom the agents can obtain this product at price $c(y_\emptyset)$. In other words, $u_\emptyset(x) := b(x, y_\emptyset) - c(y_\emptyset)$ acts as the reservation utility of agent $x \in X$, below which he will reject the principal's offers and decline to participate, whence $u \geqslant u_\emptyset$. The map $y_{b,v} : \overline{X} \longmapsto \overline{Y}$ from agents to products they select will not be continuous except possibly if the price menu v is b^*-convex; when $y_{b,v}(x)$ depends continuously on $x \in \overline{X}$ we say v is *strictly* b^*-convex.

Knowing b, c and a (Borel) probability measure μ on X — representing the relative frequency of different types of agents in the population — the principal's problem is to decide which lower semicontinuous price menu $v : \overline{Y} \longmapsto \mathbf{R} \cup \{+\infty\}$ maximizes her profits, or equivalently, minimizes her net losses:

$$\int_X \big[c\big(y_{b,v}(x)\big) - v\big(y_{b,v}(x)\big)\big]\, d\mu(x). \tag{3.4}$$

Note the integrand vanishes (3.3)–(3.4) for any agent x who elects not to participate (i.e., who chooses the outside option $y_\emptyset \in \overline{Y}$).

For absolutely continuous distributions of agents — or more generally if μ vanishes on Lipschitz hypersurfaces — it is known that the principal's losses (3.4) depend on v only through the indirect utility $u = v^b$, an observation which can be traced back to Mirrlees [24] in one dimension and Rochet [29] more generally; see also Carlier [8]. This indirect utility $u \geqslant u_\emptyset$ is b-convex, due to the well-known identity $((v^b)^{b^*})^b = v^b$ (e.g. Proposition 5.8 of [38]). Conversely, the principal can design any b-convex function $u \geqslant u_\emptyset$ that she wishes simply by choosing price strategy $v = u^{b^*}$. Thus, as detailed below, the principal's problem can be reformulated as a minimization problem (4.5) on the set $\mathcal{U}_0 := \{u \in \mathcal{V}^b_{\overline{Y}} \mid u \geqslant u_\emptyset\}$. Under hypotheses **(B0)–(B3)**, our Theorem 3.2 shows the set $\mathcal{V}^b_{\overline{Y}}$ of such utilities u to be convex, in the usual sense. This represents substantial progress, even though the minimization problem (3.4) still depends nonlinearly on $v = u^{b^*}$. If, in addition, the principal's cost $c(y)$ is a b^*-convex function, then Proposition 4.4 and its corollary show her minimization problem (3.4) becomes a convex functional of u on $\mathcal{U}_0$, so the principal's problem reduces to a convex program. Necessary and sufficient conditions for a minimum can in principle then be expressed using Kuhn–Tucker type conditions, and numerical examples could be solved using standard algorithms. However we do not do this here: unless μ is taken to be a finite combination of Dirac masses, the infinite dimensionality of the convex set $\mathcal{V}^b_{\overline{Y}}$ leads to functional analytic subtleties even for the bilinear valuation function $b(x, y) = \langle x, y\rangle$, which have only been resolved with partial success by Rochet and Choné in that case [31]. If the b^*-convexity of $c(y)$ is *strict* however, or if the valuation function is positively cross-curved $\mathbf{(B3)_u}$, we shall show the principal's program has enough strict convexity to yield unique optimal strategies for both the principal and the agents in a sense made precise by Theorem 4.6. These optimal strategies represent a Stackelberg (rather than a Nash) equilibrium, in the sense that no party has any incentive to change his or her strategies, given that the principal must commit to and declare her strategy before the agents select theirs.

Of course, it is of practical interest that the principal be able to anticipate not only her optimal price menu $v : \overline{Y} \longmapsto \mathbf{R} \cup \{+\infty\}$ — also known as the *equilibrium* prices — but the corresponding distribution of goods which she will be called on to manufacture. This can be represented as a Borel probability measure ν on $\overline{Y}$, which we call the *optimal production measure*. It quantifies the relative frequency of goods to be produced, and is the image of μ under the agents' best response function $y_{b,v} : \overline{X} \longmapsto \overline{Y}$ to the principal's optimal strategy v. This image $\nu = (y_{b,v})_{\#}\mu$

is a Borel probability measure on $\overline{Y}$ known as the *push-forward* of μ by $y_{b,v}$, and is defined by the formula

$$\nu(W) := \mu\big[y_{b,v}^{-1}(W)\big] \tag{3.5}$$

for each $W \subset Y$. Theorem 4.6 asserts the optimal production measure ν is unique and the optimal price menu v is uniquely determined ν-a.e.; the same theorem gives a sharp lower bound for v throughout $\overline{Y}$. If the convex domain $X_{y_\emptyset}$ is strictly convex and the density of agents is Lipschitz continuous on X, Theorem 4.8 goes on to assert that these prices will be high enough to drive a positive fraction of agents out of the market, extending Armstrong's desirability of exclusion [2] to a rich class of multidimensional models. Thus the goods to be manufactured and their prices are uniquely determined at equilibrium, and the principal can price the goods she prefers not to trade arbitrarily high but not arbitrarily low. Theorem 4.6 goes on to assert that the optimal strategy $y_{b,v}(x)$ is also uniquely determined for μ-almost every agent x by b, c and μ, for each Borel probability measure μ on X. Apart from Theorem 4.8, these conclusions apply to singular and discrete measures as well as to continuous measures μ, assuming the tie-breaking conventions of Remark 4.2 are adopted whenever μ fails to vanish on each Lipschitz hypersurface.

A number of examples of valuation functions $b(x, y)$ which satisfy our hypotheses are developed in the works by Delanoë, Ge, Figalli, Kim, Lee, Li, Loeper, Ma, McCann, Rifford, Trudinger and Wang; see [17] and [38] for references. Here we mention a few which have relevance to economics:

Example 3.3. For single dimensional type and allocation spaces $n = 1$, hypotheses **(B1)–(B2)** are equivalent to asserting that the valuation function $b(x, y)$ be defined on a product of two intervals where its cross-partial derivatives $D^2_{xy}b$ do not vanish. Positive cross-curvature $\mathbf{(B3)_u}$ asserts that $D^2_{xy}b$ in turn satisfies a Spence–Mirrlees condition, by having positive cross-partial derivatives: $D^2_{xy}(D^2_{xy}b) > 0$.

Example 3.4. The bilinear valuation function $b(x, y) = x \cdot y$ of Armstrong, Rochet and Choné satisfies **(B0)–(B3)** provided only that $X, Y \subset \mathbf{R}^n$ are convex bodies. In this case b-convexity coincides with ordinary convexity of u together with a constraint on its gradient's range: $Du(x) \in \overline{Y}$ at each point where u is differentiable. Thus Theorem 4.6 asserts that any strictly convex manufacturing cost $c(y)$ leads to unique optimal strategies for the principal and for μ-almost every agent. This uniqueness is well known for absolutely continuous measures $d\mu \ll d\,\mathrm{vol}$ [31], and Carlier and Lachand–Robert have extended Mussa and Rosen's differentiability result $u \in C^1(\overline{X})$ to $n \geqslant 1$ in that case [9,26], but the uniqueness of optimal strategies under the tie-breaking rules described in Remark 4.2 may be new results when applied, for example, to discrete distributions μ concentrated on finitely many agent types.

Example 3.5. Ma, Trudinger and Wang's perturbation $b(x, y) = x \cdot y + F(x)G(y)$ of the bilinear valuation function is non-negatively cross-curved **(B3)** provided $F \in C^4(\overline{X})$ and $G \in C^4(\overline{Y})$ are both convex [21,16]; it is positively cross-curved if the convexity is strong, meaning both $F(x) - \epsilon|x|^2$ and $G(y) - \epsilon|y|^2$ remain convex for some $\epsilon > 0$. It satisfies **(B0)–(B1)** provided $\sup_{x \in X} |DF(x)| < 1$ and $\sup_{y \in Y} |DG(y)| < 1$, and **(B2)** if the convex domains X and $Y \subset \mathbf{R}^n$ are sufficiently convex, meaning all principal curvatures of these domains are sufficiently large at each boundary point [21]. On the other hand, $b(x, y) = x \cdot y + F(x)G(y)$ will violate **(B3)** if $D^2F(x_0) > 0$ holds but $D^2G(y_0) \geqslant 0$ fails at some $(x_0, y_0) \in \overline{X} \times \overline{Y}$.

Example 3.6 *(On geometry in spatial economics).* Consider now a valuation function such as $b(x, y) = -\frac{1}{2}|x - y|^2$, modeling a family of buyers $X \subset \mathbf{R}^n$ each of whom prefers to choose products corresponding as closely as possible to their own type. Such a function might model a geographical distribution of otherwise identical consumers who must decide whether to pay a high fee to have a certain service (such as deliveries or waste removal) provided directly to their home, an intermediate fee to obtain this service at a nearby depot, or no fee to obtain the service from a more remote source $y_\emptyset = 0$ (such as a centralized warehouse in the first case or a public landfill in the second). If the monopolist's costs for providing this service are independent of location $c(y) = \text{const}$, then the problem becomes mathematically equivalent to Rochet and Choné's (Example 3.4 above): each b-convex strategy u satisfies $D^2 u \geqslant -I$, and corresponds to the convex strategy $u(x) + |x|^2/2$ of [31]. From their results concerning product and buyer types in the unit disc or unit square X, we infer that only buyers sufficiently far from the source $y_\emptyset$ can be induced to pay a positive price for the convenience of obtaining this service nearby. In the case of the square, there will be a strip of buyer types who select delivery points concentrated along the diagonal, followed by a region of buyers who pay more for individually customized delivery points. Since the products consumed lie in a subset of the square, the constraints $Du \in [0, \infty[^2$ of [31] are not binding, allowing us to take $Y = \mathbf{R}^2$ instead of $Y = [0, \infty[^2$. This in turn allows us to reflect their solution in both the horizontal and vertical axes, to yield an example in which the outside option $y_\emptyset$ lies in the center of a two-by-two square, rather than the corner of a one-by-one square. Comparison with the case of the unit disc X centered at $y_\emptyset$, for which the solution is rotationally symmetric, shows the bunching observed by Rochet and Choné depends dramatically on the domain's geometry, and its lack of strict convexity.

Now consider instead the possibility that the geographical region X, instead of being flat, is situated either at the bottom of a valley, or at a pass in the mountains, and that the valuation function $b(x, y) = -\frac{1}{2}d(x, y)^2$ reflects this geography, by depending on the distance d as measured along a spherical cap in the case of the valley or along a piece of a saddle in the case of the mountain pass. Then **(B0)**–**(B1)** are both satisfied, and **(B2)** will be too provided the domain $X = Y$ is convex enough... consisting for example of all points sufficiently close to the free source $y_\emptyset$ in the spherical or saddle geometry. According to results of Loeper [19,20] and Kim and McCann [17], **(B3)** will be satisfied in the case of the sphere [17], but violated in the case of a saddle [19]. Thus for a town in a valley (or on a mountain top), our results show that the screening problem remains convex, while for a town located on a pass in the mountains the problem becomes non-convex. This calls into question the uniqueness, stability, and structure of its solution(s) in the latter case, and displays how geometry and geography can affect the solubility of economic problems.

