JOTA manuscript No.
(will be inserted by the editor)

Maximum Entropy Estimates for Risk-Neutral Probability Measures with non-Strictly-Convex Data

Christopher Bose · Rua Murray

Abstract This article investigates use of the Principle of Maximum Entropy for approximation of the risk-neutral probability density on the price of a financial asset as inferred from market prices on associated options. The usual strict convexity assumption on the market-price to strike-price function is relaxed, provided one is willing to accept a partially supported risk-neutral density. This provides a natural and useful extension of the standard theory. We present a rigorous analysis of the related optimization problem via convex duality and constraint qualification on both bounded and unbounded price domains. The relevance of this work for applications is in explaining precisely the consequences of any gap between convexity and strict convexity in the price function. The computational feasibility of the method and analytic consequences arising from non-strictly-convex price functions are illustrated with a numerical example.

Keywords Financial mathematics, risk-neutral probability density, maximum entropy method, moment constraint, Lagrangian duality.

Mathematics Subject Classification (2000) Primary 91G80 Secondary 49K21, 41A46

This work has been supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

C. Bose
Department of Mathematics and Statistics, University of Victoria, PO Box 3060, STN CSC, Victoria, B.C. Canada V8P 5C2
E-mail: cbose@uvic.ca

R. Murray (Corresponding author)
Department of Mathematics