In the next section we formulate the results mathematically. Let us first highlight a further implication of our results concerning robustness of the phenomena observed by Rochet and Choné. The quadratic functions $b(x, y) = x \cdot y$ and $b(x, y) = -\frac{1}{2}|x - y|^2$ both lie on the boundary of the set of non-negatively cross-curved valuations, since their cross-curvatures (2.1) vanish identically. Our results show non-negative cross-curvature **(B3)** to be a necessary and sufficient condition for the principal–agent problem to be a convex program: the feasible set $\mathcal{V}^b_{\overline{Y}}$ becomes non-convex otherwise, and it is reasonable to expect that uniqueness of the solution among other phenomena observed in [31] may be violated in that case. In analogy with the discontinuities discovered by Loeper [19], we therefore conjecture that the bundling discovered by Rochet and Choné is robust with respect to perturbations of the quadratic valuation functions which respect **(B0)–(B3)**, but not generally with respect to perturbations violating **(B3)**. (See [3] however, for a different robustness result.)

4. Mathematical formulation

Any price menu $v : \overline{Y} \longmapsto \mathbf{R} \cup \{\infty\}$ satisfies

$$v^b(x) + v(y) - b(x, y) \geqslant 0 \tag{4.1}$$

for all $(y, x) \in \overline{Y} \times \overline{X}$, according to definition (3.1). Comparison with (3.2) makes it clear that a (product, agent) pair produces equality in (4.1) if and only if selecting product y is among the best responses of agent x to this menu; the set of such best-response pairs is denoted by $\partial^{b^*} v \subset \overline{Y} \times \overline{X}$; see also (A.2). We think of this relation as giving a multivalued correspondence between products and agents: given price menu v the set of agents (if any) willing to select product y is denoted by $\partial^{b^*} v(y)$. It turns out $\partial^{b^*} v(y)$ is non-empty for all $y \in \overline{Y}$ if and only if v is b^*-convex. Thus b^*-convexity of v — or of c — means precisely that each product is priced low enough to be included among the best responses of some agent or limiting agent type $x \in \overline{X}$. As we shall see in Remark 4.2, assuming b^*-convexity of v costs little or no generality; however, the b^*-convexity of c is a real restriction — but plausible when the product types $Y \subset \mathbf{R}^n$ represent mixtures (weighted combinations of pure products) which the principal could alternately choose to purchase separately and then bundle together; this becomes natural in the context of the bilinear valuation $b(x, y) = x \cdot y$ assumed by Rochet and Choné [31].

Let $\operatorname{Dom} Du \subset \overline{X}$ denote the set where u is differentiable. If y is among the best responses of agent $x \in \operatorname{Dom} Dv^b$ to price menu v, the equality in (4.1) implies

$$Dv^b(x) = D_x b(x, y). \tag{4.2}$$

In other words $y = y_b(x, Dv^b(x))$, where y_b is defined as follows:

Definition 4.1. For each $q \in \overline{Y}_x$, define $y_b(x, q)$ to be the unique product selected by an agent $x \in \overline{X}$ whose marginal utility with respect to his type is q; i.e. y_b is the unique solution to

$$D_x b\big(x, y_b(x, q)\big) = q \tag{4.3}$$

guaranteed by **(B1)**. The map y_b (which is defined on a subset of the cotangent bundle $T^*\overline{X}$ and takes values in $\overline{Y}$) has also been called the b-exponential map [19], and denoted by $y_b(x, q) = b\text{-}Exp_x q$.

The fact that the best response function takes the form $y = y_b(x, Dv^b(x))$, and that $\operatorname{Dom} Dv^b$ exhausts $\overline{X}$ except for a countable number of Lipschitz hypersurfaces, are key observations exploited throughout both the economic and optimal transportation literature. Indeed, v^b is well known to be a b-convex function. It is therefore Lipschitz and semiconvex, satisfying the bounds

$$\big|Dv^b\big| \leqslant \|c\|_{C^1(X \times Y)}, \qquad D^2 v^b \geqslant -\|c\|_{C^2(X \times Y)} \quad \text{inside } \overline{X}. \tag{4.4}$$

The second inequality above holds in the distributional sense, and implies the differentiability of v^b outside a countable number of Lipschitz hypersurfaces [14].

Assuming μ assigns zero mass to each Lipschitz hypersurface (and so also to a countable number of them), the results just summarized allow the principal's problem (3.4) to be re-expressed in the form $\min\{L(u) \mid u \in \mathcal{U}_0\}$, where the principal's net losses are given by

$$L(u) := \int_X \big[u(x) + c\big(y_b\big(x, Du(x)\big)\big) - b\big(x, y_b\big(x, Du(x)\big)\big)\big]\, d\mu(x) \tag{4.5}$$

as is by now well known [8]. Here $\mathcal{U}_0 = \{u \in \mathcal{V}^b_{\overline{Y}} \mid u \geqslant u_\emptyset\}$ denotes the set of b-convex functions on $\overline{X}$ dominating the reservation utility $u_\emptyset(x) = b(x, y_\emptyset) - c(y_\emptyset)$, and the equality produced in (4.1) by the response $y_{b,v}(x) = y_b(x, Dv^b(x))$ for μ-a.e. x has been exploited. Our hypothesis on the distribution of agent types holds a fortiori whenever μ is absolutely continuous with respect to Lebesgue measure on X. If no such hypothesis is satisfied, the reformulation (4.5) of the principal's net losses may not be well defined, unless we extend the definition of $Du(x)$ to all of X by making a measurable selection from the relation

$$\partial u(x) := \big\{q \in \mathbf{R}^n \bigm| u(z) \geqslant u(x) + q \cdot (z - x) + o\big(|z - x|\big)\ \forall z \in X\big\}$$

consistent with the following tie-breaking rule, analogous to one adopted, e.g., by Buttazzo and Carlier in a similar context [6]:

Remark 4.2 *(Tie-breaking rules for singular measures).* When an agent x remains indifferent between two or more products, it is convenient to reduce the ambiguity in the definition of his best response by insisting that $y_{b,v}(x)$ be chosen to maximize the principal's profit $v(y) - c(y)$, among those products y which maximize (3.2). We retain the result $y_{b,v}(x) = y_b(x, Dv^b(x))$ by a corresponding (measurable) selection $Dv^b(x) \in \partial v^b(x)$. This convention costs no generality when the distribution μ of agent types vanishes on Lipschitz hypersurfaces in X, since $u = v^b$ is then differentiable μ-a.e.; in the remaining cases it may be justified by assuming the principal has sufficient powers of persuasion to sway an agent's choice to her own advantage whenever some indifference would otherwise persist between his preferred products [24]. After adopting this convention, it costs the principal none of her profits to restrict her choice of strategies to b^*-convex price menus $v = (v^b)^{b^*}$, a second convention we also choose to adopt whenever μ fails to vanish on each Lipschitz hypersurface.

The relevance of Theorem 3.2 to the principal–agent problem should now be clear: it guarantees convexity of the feasible set $\mathcal{U}_0$ in (4.5). Our next proposition addresses the convexity properties of the principal's objective functional (i.e. her net losses). Should convexity of this objective be strict, then the best response $y_{b,v}(x)$ selected by the tie-breaking rule above becomes unique — which it need not be otherwise.

We precede this proposition with a lemma containing a more intuitive characterization of non-negative cross-curvature found by two of us [17]. After **(B2)** and the second part of hypothesis **(B1)** are used to define $y_b(x, q)$, notice the first part of **(B1)** becomes equivalent to the absence of critical points for the functions (4.6) of q. Inspired by Loeper's characterization [19] of (A3w), the next lemma asserts **(B3)** is equivalent to convexity of these non-critical functions.

Lemma 4.3 *(Characterizing non-negative cross-curvature [17]). A valuation b satisfying* **(B0)**–**(B2)** *is non-negatively cross-curved* **(B3)** *if and only if for each $x \neq x_1$ in $\overline{X}$,*

$$q \in \overline{Y}_x \longmapsto b\big(x_1, y_b(x, q)\big) - b\big(x, y_b(x, q)\big) \tag{4.6}$$

is a convex function. If the valuation is positively cross-curved, then (4.6) *will be strongly convex* (*meaning its Hessian will be positive definite*).

This lemma plays a key role in establishing the proposition which follows.

Proposition 4.4 *(Convexity of the principal's objective). If $b \in C^4(\overline{X} \times \overline{Y})$ satisfies* **(B0)** –**(B3)** *and $c : \overline{Y} \longmapsto \mathbf{R}$ is b^*-convex, then for each $x \in \overline{X}$, definition* (4.3) *makes $a(q) := c(y_b(x, q)) -$*

$b(x, y_b(x,q))$ a convex function of q on the convex set $\overline{Y}_x := D_x b(x, \overline{Y}) \subset \mathbf{R}^n$. This convexity is strict (i) *if c is strictly b^*-convex, i.e., if* $\operatorname{Dom} Dc^b = \overline{X}$*, or equivalently, if the allocation $y_{b,c} : \overline{X} \longmapsto \overline{Y}$ is continuous. Alternately, this convexity is strict* (ii) *for each fixed $x \in \operatorname{Dom} Dc^b$ such that*

$$q \in \overline{Y}_x \longmapsto b\big(x_0, y_b(x,q)\big) - b\big(x, y_b(x,q)\big) \tag{4.7}$$

is a strictly convex function of q for all $x_0 \in \overline{X} \setminus \{x\}$.

The strict convexity of (4.7) may subsequently be denoted by **(B3)$_\mathbf{s}$**, a condition which Lemma 4.3 shows to be intermediate in strength between non-negative cross-curvature **(B3)** and positive cross-curvature **(B3)$_\mathbf{u}$**. As an immediate corollary to Theorem 3.2 and Proposition 4.4, we have convexity of the principal's optimization problem.

Corollary 4.5 *(Convexity of the principal's minimization). Let the distribution of agent types be given by a Borel probability measure μ on $X \subset \mathbf{R}^n$. Unless μ vanishes on all Lipschitz hypersurfaces, adopt the tie-breaking conventions of Remark* 4.2*. If the value $b(x,y)$ of product $y \in \overline{Y}$ to agent $x \in \overline{X}$ satisfies* **(B0)–(B3)** *and the principal's manufacturing cost $c : \overline{Y} \longmapsto \mathbf{R}$ is b^*-convex, then the principal's problem* (4.5) *becomes a convex minimization over the convex set $\mathcal{U}_0$.*

As a consequence, we obtain criteria guaranteeing uniqueness of the principal's best strategy.

Theorem 4.6 *(Criteria for uniqueness of optimal strategies). Assume the notation and hypotheses of Corollary* 4.5*. Suppose, in addition, either* (i) *that the manufacturing cost c is strictly b^*-convex, or else* (ii) *that the valuation function b satisfies the strengthened hypothesis* **(B3)$_\mathbf{s}$** *of* (4.7)*. In case* (ii) *assume also* $\mu[\operatorname{Dom} Dc^b] = 1$ (*it holds automatically unless μ concentrates mass on some Lipschitz hypersurface*)*. Then the equilibrium response of μ-almost every agent is uniquely determined, as is the optimal measure ν from* (3.5); (*always assuming the tie-breaking conventions of Remark* 4.2 *to be in effect if μ does not vanish on each Lipschitz hypersurface*)*. Moreover, the principal has two optimal strategies $u_\pm \in \mathcal{U}_0$ which coincide at least μ-almost everywhere, and sandwich all other optimal strategies $u \in \mathcal{U}_0$ between them: $u_- \leqslant u \leqslant u_+$ on $\overline{X}$. Finally, a lower semicontinuous $v : \overline{Y} \longmapsto \mathbf{R} \cup \{+\infty\}$ is an optimal price menu if and only if $v \geqslant u_+^{b^*}$ throughout $\overline{Y}$, with equality holding ν-almost everywhere.*

This theorem gives hypotheses which guarantee — even for discrete measures μ corresponding to finitely many agent types — that the solution to the principal's problem is unique in the sense that optimality determines how many of each type of product the principal should manufacture, what price she should charge for each of them, and which product will be selected by almost every agent. A lower bound is specified on the price of each product which she does not wish to produce, to ensure that it does not tempt any agent. When μ vanishes on Lipschitz hypersurfaces, this solution represents the only Stackelberg equilibrium balancing the interests of the principal with those of the agents; for more singular μ, it is possible that other Stackelberg equilibria exist, but if so they violate the restrictions imposed on the behaviour of the principal and the agents in Remark 4.2.

The uniqueness theorem has as its corollary the following stability result concerning optimal strategies. Recall that a sequence $\{\mu_i\}_{i=1}^\infty$ of Borel probability measures on a compact set $\overline{X} \subset \mathbf{R}^n$ is said to converge *weakly-$*$* to μ_∞ if

$$\int_{\overline{X}} g(x)\,d\mu_\infty(x) = \lim_{i\to\infty} \int_{\overline{X}} g(x)\,d\mu_i(x) \tag{4.8}$$

for each continuous test function $g : \overline{X} \longmapsto \mathbf{R}$. This notion of convergence makes the Borel probability measures $\mathcal{P}(\overline{X})$ on $\overline{X}$ into a compact set, as a consequence of the Riesz–Markov and Banach–Alaoglu theorems.

Corollary 4.7 *(Stability of optimal strategies). For each $i \in \mathbf{N} \cup \{\infty\}$, let the triple (b_i, c_i, μ_i) consist of a valuation $b_i : \overline{X} \times \overline{Y} \longmapsto \mathbf{R}$, manufacturing cost $c_i : \overline{Y} \longmapsto \mathbf{R}$, and a distribution of agent types μ_i on X satisfying the hypotheses of Theorem* 4.6. *Let $u_i : \overline{X} \longmapsto \mathbf{R}$ denote a b_i-convex utility function minimizing the losses of a principal faced with data (b_i, c_i, μ_i). Suppose that $b_i \to b_\infty$ in $C^2(\overline{X} \times \overline{Y})$, $c_i \to c_\infty$ uniformly on $\overline{Y}$, and $\mu_i \rightharpoonup \mu_\infty$ weakly-$*$ as $i \to \infty$. Assume finally that μ_∞ vanishes on all Lipschitz hypersurfaces. For μ_∞-a.e. agent $x \in X$, the product $G_i(x) := y_{b_i}(x, Du_i(x))$ selected then converges to $G_\infty(x)$. The optimal measures $\nu_i := (G_i)_\#\mu_i$ converge weakly-$*$ to ν_∞ as $i \to \infty$. And the principal's strategies converge uniformly in the sense that* $\lim_{i\to\infty} \|u_i - u_\infty\|_{L^\infty(X, d\mu_\infty)} = 0$.

Finally as evidence for the robustness of bunching phenomena displayed by our models, we show the desirability of exclusion phenomenon found by Armstrong for valuations $b(x, y) = \sum_{i=1}^n x_i b_i(y)$ which are linear — or more generally homogeneous of degree one — in x [2], extends to the full range of non-negatively cross-curved models. We assume strict convexity on the domain $X_{y_\emptyset} := D_y b(X, y_\emptyset)$ (see Remark 4.9), and that the distribution of agent types $d\mu(x) = f(x)dx$ has a Sobolev density — denoted $f \in W^{1,1}(\overline{X})$ and meaning both the function and its distributional derivative Df are given by Lebesgue integrable densities. This is satisfied a fortiori if f is Lipschitz or continuously differentiable (as Armstrong assumed). The exclusion phenomenon is of interest, since it confirms that a positive fraction of customers must be excluded from participation at equilibrium, thus ensuring elasticity of demand.

Theorem 4.8 *(The desirability of exclusion). Let the distribution $d\mu(x) = f(x)dx$ of agent types be given by a density $f \in W^{1,1}$ on $\overline{X} \subset \mathbf{R}^n$. Assume that the value $b(x, y)$ of product $y \in \overline{Y}$ to agent $x \in \overline{X}$ satisfies* **(B0)**–**(B3)** *and the principal's manufacturing cost $c : \overline{Y} \longmapsto \mathbf{R}$ is b^*-convex. Suppose further that the convex domain $X_{y_\emptyset} = D_x b(X, y_\emptyset)$ has no $n-1$ dimensional facets in its boundary. Then any minimizer $u \in \mathcal{U}_0$ of the principal's losses* (4.5) *coincides with the reservation utility on a set $U_0 := \{x \in \overline{X} \mid u(x) = b(x, y_\emptyset) - c(y_\emptyset)\}$ whose interior contains a positive fraction of the agents. Such agents select the outside option $y_\emptyset$.*

Remark 4.9 *(Facets and exclusion in different dimensions).* A convex domain $X \subset \mathbf{R}^n$ fails to be strictly convex if it has line segments in its boundary. These segments belong to facets of dimension 1 or higher, up to $n-1$ if the domain has a flat side (meaning a positive fraction of its boundary coincides with a supporting hyperplane). Thus strict convexity of $X_{y_\emptyset}$ is sufficient for the hypothesis of the preceding theorem to be satisfied — except in dimension $n = 1$. In a single dimension, every convex domain $X \subset \mathbf{R}$ is an interval — hence strictly convex — whose endpoints form zero-dimensional facets. Thus Theorem 4.8 is vacuous in dimension $n = 1$, which is consistent with Armstrong's observation the necessity of exclusion is a hallmark of higher dimensions $n \geqslant 2$. More recently, Deneckere and Severinov [10] have argued that necessity of exclusion is specific to the case in which the dimensions m and n of agent and product types coincide. When $(m, n) = (2, 1)$ they give necessary and sufficient conditions for the desirability of exclusion, yielding a result quite different from ours in that exclusion turns out to be more frequently the exception than the rule. However, for another setting in which exclusion is generic, see [3].

5. Discussion, extension, and conclusions

The role of private information in determining market value has a privileged place in economic theory. This phenomenon has been deeply explored in the principal–agent framework, where a single seller (or single buyer) transacts business with a collection of anonymous agents. In this context, the private (asymmetric) information takes the form of a characteristic $x \in X$ peculiar to each individual buyer which determines his valuation $b(x, y)$ for different products $y \in Y$ offered by the principal; x remains concealed from the principal by anonymity of the buyer — at least until a purchase is made. Knowing only the valuation function $b(x, y)$, the statistical distribution $d\mu(x)$ of buyer types, and her own manufacturing costs $c(y)$, the principal's goal is to fix a price menu for different products which maximizes her profits.

Many studies involving finite spaces of agent and product types X and Y have been carried out, including Spence's initial work on labour market signalling. However for a principal who transacts business with a one-dimensional continuum of agents $X \subset \mathbf{R}$, the problem was solved in Mirrlees' celebrated work on optimal taxation [24], and in Spence's study [35], assuming the contract types $y \in Y \subset \mathbf{R}$ are also parameterized by a single real variable. (For Mirrlees, $y \in \mathbf{R}$ represented the amount of labour an individual chooses to do facing a given tax schedule, while for Spence it represented the amount of education he chooses to acquire facing a given range of employment possibilities, $x \in \mathbf{R}$ being his intrinsic ability in both cases.) In the context of nonlinear pricing discussed above, the one-dimensional model was studied by Mussa and Rosen [26]. The challenge of resolving the multidimensional version $X, Y \subset \mathbf{R}^n$ of this archetypal problem in microeconomic theory has been highlighted by many authors [27,22,32,5]. When only one side of the market displays multidimensional types, analyses have been carried out by Mirman and Sibley [23], Roberts [28] and Spence [36], who allow multidimensional products, and by Laffont, Maskin and Rochet [18], and Deneckere and Severinov [10] who model two-dimensional agents choosing from a one-dimensional product line. When both sides of the market display multidimensional types, existence of an equilibrium has been established by Monteiro and Page [25] and by Carlier [8], who employed a variational formulation; see also the control-theoretic approach of Basov [5]. However, non-convexities have rendered the behaviour of this optimization problem largely intractable [15] — unless the valuation function $b(x, y) = x \cdot G(y)$ is assumed to depend linearly on agent type [39,2,31]. Moreover, the presence of convexity typically depends on a correct choice of coordinates, so is not always easy to discern. The present study treats general Borel probability measures μ on $X \subset \mathbf{R}^n$, and provides a unified framework for dealing with discrete and continuous type spaces, by invoking the tie-breaking rules of Remark 4.2 in case μ is discrete. Assuming b^*-convexity of c, we consider valuations linear in price (3.2) (sometimes called quasilinear), which satisfy a generalized Spence–Mirrlees single crossing condition **(B0)**–**(B1)** and appropriate convexity conditions on its domain **(B2)**, and we identify a criterion **(B3)** equivalent to convexity of the principal's optimization problem (Theorem 3.2). This criterion is a strengthening of Ma, Trudinger and Wang's necessary [19] and sufficient [21,37] condition for continuity of optimal mappings. Like all of our hypotheses, it is independent of the choice of parameterization of agent and/or product types — as emphasized in [16]. We believe the resulting convexity is a fundamental property which will eventually enable a more complete theoretical and computational analysis of the multidimensional principal–agent problem, and we indicate some examples of valuation functions which satisfy it in Examples 3.3–3.6; the bilinear example $b(x, y) = x \cdot y$ of Rochet and Choné lies on the boundary of such valuation functions. If either the cross-curvature inequality **(B3)** holds strictly or the b^*-convexity of $c(y)$ is strict — meaning the *efficient* solution $y_{b,c}(x)$ depends continuously on $x \in \overline{X}$ — we go on to derive uniqueness and stability of optimal strategies (Theorem 4.6 and its corollary). Under mild additional hypotheses we confirm that a positive fraction of agents must be priced out of the market when the type spaces are multidimensional (Theorem 4.8). We conjecture that non-negative cross-curvature **(B3)** is likely to be necessary and sufficient for robustness of Armstrong's desirability of exclusion [2] and the other bunching phenomena observed by Rochet and Choné [31].

Remark 5.1 *(Maximizing social welfare under profitability constraints).* Before concluding this paper, let us briefly mention an important class of related models to which the same considerations apply: namely, the problem of maximizing the expected welfare of the agents under a profitability constraint on the principal. Such a model has been used by Roberts [28] to study energy pricing by a public utility, and explored by Spence [36] and Monteiro and Page [25] in other contexts. Suppose the welfare of agent $x \in X$ is given by a function $w(x, u(x))$ of his indirect utility (3.2) which is concave with respect to its second variable: $\partial^2 w/\partial u^2 \leqslant 0$. Introducing a Lagrange multiplier λ for the profitability constraint $L(u) \leqslant 0$, the problem of maximizing the net social welfare over all agents becomes equivalent to the maximization

$$W(\lambda) := \max_{u \in \mathcal{U}_0} -\lambda L(u) + \int_X w\big(x, u(x)\big)\, d\mu(x)$$

for some choice of $\lambda \geqslant 0$. Assuming **(B0)–(B3)**, and b^*-convexity of c, for each $\lambda \geqslant 0$ this amounts to a concave maximization on a convex set, as a consequence of Theorem 3.2, Proposition 4.4 and the concavity of w. Theorem 4.6 and its corollary give hypotheses which guarantee uniqueness and stability of its solution u_λ; if the concavity of $w(x, \cdot)$ is strict, we obtain uniqueness μ-a.e. of u_λ more directly under the weaker hypotheses of Corollary 4.5. Either way, once the uniqueness of u_λ has been established, standard arguments in the calculus of variations show the convex function $W(\lambda)$ to be continuously differentiable, and that each value of its derivative $W'(\lambda) = -L(u_\lambda)$ corresponds to a possibly degenerate interval $\lambda \in [\lambda_1, \lambda_2]$ on which u_λ is constant; see e.g. Corollary 2.11 of [7]. Uniqueness of a social welfare maximizing strategy subject to any budget constraint in the range $]L(u_0), L(u_\infty)[$ is therefore established; this range contains the vanishing budget constraint as long as $L(u_0) > 0 > L(u_\infty)$; here u_0 represents the unconstrained maximizer whereas $u_\infty \in \mathcal{U}_0$ minimizes the principal's losses (4.5). All of our results — except for the desirability of exclusion (Theorem 4.8) — extend immediately to this new setting. This sole exception is in accord with the intuition that it need not be necessary to exclude any potential buyers if one aims to maximize social welfare instead of the monopolist's profits.

6. Proofs

The first sentence of Lemma 4.3 comes from Theorem 2.11 of [17]. We recall its proof partly for the sake of completeness, but also to establish the second sentence, which asserts strong convexity.

Proof of Lemma 4.3. Fixing $x, x_1 \in \overline{X}$ and $q_0, q_1 \in \overline{Y}_x$, the second claim in **(B2)** guarantees the line segment $q_t := (1 - t)q_0 + tq_1$ belongs to $\overline{Y}_x$. Use (4.3) to define $y_t := y_b(x, q_t)$ and

$f(\cdot,t) := b(\cdot, y_t) - b(x, y_t)$ for $t \in [0,1]$. Given $t_0 \in [0,1]$, use the first claims in **(B1)–(B2)** similarly, to define the curve $s \in [0,1] \longmapsto x_s \in \overline{X}$ for which

$$D_y b(x_s, y_{t_0}) = (1-s) D_y b(x, y_{t_0}) + s D_y b(x_1, y_{t_0}), \tag{6.1}$$

and set $g(s) = \frac{\partial^2 f}{\partial t^2}(x_s, t_0)$. The convexity of (4.6) will be verified by checking $g(1) \geqslant 0$. Let us start by observing $s \in [0,1] \longmapsto g(s)$ is a convex function, as a consequence of property **(B3)** and (6.1). We next claim $g(s)$ is minimized at $s = 0$, since

$$g'(s) = \frac{\partial^2}{\partial t^2}\bigg|_{t=t_0} \big\langle D_x b\big(x_s, y_b\big(x_0, (1-t)q_0 + t q_1\big)\big), \dot{x}_s \big\rangle$$

vanishes at $s = 0$, since $x_0 = x$ in the definition (4.3) of y_b. Thus $g(1) \geqslant g(0) = 0$, establishing the convexity of (4.6). If b is positively cross-curved, then $g''(s) > 0$ and the desired strong convexity follows from $g(1) > g(0) = 0$ since $x \neq x_1$ implies the curve x_s does not degenerate to a single point.

Conversely, if the convexity of (4.6) fails we can find $x_1 \in \overline{X}$ and $s_0, t_0 \in [0,1]$ for which the construction above yields $g''(s_0) < 0$. In view of Lemma 4.5 of [16], this provides a contradiction to (2.1). □

We shall also need to recall two basic facts about b-convex functions from e.g. [14]: any supremum of b-convex functions is again b-convex, unless it is identically infinite; and for each $y \in \overline{Y}$ and $\lambda \in \mathbf{R}$, the function

$$x \in \overline{X} \longmapsto b(x,y) - \lambda \tag{6.2}$$

is b-convex. Functions of the form either $y \in \overline{Y} \longmapsto b(x,y) - \lambda$ or (6.2) are sometimes called *mountains* below.

Proof of Proposition 4.4. The b^*-convexity of the manufacturing cost $c = (c^b)^{b^*}$ asserts

$$c(y) = \sup_{x \in \overline{X}} b(x,y) - c^b(x)$$

is a supremum of mountains, whence

$$a(q) := c\big(y_b(x,q)\big) - b\big(x, y_b(x,q)\big) = \sup_{x_0 \in \overline{X}} b\big(x_0, y_b(x,q)\big) - b\big(x, y_b(x,q)\big) - c^b(x_0)$$

for all $x \in \overline{X}$ and $q \in \overline{Y}_x$. According to Lemma 4.3, we have just expressed $a(q)$ as a supremum of convex functions, thus establishing convexity of $a(q)$. The remainder of the proof will be devoted to deducing strict convexity of $a(q)$ under the additional hypotheses (i) or (ii).

In case (ii), $\mathbf{(B3)_s}$ implies all but one of the functions of $q \in Y_x$ under the supremum above are strictly convex, the exception being the constant function $-c^b(x)$ corresponding to $x_0 = x$. Thus $a(q)$ is strictly convex, except possibly on the set $\{q \in \overline{Y}_x \mid a(q) = -c^b(x)\}$ where its lower bound is attained. However, if q_0 belongs to this set, differentiating the function under the supremum with respect to x_0 yields $D_x b(x, y_b(x, q_0)) \in \partial c^b(x)$. Since (ii) assumes differentiability of c^b at x, **(B1)** then implies the minimum of $a(q)$ is attained uniquely at $q_0 = Dc^b(x)$, to establish strict convexity of $a(q)$.

The remainder of the proof will be devoted to case (i): deducing strict convexity of $a(q)$ from strict b^*-convexity of $c(y)$ assuming only **(B3)**. Recall that strict b^*-convexity was defined by continuity of the agents' responses $y_{b,c} : \overline{X} \longmapsto \overline{Y}$ to the principal's manufacturing costs (as opposed to the prices the principal would prefer to select). Fix $x \in \overline{X}$ and use the C^3 change of variables $q \in \overline{Y}_x \longmapsto y_b(x,q) \in \overline{Y}$ to define $\tilde{b}(\cdot, q) := b(\cdot, y_b(x,q)) - b(x, y_b(x,q))$ and $\tilde{c}(q) := c(y_b(x,q)) - b(x, y_b(x,q)) = a(q)$. As in [13], it is easy to deduce that $\tilde{b}$ satisfies the same hypotheses **(B0)–(B3)** on $\overline{X} \times \overline{Y}_x$ as the original valuation function — except for the fact that $\tilde{b} \in C^3$ whereas $b \in C^4$. For the reasons explained in [13] this discrepancy shall not trouble us here: we still have continuous fourth derivatives of $\tilde{b}$ as long as at least one of the four derivatives is with respect to a variable in $\overline{X}$, and at most three derivatives are with respect to variables in $\overline{Y}_x$. Note also that $\tilde{c}^{\tilde{b}} = c^b$ and the continuity of the agents' responses $y_{\tilde{b},\tilde{c}}$ in the new variables follows from their presumed continuity in the original variables, since $y_{\tilde{b},\tilde{c}}(\cdot) = D_x c(x, y_{b,c}(\cdot))$.

The advantage of the new variables is that for each $x_0 \in \overline{X}$, the mountain $q \in \overline{Y}_x \longmapsto \tilde{b}(x_0, q)$ is a convex function, according to Lemma 4.3; (alternately, Theorem 4.3 of [13]). To produce a contradiction, assume convexity of $\tilde{c}(q)$ fails to be strict, so there is a segment $t \in [0,1] \longmapsto q_t \in \overline{Y}_x$ given by $q_t = (1-t)q_0 + tq_1$ along which $\tilde{c}$ is affine with the same slope $p \in \partial\tilde{c}(q_t)$ for each $t \in [0,1]$. In fact, the compact convex set $\partial\tilde{c}(q_t)$ is independent of $t \in]0,1[$, so taking p to be an extreme point of $\partial\tilde{c}(q_t)$ allows us to find a sequence $q_{t,k} \in Y_x \cap \operatorname{Dom} D\tilde{c}$ converging to q_t such that $p = \lim_{k\to\infty} D\tilde{c}(q_{t,k})$, by Theorem 25.6 of Rockafellar [33]. On the other hand, b^*-convexity implies $\tilde{c}(q)$ is a supremum of mountains: thus to each $t \in [0,1]$ and integer k corresponds some $x_{t,k} \in \overline{X}$ such that $(x_{t,k}, q_{t,k}) \in \partial^{\tilde{b}^*}\tilde{c}$, meaning

$$\tilde{c}(q) \geqslant \tilde{b}(x_{t,k}, q) - \tilde{b}(x_{t,k}, q_{t,k}) + \tilde{c}(q_{t,k}) \tag{6.3}$$

for all $q \in \overline{Y}_x$. Since $q_{t,k} \in \operatorname{Dom} D\tilde{c}$, saturation of this bound at $q_{t,k}$ implies $D\tilde{c}(q_{t,k}) = D_q\tilde{b}(x_{t,k}, q_{t,k})$. Compactness of $\overline{X}$ allows us to extract a subsequential limit $(x_{t,k}, q_{t,k}) \to (x_t, q_t) \in \partial^{\tilde{b}^*}\tilde{c}$ satisfying $p = D_q\tilde{b}(x_t, q_t)$. This first order condition shows the curve $t \in [0,1] \longmapsto x_t \in \overline{X}$ to be differentiable, with derivative

$$\dot{x}_t = -D^2_{qx}\tilde{b}(x_t, q_t)^{-1} D^2_{qq}\tilde{b}(x_t, q_t)\dot{q}_t, \tag{6.4}$$

by the implicit function theorem and **(B1)**. On the other hand, both $\tilde{c}(\cdot)$ and $\tilde{b}(x_t, \cdot)$ are convex functions of $q \in \overline{Y}_x$ in (6.3), so both must be affine along the segment q_t. This implies $\dot{q}_t = q_1 - q_0$ is a zero eigenvector of $D^2_{qq}\tilde{b}(x_t, q_t)$, which in turn implies $x_t = \text{const}$ from (6.4). On the other hand, the efficient response $q_t = y_{\tilde{b},\tilde{c}}(x_t)$ of agent x_t to price menu $\tilde{c}$ is not constant, since the endpoints $q_0 \neq q_1$ of the segment are distinct. This produces the desired contradiction and establishes strict convexity of $\tilde{c}$. □

Combining Proposition 4.4 with the following standard lemma will allow us to establish our necessary and sufficient criteria for convexity of the feasible set $\mathcal{U}_0$.

Lemma 6.1 *(Identification of supporting mountains). Let u be a b-convex function on $\overline{X}$. Assume u is differentiable at $x_0 \in X$ and $D_x u(x_0) = D_x b(x_0, y)$ for some $y \in \overline{Y}$. Then, $u(x) \geqslant m(x)$ for all $x \in X$, where $m(\cdot) = b(\cdot, y) - b(x_0, y) + u(x_0)$.*

Proof. By b-convexity of u, there exists $y_0 \in \overline{Y}$ such that $u(x_0) = b(x_0, y_0) - u^{b^*}(y_0)$ and also $u(x) \geqslant b(x, y_0) - u^{b^*}(y_0)$ for all $x \in X$. Since u is differentiable at x_0, this implies $D_x u(x_0) = D_x b(x_0, y_0)$. By the assumption **(B1)**, we conclude $y = y_0$. This completes the proof since $m(\cdot) = b(\cdot, y_0) - u^{b^*}(y_0)$. □

Proof of Theorem 3.2. Let us first show the sufficiency. It is enough to show that for any two b-convex functions u_0 and u_1, the linear combination $u_t := (1-t)u_0 + tu_1$ is again b-convex, for each $0 \leqslant t \leqslant 1$. Fix $x_0 \in \overline{X}$. Since b-convex functions are defined as suprema of mountains, there exist $y_0, y_1 \in \overline{Y}$ such that

$$m_i^{x_0}(\cdot) := b(\cdot, y_i) - b(x_0, y_i), \quad i = 0, 1,$$

satisfy $u_i(x) \geqslant m_i^{x_0}(x) + u_i(x_0)$ for all $x \in \overline{X}$. Clearly equality holds when $x = x_0$. Let us consider the function

$$m_t^{x_0}(\cdot) = b(\cdot, y_t) - b(x_0, y_t),$$

where y_t defines a line segment

$$t \in [0,1] \longmapsto D_x b(x_0, y_t) = (1-t) D_x b(x_0, y_0) + t D_x b(x_0, y_1) \in \mathbf{R}^n.$$

Note that (i) $m_t^{x_0}(x_0) = 0$. We claim that (ii) $u_t(\cdot) \geqslant m_t^{x_0}(\cdot) + u_t(x_0)$. Notice that

$$u_t(\cdot) \geqslant (1-t) m_0^{x_0}(\cdot) + t m_1^{x_0}(\cdot) + u_t(x_0).$$

Thus the claim follows from the inequality $(1-t)m_0^{x_0} + t m_1^{x_0} \geqslant m_t^{x_0}$, which is implied by **(B3)** according to Lemma 4.3. The last two properties (i) and (ii) enable one to express u_t as a supremum of mountains

$$u_t(\cdot) = \sup_{x_0 \in X} m_t^{x_0}(\cdot) + u_t(x_0),$$

hence u_t is b-convex by the remark immediately preceding (6.2).

Conversely, let us show the necessity of **(B3)** for convexity of $\mathcal{V}_{\overline{Y}}^b$. Using the same notation as above, recall that each mountain $m_i^{x_0}$, $i = 0, 1$ is b-convex. Assume the linear combination $h_t := (1-t)m_0^{x_0} + t m_1^{x_0}$ is b-convex. Since $D_x h_t(x_0) = (1-t) D_x b(x_0, y_0) + t D_x b(x_0, y_1)) = D_x m_t(x_0)$, Lemma 6.1 requires that $m_t^{x_0} \leqslant h_t$ for every $0 \leqslant t \leqslant 1$. This last condition is equivalent the property characterizing non-negative cross-curvature in Lemma 4.3. This completes the proof of necessity and the proof of the theorem. □

Let us turn now to the convexity of the principal's problem.

Proof of Corollary 4.5. Corollary 4.5 follows by combining the convexity of the set $\mathcal{U}_0$ of feasible strategies proved in Theorem 3.2 with the convexity of $a(q)$ from Proposition 4.4. If μ fails to vanish on each Lipschitz hypersurface, a little care is needed to deduce convexity of the principal's objective $L(u)$ from that of $a(q)$, by invoking the conventions adopted in Remark 4.2 as follows. Let $t \in [0,1] \longmapsto u_t = (1-t)u_0 + tu_1$ denote a line segment in the convex set $\mathcal{U}_0$. If $q \in \partial u_t(x)$ for some $x \in X$, then $y_b(x,q) \in \partial^b u_t(x)$ by Theorem 3.1 of Loeper [19]; (a direct proof along the lines of Lemma 4.3 may be found in [16]). So $y_b(x,q)$ is among the best responses of x to price menu $v_t = u_t^{b^*}$. For each $t \in [0,1]$ select $Du_t(x) \in \partial u_t(x)$ measurably to ensure $\min\{c(y_b(x,q)) - b(x, y_b(x,q)) \mid q \in \partial u_t(x)\}$ is achieved at $q = Du_t(x)$. Then $a(Du_t(x)) \leqslant a((1-t)Du_0(x) + tDu_1(x))$ since $(1-t)Du_0(x) + tDu_1(x) \in \partial u_t(x)$. The desired convexity of $L(u)$ follows. □

Next we establish uniqueness of the principal's strategy.

Proof of Theorem 4.6. Suppose both u_0 and u_1 minimize the principal's net losses $L(u)$ on the convex set $\mathcal{U}_0$. Define the line segment $u_t = (1-t)u_0 + tu_1$ and — in case μ fails to vanish on each Lipschitz hypersurface — the measurable selection $Du_t(x) \in \partial u_t(x)$ as in the proof of Corollary 4.5. The strict convexity of $a(q)$ asserted by Proposition 4.4 combines with the tie-breaking rule to remove all freedom from this selection. Under the hypotheses of Theorem 4.6, the same strict convexity implies the contradiction $L(u_{1/2}) < \frac{1}{2}L(u_0) + \frac{1}{2}L(u_1) = L(u_1)$ unless $Du_0 = Du_1$ holds μ-a.e. This establishes the uniqueness μ-a.e. of the agents' equilibrium strategies $y_{b,v}(x) := y_b(x, Du_1(x))$, and of the principal's optimal measure $\nu := (y_{b,v})_{\#}\mu$ in (3.5).

Let $\operatorname{spt}\mu$ denote the smallest closed subset of $\overline{X}$ containing the full mass of μ. To identify $u_0 = u_1$ on $\operatorname{spt}\mu$ and establish the remaining assertions is more technical. First observe that the participation constraint $u_{1/2}(x) \geqslant b(x, y_\emptyset) - c(y_\emptyset) =: u_\emptyset(x)$ on the continuous function $u_{1/2} \in \mathcal{U}_0$ must bind for some agent type $x_0 \in \operatorname{spt}\mu$; otherwise, for $\epsilon > 0$ sufficiently small, $u_{1/2} - \epsilon$ would belong to $\mathcal{U}_0$ and reduce the principal's losses by ϵ, contradicting the asserted optimality of $u_{1/2}$. Since $u_{1/2}$ is a convex combination of two other functions obeying the same constraint, we conclude $u_0(x_0) = u_1(x_0)$ coincides with the reservation utility $u_\emptyset(x_0)$ for type x_0. Now use the map $y_{b,v} := y_b \circ Du_1$ from the first paragraph of the proof to define a joint measure $\gamma := (id \times y_{b,v})_{\#}\mu$ given by $\gamma[U \times V] = \mu[U \times y_{b,v}^{-1}(V)]$ for Borel $U \times V \subset \overline{X} \times \overline{Y}$, and denote by $\operatorname{spt}\gamma$ the smallest closed subset $S \subset \overline{X} \times \overline{Y}$ carrying the full mass of γ. Notice $\operatorname{spt}\gamma$ does not depend on $t \in [0,1]$, nor in fact on u_0 or u_1; any other optimal strategy for the principal would lead to the same γ.

Since the graph of $y_{b,v}$ lies in the closed set $\partial^b u_1 \subset \overline{X} \times \overline{Y}$, the same is true of $S := \{(x_0, y_\emptyset)\} \cup \operatorname{spt}\gamma$. Thus S is b-cyclically monotone (A.1) by the result of Rochet [30] discussed immediately before Lemma A.1. Lemma A.1 then yields a minimal b-convex function u_- satisfying $u_-(x_0) = b(x_0, y_\emptyset) - c(y_\emptyset)$ for which $S \subset \partial^b u_-$. The fact that $(x_0, y_\emptyset) \in S$ implies some mountain $b(\cdot, y_\emptyset) + \lambda$ bounds $u_-(\cdot)$ from below with contact at x_0. Clearly $\lambda = -c(y_\emptyset)$ whence $u_- \in \mathcal{U}_0$.

Now we have $u_i \geqslant u_-$ for $i = 0, 1$ with equality at x_0. Also, $y_{b,v}(x) \in \partial^b u_-(x)$ for μ almost all x, whence u_- must be an optimal strategy: it is smaller in value than u_i and produces at least as favorable a response as u_i from almost all agents. Finally since

$$L(u_i) - L(u_-) \geqslant \int_X \big(u_i(x) - u_-(x)\big)\, d\mu(x) \geqslant 0,$$

the fact that u_i minimizes the losses of the principal implies the continuous integrand vanishes μ-almost everywhere. Thus $u_i \geqslant u_-$ on $\overline{X}$, with equality holding throughout $\operatorname{spt}\mu$ as desired.

Since u_0 was arbitrary, we have now proved that all optimal $u \in \mathcal{U}_0$ coincide with u_1 on $\operatorname{spt}\mu$. Optimality of u also implies $\operatorname{spt}\gamma \subset \partial^b u$; if in addition the participation constraint $u(x) \geqslant b(x, y_\emptyset) - c(y_\emptyset)$ binds at x_0, then $u \geqslant u_-$ on $\overline{X}$. Although u_- appears to depend on our choice of $x_0 \in \operatorname{spt}\mu$ in the construction above this is not actually the case: $u(x_0) = u_1(x_0)$ shows the participation constraint binds at x_0 for every optimal strategy and u_- is therefore uniquely determined by its minimality among optimal strategies $u \in \mathcal{U}_0$.

Now, since any supremum of b-convex functions (not identically infinite) is again b-convex, define $u_+ \in \mathcal{U}_0$ as the pointwise supremum among all of the principal's equilibrium strategies $u \in \mathcal{U}_0$. The foregoing shows $u_+ = u_-$ on $\operatorname{spt}\mu$, while $(x, y) \in \operatorname{spt}\gamma \subset \partial^b u$ implies

$$\begin{aligned} u_+(\cdot) &\geqslant u(\cdot) \geqslant u(x) + b(\cdot, y) - b(x, y) \\ &= u_+(x) + b(\cdot, y) - b(x, y) \end{aligned}$$

on $\overline{X}$, whence $\operatorname{spt}\gamma \subset \partial^b u_+$. From here we deduce $L(u_+) \leqslant L(u)$, hence u_+ is itself an optimal strategy for the principal.

Finally, $v : \overline{Y} \longmapsto \mathbf{R} \cup \{+\infty\}$ is an equilibrium price menu in Carlier's reformulation [8] if and only if $u := v^b$ minimizes $L(u)$ on $\mathcal{U}_0$, in which case $u_- \leqslant u \leqslant u_+$ throughout $\overline{X}$ implies $u_+^{b^*} \leqslant (v^b)^{b^*} \leqslant u_-^{b^*}$ throughout $\overline{Y}$. Moreover, $u_- = u_+$ on $\operatorname{spt}\mu$ implies $u_+^{b^*} = u_-^{b^*}$ on $\operatorname{spt}\nu$, since $y_{b,v}(x) \in \partial^b u_\pm(x)$ for μ-a.e. x implies $u_\pm^{b^*}(y_{b,v}(x)) = b(x, y_{b,v}(x)) - u_\pm(x)$. We therefore conclude that if v is an equilibrium price menu, then $v \geqslant (v^b)^{b^*} \geqslant u_+^{b^*}$ on $\overline{Y}$, with both equalities holding ν-a.e. Conversely, if $v : \overline{Y} \longmapsto \mathbf{R} \cup \{+\infty\}$ satisfies $v \geqslant u_+^{b^*}$ with equality ν-a.e., we deduce the same must be true for its b-convex hull $(v^b)^{b^*}$, the latter being the largest b-convex function dominated by v. Thus $(v^b)^{b^*}(y_\emptyset) = c(y_\emptyset)$ and $v^b \in \mathcal{U}_0$ and $v^b \leqslant u^+$ throughout $\overline{X}$ with equality holding μ-a.e. If μ vanishes on Lipschitz hypersurfaces, then $Dv^b = Du^+$ agree μ-a.e., so $L(v^b) = L(u^+)$ and v^b is a optimal strategy for the principal as desired. If, on the other hand, μ does not vanish on all Lipschitz hypersurfaces, then we may assume v is its own b^*-convex hull by Remark 4.2. Any mountain which touches $u_+^{b^*}$ from below on $\operatorname{spt}\nu$ also touches $v \geqslant u_+^{b^*}$ from below at the same point, thus $\partial^{b^*} u_+^{b^*} \subset \partial^{b^*} v$; since v is b-convex this is equivalent to $\partial^b u_+ \subset \partial^b v^b$. This shows the best response of x facing price menu $u_+^{b^*}$ is also one of his best responses facing price menu v: he cannot have a better response since his indirect utility $v^b \leqslant u^+$. The constraint on the agent's behaviour imposed by Remark 4.2 now implies $L(v^b) \leqslant L(u^+)$; equality must hold since u^+ is one of the principal's optimal strategies. This confirms optimality of v^b and concludes the proof of the theorem. □

To show stability of the equilibrium requires the following convergence result concerning Borel probability measures $\mathcal{P}(\overline{X} \times \overline{Y})$ on the product space.

Proposition 6.2 *(Convergence of losses and mixed strategies). Suppose a sequence of triples* $(b_\infty, c_\infty, \mu_\infty) = \lim_{i\to\infty}(b_i, c_i, \mu_i)$ *satisfy the hypotheses of Corollary* 4.7. *Let* $L_i(u)$ *denote the net losses* (4.5) *by a principal who adopts strategy u facing data* (b_i, c_i, μ_i). *If any sequence* u_i *of* b_i*-convex functions converge uniformly on X, then their limit* u_∞ *is* b_∞*-convex and* $L_\infty(u_\infty) = \lim_{i\to\infty} L_i(u_i)$. *Furthermore, there is a unique joint measure* $\gamma_\infty \in \mathcal{P}(\overline{X} \times \overline{Y})$ *supported in* $\partial^{b_\infty} u_\infty$ *with left marginal* μ_∞, *and any sequence of joint measures* $\gamma_i \in \mathcal{P}(\overline{X} \times \overline{Y})$ *vanishing outside* $\partial^{b_i} u_i$ *and with left marginal* μ_i, *must converge weakly-$*$ to* γ_∞.

Proof. Assume a sequence $u_i \to u_\infty$ of b_i-convex functions converges uniformly on $\overline{X}$. Topologizing the continuous functions $C(\overline{Z})$ by uniform convergence, where $Z = X, Y$ or $X \times Y$, makes the transformation $(b, u) \longmapsto u^{b^*}$ given by (3.1) continuous on $C(\overline{X} \times \overline{Y}) \times C(\overline{X})$. This fact allows us to take $i \to \infty$ in the relation $u_i^{b_i^* b_i} = u_i$ to conclude b_∞-convexity of u_∞. From the semiconvexity (4.4) of u_∞ we infer its domain of differentiability $\operatorname{Dom} Du_\infty$ exhausts X apart from a countable collection of Lipschitz hypersurfaces, which are μ_∞-negligible by hypothesis. Define the map $G_\infty(x) = y_{b_\infty}(x, Du_\infty(x))$ on $\operatorname{Dom} Du_\infty$. Since $\partial^{b_\infty} u_\infty \cap (\operatorname{Dom} Du_\infty \times \overline{Y})$ coincides with the graph of G_∞, any measure γ_∞ supported in $\partial^{b_\infty} u_\infty$ with left marginal μ_∞ is given (6.5) by $\gamma_\infty := (id \times G_\infty)_\# \mu_\infty$ as in, e.g., Lemma 2.1 of Ahmad et al. [1]. This specifies γ_∞ uniquely.

Now suppose $\gamma_i \geqslant 0$ is a sequence of measures supported in $\partial^{b_i} u_i$ having left marginal μ_i. Compactness allows us to extract from any subsequence of γ_i a further subsequence which converges weakly-$*$ to some limit $\bar\gamma \in \mathcal{P}(\overline{X} \times \overline{Y})$. Since $\mu_i \rightharpoonup \mu_\infty$ the left marginal of $\bar\gamma$ is given by μ_∞. Moreover, since $u_i(x) + u_i^{b_i^*}(y) \geqslant b_i(x, y)$ throughout $\overline{X} \times \overline{Y}$ with equality on $\operatorname{spt}\gamma_i$, uniform convergence of this expression yields $\operatorname{spt}\bar{\gamma} \subset \partial^{b_\infty} u_\infty$. The uniqueness result of the preceding paragraph then asserts $\bar{\gamma} = \gamma_\infty$ independently of the choice of subsequence, so the full sequence $\gamma_i \rightharpoonup \gamma_\infty$ converges weakly-$*$.

Finally, use the measurable selection $Du_i(x) \in \partial u_i(x)$ of Remark 4.2 to extend $Du_i(x)$ from $\operatorname{Dom} Du_i$ to X so as to guarantee that $G_i(x) := y_{b_i}(x, Du_i(x))) \in \partial^{b_i} u_i(x)$. Use the Borel map $G_i : X \longmapsto \overline{Y}$ to push μ_i forward to the joint probability measure $\gamma_i := (id \times G_i)_\# \mu_i$ on $X \times \overline{Y}$ defined by

$$\gamma_i[U \times V] := \mu_i\big[U \cap G_i^{-1}(V)\big] \tag{6.5}$$

for each Borel $U \times V \subset X \times \overline{Y}$. Notice γ_i is supported in $\partial^{b_i} u_i$ and has μ_i for its left marginal, hence converges weakly-$*$ to γ_∞. Moreover, our choice of measurable selection guarantees that the net losses (4.5) of the principal choosing strategy u_i coincide with

$$L_i(u_i) = \int\limits_{X \times \overline{Y}} \big(c_i(y) - u_i^{b_i^*}(y)\big)\, d\gamma_i(x, y). \tag{6.6}$$

Weak-$*$ convergence of the measures $\gamma_i \rightharpoonup \gamma_\infty$ couples with uniform convergence of the integrands to yield the desired limit

$$\lim_{i \to \infty} L_i(u_i) = \int\limits_{X \times \overline{Y}} \big(c_\infty(y) - u_\infty^{b_\infty^*}(y)\big)\, d\gamma_\infty(x, y) = L_\infty(u_\infty)$$

and establish the proposition. □

Proof of Corollary 4.7. Let $\mathcal{U}_0^i$ denote the space of b_i-convex functions $u(\cdot) \geqslant b_i(\cdot, y_\emptyset) - c_i(y_\emptyset)$, and $L_i(u)$ denote the net loss of the principal who chooses strategy u facing the triple (b_i, c_i, μ_i). The L_i-minimizing strategies $u_i \in \mathcal{U}_0^i$ are Lipschitz and semiconvex, with upper bounds (4.4) on $|Du_i|$ and $-D^2 u_i$ which are independent of i since $\|b_i - b_\infty\|_{C^2} \to 0$. The Ascoli–Arzelà theorem therefore yields a subsequence $u_{i(j)}$ which converges uniformly to a limit $\bar{u}$ on the compact set $\overline{X}$. Since the functions u_i have a semiconvexity constant independent of i, it is a well-known corollary that their gradients also converge $Du_{i(j)}(x) \to D\bar{u}(x)$ pointwise on the set of common differentiability $(\operatorname{Dom} D\bar{u}) \cap (\bigcap_{i=1}^\infty \operatorname{Dom} Du_i)$. This set exhausts $\overline{X}$ up to a countable union of Lipschitz hypersurfaces — which is μ_∞-negligible by hypothesis. Setting $G_i(x) = y_{b_i}(x, Du_i(x))$, it is not hard to deduce $y_{b_\infty}(x, D\bar{u}(x)) = \lim_{j \to \infty} G_{i(j)}(x)$ on this set from Definition 4.1. If we can now prove $\bar{u}$ minimizes $L_\infty(u)$ on $\mathcal{U}_0^\infty$, the uniqueness of equilibrium product selected by μ_∞-a.e. agent $x \in X$ in Theorem 4.6 will then imply that $\lim_{j \to \infty} G_{i(j)}(x) = G_\infty(x)$ converges to a limit independent of the subsequence chosen, hence the full sequence $G_i(x)$ converges μ_∞-a.e.

To see that $\bar{u}$ minimizes $L_\infty(u)$ on $\mathcal{U}_0^\infty$, observe $u \in \mathcal{U}_0^\infty$ implies $u^{b_\infty^* b_i} \in \mathcal{U}_0^i$ is L_i-feasible, being the b_i-transform of a price menu $u^{b_\infty^*}(\cdot)$ agreeing with $c_\infty(\cdot)$ at $y_\emptyset$. Moreover, $u^{b_\infty^* b_i} \to u^{b_\infty^* b_\infty}$ uniformly as $i \to \infty$ (by continuity of the b-transform asserted in the first paragraph of the preceding proof). The optimality of u_i therefore yields $L_i(u_i) \leqslant L_i(u^{b_\infty^* b_i})$. Proposition 6.2 allows us to deduce $L_\infty(\bar{u}) \leqslant L_\infty(u)$ by taking the subsequential limit $j \to \infty$. Since the same proposition asserts b_∞-convexity of $\bar{u}$, we find $\bar{u} \in \mathcal{U}_0^\infty$ is the desired minimizer after taking the limit $j \to \infty$ in $u_{i(j)}(\cdot) \geqslant b_{i(j)}(\cdot, y_\emptyset) - c_{i(j)}(y_\emptyset)$. This concludes the proof of μ_∞-a.e. convergence of the maps $G_\infty(x) = \lim_{i \to \infty} G_i(x)$.

Turning to the optimal measures: as in the preceding proof, a measurable selection $Du_i(x) \in \partial u_i(x)$ consistent with the tie-breaking hypotheses of Remark 4.2 may be used to extend the Borel map $G_i(x) = y_b(x, Du_i(x))$ from $\operatorname{Dom} Du_i$ to X and define a joint measure $\gamma_i := (id \times G_i)_{\#}\mu_i$ supported on $\partial^{b_i} u_i$ as in (6.5). The left marginal of γ_i is obviously given by μ_i, and its right marginal coincides with the unique optimal measure ν_i given by Theorem 4.6. Proposition 6.2 then yields weak-$*$ convergence of $\gamma_i \rightharpoonup \gamma_\infty$ and hence of $\nu_i \rightharpoonup \nu_\infty$. Theorem 4.6 also asserts the two minimizers $u_\infty = \bar{u}$ agree μ_∞-a.e. In this case the uniform limit $\bar{u}$ is independent of the Ascoli–Arzelà subsequence, hence we recover convergence of the full sequence $u_i \to u_\infty$ in $L^\infty(X, d\mu_\infty)$. $\square$

Finally, let us extend Armstrong's desirability of exclusion to our model. Our proof is inspired by Armstrong's [2], but differs from his in a number of ways.

Proof of Theorem 4.8. Use the C^3-smooth diffeomorphism $x \in \overline{X} \longmapsto p = D_y b(x, y_\emptyset) \in \overline{X}_{y_\emptyset}$ provided by **(B0)–(B2)** and its inverse $p \in \overline{X}_{y_\emptyset} \longmapsto x = x_b(y_\emptyset, p) \in \overline{X}$ to reparameterize the space of agents over the strictly convex set $\overline{X}_{y_\emptyset}$. Then $\tilde{u}(p) := u(x_b(y_\emptyset, p)) - b(x_b(y_\emptyset, p), y_\emptyset) + c(y_\emptyset)$ defines a non-negative $\tilde{b}$-convex function, where $\tilde{b}(p, y) := b(x_b(y_\emptyset, p), y) - b(x_b(y_\emptyset, p), y_\emptyset) + c(y_\emptyset)$. In other words, the space $\mathcal{U}_0$ corresponds to the space $\tilde{\mathcal{U}}_0$ of non-negative $\tilde{b}$-convex functions on $\overline{X}_{y_\emptyset}$ in the new parameterization. This subtraction of the reservation utility from the valuation function does not change any agent's response to a price menu v offered by the principal, since valuations by different agent types are never compared. However, it does make the valuation function $\tilde{b}(p, y)$ a convex function of $p \in \overline{X}_{y_\emptyset}$, as is easily seen by interchanging the roles of x and y in Lemma 4.3. The indirect utility $\tilde{u}(p) = v^{\tilde{b}}(p)$ is then also convex, being a supremum (3.1) of such valuation functions.

In the new variables, the distribution of agents $\tilde{f}(p)dp = f(x)dx$ is given by $\tilde{f}(p) = f(x_b(y_\emptyset, p)) \det[\partial x_b^i(y_\emptyset, p)/\partial p_j]$. The principal's net losses $\tilde{L}(\tilde{u}) = L(u)$ are given as in (4.5) by

$$\tilde{L}(\tilde{u}) = \int\limits_{X_{y_\emptyset}} \tilde{a}\big(D\tilde{u}(p), \tilde{u}(p), p\big) \tilde{f}(p)\, dp,$$

where $\tilde{a}(q, s, p) = c(y_{\tilde{b}}(p, q)) - \tilde{b}(p, y_{\tilde{b}}(p, q)) + s$ is a convex function of q on $\tilde{Y}_p := D_p\tilde{b}(p, Y)$ for each fixed p and s, according to Proposition 4.4; (recall that $\tilde{b} \in C^3(\overline{X}_{y_\emptyset} \times \overline{Y})$ satisfies the same hypotheses **(B0)–(B3)** as $b \in C^4(\overline{X} \times \overline{Y})$, except for the possibility that four continuous derivatives with respect to variables in $X_{y_\emptyset}$ fail to exist, which is irrelevant as already discussed). This convexity implies

$$\tilde{a}(q, s, p) \geqslant \tilde{a}(q_0, s, p) + \big\langle D_q\tilde{a}(q_0, s, p), q - q_0\big\rangle$$

for all $q, q_0 \in \overline{\tilde{Y}}_p$. With p still fixed, the choice $q_0 = D_p\tilde{b}(p, y_\emptyset) = \mathbf{0}$ shows $\tilde{a}(\mathbf{0}, s, p) = s$ whence $\tilde{a}(q, s, p) \geqslant \langle D_q\tilde{a}(\mathbf{0}, s, p), q\rangle$ for $s = \tilde{u}(x) \geqslant 0$.

Now suppose $\tilde{u} \in \tilde{\mathcal{U}}_0$ minimizes $\tilde{L}(\tilde{u})$. For $\epsilon \geqslant 0$, define the continuously increasing family of compact convex sets $\tilde{U}_\epsilon := \{p \in \overline{X}_{y_\emptyset} \mid \tilde{u}(p) \leqslant \epsilon\}$. Observe that $\tilde{U}_0$ must be non-empty, since otherwise for $\epsilon > 0$ small enough $\tilde{U}_\epsilon$ would be empty, and then $\tilde{u} - \epsilon \in \tilde{\mathcal{U}}_0$ is a better strategy, reducing the principal's losses by ϵ. We now claim the interior of the set $\tilde{U}_0$ — which corresponds to agents who decline to participate — contains a non-zero fraction of the total population of agents. Our argument is inspired by the strategy Armstrong worked out in a special case [2], which was to show that unless this conclusion is true, the profit the principal extracts from agents in $\tilde{U}_\epsilon$ would vanish at a higher order than $\epsilon > 0$, making $\tilde{u}_\epsilon := \max\{\tilde{u} - \epsilon, 0\} \in \tilde{\mathcal{U}}_0$ a better strategy than $\tilde{u}$ for the principal when ϵ is sufficiently small.

For $\epsilon > 0$, the contribution of $\tilde{U}_\epsilon$ to the principal's profit is given by

$$\begin{aligned} -\tilde{L}_\epsilon(\tilde{u}) &:= -\int_{\tilde{U}_\epsilon} \tilde{a}\big(D\tilde{u}(p), \tilde{u}(p), p\big)\tilde{f}(p)\,dp \\ &\leqslant -\int_{\tilde{U}_\epsilon} \big\langle D_q\tilde{a}\big(\mathbf{0}, \tilde{u}(p), p\big), D\tilde{u}(p)\big\rangle \tilde{f}(p)\,dp \\ &= \int_{\tilde{U}_\epsilon} \tilde{u}(p)\nabla_p \cdot \big(\tilde{f}(p)D_q\tilde{a}\big(\mathbf{0}, \tilde{u}(p), p\big)\big)\,dp - \int_{\partial\tilde{U}_\epsilon} \tilde{u}(p)\langle D_q\tilde{a}, \hat{n}\rangle \tilde{f}(p)\,dS(p) \end{aligned} \tag{6.7}$$

where $\hat{n} = \hat{n}_{\tilde{U}_\epsilon}(p)$ denotes the outer until normal to $\tilde{U}_\epsilon$ at p, and the divergence theorem has been used. Here $\partial\tilde{U}_\epsilon$ denotes the boundary of the convex set $\tilde{U}_\epsilon$, and $dS(p)$ denotes the $n-1$ dimensional surface (i.e. Hausdorff) measure on this boundary. (For Sobolev functions, the integration by parts formula that we need is contained in §4.3 of [12] under the additional restriction that the vector field $\tilde{u}(\cdot)D_q a(\mathbf{0}, \tilde{u}(\cdot), \cdot)$ be C^1 smooth, but extends immediately to Lipschitz vectors fields by approximation; the operation of restricting $\tilde{f}$ to the boundary of $\tilde{U}_\epsilon$ is there shown to give a bounded linear map from $W^{1,1}(U_\epsilon, dp)$ to $L^1(\partial U_\epsilon, dS)$ called the boundary *trace*.) As $\epsilon \to 0$, we claim both integrals in (6.7) vanish at rate $o(\epsilon)$ if the interior of $\tilde{U}_0$ is empty. To see this, note $\tilde{u} = \epsilon$ on $\partial\tilde{U}_\epsilon \cap \operatorname{int} X_{y_\emptyset}$, so

$$\begin{aligned} &\int_{\partial\tilde{U}_\epsilon} \tilde{u}(p)\langle D_q\tilde{a}, \hat{n}\rangle \tilde{f}(p)\,dS(p) \\ &\quad= \epsilon \int_{\partial\tilde{U}_\epsilon} \langle D_q\tilde{a}, \hat{n}\rangle \tilde{f}(p)\,dS(p) + \int_{\partial\tilde{U}_\epsilon \cap \partial X_{y_\emptyset}} \big[\tilde{u}(p) - \epsilon\big]\langle D_q\tilde{a}, \hat{n}\rangle \tilde{f}(p)\,dS(p) \\ &\quad= \epsilon \int_{\tilde{U}_\epsilon} \nabla_p \cdot \big(\tilde{f}(p)D_q\tilde{a}\big(\mathbf{0}, \tilde{u}(p), p\big)\big)\,dp + \int_{\tilde{U}_\epsilon \cap \partial X_{y_\emptyset}} \big[\tilde{u}(p) - \epsilon\big]\langle D_q\tilde{a}, \hat{n}\rangle \tilde{f}(p)\,dS(p). \end{aligned}$$

Since $0 \leqslant \tilde{u} \leqslant \epsilon$ in $\tilde{U}_\epsilon$, we combine the last inequality with (6.7) to obtain

$$-\frac{\tilde{L}_\epsilon(\tilde{u})}{\epsilon} \leqslant \int_{\tilde{U}_\epsilon} \big|\nabla_p \cdot \big(\tilde{f}(p)D_q\tilde{a}\big(\mathbf{0}, \tilde{u}(p), p\big)\big)\big|\,dp + \int_{\tilde{U}_\epsilon \cap \partial X_{y_\emptyset}} \big|\langle D_q\tilde{a}, \hat{n}\rangle \tilde{f}(p)\big|\,dS(p). \tag{6.8}$$

Notice that domain monotonicity implies the $\epsilon \to 0$ limit of the last expressions above is given by integrals over the limiting domain $\tilde{U}_0 = \cap_{\epsilon > 0}\tilde{U}_\epsilon$. Assume now the interior of the convex set $\tilde{U}_0$ is empty, so that $\tilde{U}_0$ has dimension at most $n-1$. Then the volume $|\tilde{U}_\epsilon| = o(1)$, hence the first integral in the right-hand side dwindles to zero as $\epsilon \to 0$, (recalling that $\tilde{u}$ is Lipschitz, $\tilde{f} \in W^{1,1}$ and $\tilde{a} \in C^3$). Concerning the second term, if the convex set $\tilde{U}_0$ has dimension $n-1$ then its *relative* interior must be disjoint from the boundary of the convex body $X_{y_\emptyset}$, since the latter is assumed to have no $n-1$ dimensional facets. Either way $\tilde{U}_0 \cap \partial X_{y_\emptyset}$ has dimension at most $n-2$, which, together with the fact that the boundary trace of $\tilde{f}$ is integrable, implies that

$$\int\limits_{\tilde{U}_\epsilon \cap \partial X_{y_\emptyset}} \tilde{f}(p)\, dS(p) = o(1)$$

as $\epsilon \to 0$. All in all, we have shown $L_\epsilon(\tilde{u}) = o(\epsilon)$ as $\epsilon \to 0$ whenever $\tilde{U}_0$ has empty interior, which — as was explained above — contradicts the asserted optimality of the strategy $\tilde{u}$. However, even if $\tilde{U}_0$ has non-empty interior, more must be true to avoid inferring the contradictory conclusion $L_\epsilon(\tilde{u}) = o(\epsilon)$ as $\epsilon \to 0$ from (6.8): one of the two limiting integrals

$$\int\limits_{\tilde{U}_0} \left|\nabla_p \cdot \big(\tilde{f}(p) D_q \tilde{a}\big(\mathbf{0}, \tilde{u}(p), p\big)\big)\right| dp > 0 \quad \text{or} \quad \int\limits_{\tilde{U}_0 \cap \partial X_{y_\emptyset}} \left|\langle D_q \tilde{a}, \hat{n}\rangle\right| \tilde{f}(p)\, dS(p) > 0$$

must be non-vanishing. In either case, the $W^{1,1}$ density $\tilde{f}$ must be positive somewhere in $\tilde{U}_0$, whose interior therefore includes a positive fraction of the agents. Since $\tilde{u}$ is differentiable with vanishing gradient on the interior of $\tilde{U}_0$, there is no ambiguity in the strategy of these agents: they respond to $\tilde{u}$ by choosing the outside option. □

Appendix A. Minimal *b*-convex potentials

The purpose of this appendix is to establish a mathematical result (and some terminology) needed in the last part of the uniqueness proof, Theorem 4.6. In particular, we establish a minimality property enjoyed by Rochet's construction of a b-convex function for which $\partial^b u$ contains a prescribed set [30]; Rochet's construction is modeled on the analogous construction by Rockafellar of a convex function u whose subdifferential ∂u contains a given cyclically monotone set [33].

Recall a relation $S \subset \overline{X} \times \overline{Y}$ is *b-cyclically monotone* if for each integer $k \in \mathbf{N}$ and k-tuple of points $(x_1, y_1), \ldots, (x_k, y_k) \in S$, the inequality

$$\sum_{i=1}^{k} b(x_i, y_i) - b(x_{i+1}, y_i) \geqslant 0 \tag{A.1}$$

holds with $x_{k+1} := x_1$. For a function $u : \overline{X} \longmapsto \mathbf{R} \cup \{+\infty\}$, the relation $\partial^b u \subset \overline{X} \times \overline{Y}$ consists of those points (x, y) such that

$$u(\cdot) \geqslant u(x) + b(\cdot, y) - b(x, y) \tag{A.2}$$

holds throughout $\overline{X}$. Rochet's generalization of Rockafellar's theorem asserts that $S \subset \overline{X} \times \overline{Y}$ is b-cyclically monotone if and only if there exists a b-convex function $u : \overline{X} \longmapsto \mathbf{R} \cup \{+\infty\}$ such that $S \subset \partial^b u$. Here we need to extract a certain minimality property from its proof.

Lemma A.1. *Given a b-cyclically monotone $S \subset \overline{X} \times \overline{Y}$ and $(x_0, y_0) \in S$, there is a b-convex function u vanishing at x_0 and satisfying $S \subset \partial^b u$, which is minimal in the sense that $u \leqslant \tilde{u}$ for all $\tilde{u} : \overline{X} \longmapsto \mathbf{R} \cup \{+\infty\}$ vanishing at x_0 with $S \subset \partial^b \tilde{u}$.*

Proof. Given a b-cyclically monotone $S \subset \overline{X} \times \overline{Y}$ and $(x_0, y_0) \in S$, Rochet [30] verified the elementary fact that the following formula defines a b-convex function u for which $S \subset \partial^b u$:

$$u(\cdot) = \sup_{k\in\mathbf{N}} \sup_{(x_1,y_1),\dots,(x_k,y_k)\in S} b(\cdot, y_k) - b(x_0, y_0) + \sum_{i=1}^{k} b(x_i, y_{i-1}) - b(x_i, y_i). \tag{A.3}$$

Taking $k = 0$ shows $u(x_0) \geqslant 0$, while the opposite inequality $u(x_0) \leqslant 0$ follows from b-cyclical monotonicity (A.1) of S. Now suppose $\tilde{u}(x_0) = 0$ and $S \subset \partial^b \tilde{u}$. For each $k \in \mathbf{N}$ and k-tuple in S, we claim $\tilde{u}(\cdot)$ exceeds the expression under the supremum in (A.3). Indeed, $(x_i, y_i) \in S \subset \partial^b \tilde{u}$ implies

$$\tilde{u}(x_{i+1}) \geqslant \tilde{u}(x_i) + b(x_{i+1}, y_i) - b(x_i, y_i)$$

and $\tilde{u}(x_i) < \infty$, by evaluating (A.2) at x_i and at x_0. Summing the displayed inequalities from $i = 0, \dots, k$, arbitrariness of $x_{k+1} \in \overline{X}$ yields the desired result: $\tilde{u}(x_{k+1}) \geqslant u(x_{k+1})$. □

References

[1] N. Ahmad, H.K. Kim, R.J. McCann, Optimal transportation, topology and uniqueness, Preprint at http://www.math.toronto.edu/mccann.

[2] M. Armstrong, Multiproduct nonlinear pricing, Econometrica 64 (1996) 51–75.

[3] P. Barelli, S. Basov, M. Bugarin, I.P. King, The inclusiveness of exclusion, Preprint at http://ssrn.com/abstract=1584969.

[4] D.P. Baron, R.B. Myerson, Regulating a monopolist with unknown costs, Econometrica 50 (1982) 911–930.

[5] S. Basov, Multidimensional Screening, Springer-Verlag, Berlin, 2005.

[6] G. Buttazzo, G. Carlier, Optimal spatial pricing strategies with transportation costs, in: Contemp. Math., vol. 514, 2010, pp. 105–121.

[7] L.A. Caffarelli, R.J. McCann, Free boundaries in optimal transport and Monge–Ampère obstacle problems, Ann. of Math. (2) 171 (2010) 673–730.

[8] G. Carlier, A general existence result for the principal–agent problem with adverse selection, J. Math. Econ. 35 (2001) 129–150.

[9] G. Carlier, T. Lachand-Robert, Regularity of solutions for some variational problems subject to convexity constraint, Comm. Pure Appl. Math. 54 (2001) 583–594.

[10] R. Deneckere, S. Severinov, Multi-dimensional screening with a one-dimensional allocation space, Preprint at http://www.econ.ntu.edu.tw/sem-paper/97_2/micro_980319.pdf.

[11] I. Ekeland, S. Moreno-Bromberg, An algorithm for computing solutions of variational problems with global convexity constraints, Numer. Math. 115 (2010) 45–69.

[12] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, Stud. Adv. Math., CRC Press, Boca Raton, 1992.

[13] A. Figalli, Y.-H. Kim, R.J. McCann, Continuity and injectivity of optimal maps for non-negatively cross-curved costs, Preprint at http://www.arXiv.org:0911.3952.

[14] W. Gangbo, R.J. McCann, The geometry of optimal transportation, Acta Math. 177 (1996) 113–161.

[15] R. Guesnerie, J.-J. Laffont, Taxing price makers, J. Econ. Theory 19 (1978) 423–455.

[16] Y.-H. Kim, R.J. McCann, Continuity curvature, and the general covariance of optimal transportation, J. Eur. Math. Soc. (JEMS) 12 (2010) 1009–1040.

[17] Y.-H. Kim, R.J. McCann, Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular), J. Reine Angew. Math., in press, Preprint at arXiv:0806.0351v1 [math].

[18] J.-J. Laffont, E. Maskin, J.-C. Rochet, Optimal nonlinear pricing with two dimensional characteristics, in: R. Radner, T. Groves, S. Reiter (Eds.), Information, Incentives, and Economic Mechanisms, University of Minnesota Press, Minneapolis, 1987, pp. 256–266.

[19] G. Loeper, On the regularity of solutions of optimal transportation problems, Acta Math. 202 (2009) 241–283.

[20] G. Loeper, Regularity of optimal maps on the sphere: the quadratic cost and the reflector antenna, Arch. Ration. Mech. Anal. 199 (2011) 269–289.

[21] X.-N. Ma, N. Trudinger, X.-J. Wang, Regularity of potential functions of the optimal transportation problem, Arch. Ration. Mech. Anal. 177 (2005) 151–183.

[22] R.P. McAfee, J. McMillan, Multidimensional incentive compatibility and mechanism design, J. Econ. Theory 46 (1988) 335–354.

[23] L.J. Mirman, D.S. Sibley, Optimal nonlinear prices for multiproduct monopolies, Bell J. Econ. 11 (1980) 659–670.

[24] J.A. Mirrlees, An exploration in the theory of optimum income taxation, Rev. Econ. Stud. 38 (1971) 175–208.

[25] P.K. Monteiro, F.H. Page Jr., Optimal selling mechanisms for multiproduct monopolists: incentive compatibility in the presence of budget constraints, J. Math. Econ. 30 (1998) 473–502.

[26] M. Mussa, S. Rosen, Monopoly product quality, J. Econ. Theory 18 (1978) 301–317.

[27] M. Quinzii, J.-C. Rochet, Multidimensional screening, J. Math. Econ. 14 (1985) 261–284.

[28] K.W.S. Roberts, Welfare considerations of nonlinear pricing, Econ. J. 89 (1979) 66–83.

[29] J.-C. Rochet, The taxation principle and multitime Hamilton–Jacobi equations, J. Math. Econ. 14 (1985) 113–128.

[30] J.-C. Rochet, A necessary and sufficient condition for rationalizability in a quasi-linear context, J. Math. Econ. 16 (1987) 191–200.

[31] J.-C. Rochet, P. Choné, Ironing sweeping and multidimensional screening, Econometrica 66 (1998) 783–826.

[32] J.-C. Rochet, L.A. Stole, The economics of multidimensional screening, in: M. Dewatripont, L.P. Hansen, S.J. Turnovsky (Eds.), Advances in Economics and Econometrics, Cambridge University Press, Cambridge, 2003, pp. 115–150.

[33] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1972.

[34] T. Sei, A Jacobian inequality for gradient maps on the sphere and its application to directional statistics, Preprint at arXiv:0906.0874v2.

[35] M. Spence, Competitive and optimal responses to signals: An analysis of efficiency and distribution, J. Econ. Theory 7 (1974) 296–332.

[36] M. Spence, Multi-product quantity-dependent prices and profitability constraints, Rev. Econ. Stud. 47 (1980) 821–841.

[37] N.S. Trudinger, X.-J. Wang, On the second boundary value problem for Monge–Ampère type equations and optimal transportation, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 8 (2009) 1–32.

[38] C. Villani, Optimal Transport. Old and New, Grundlehren Math. Wiss. (Fundamental Principles of Mathematical Sciences), vol. 338, Springer, New York, 2009.

[39] R. Wilson, Nonlinear Pricing, Oxford University Press, Oxford, 1993.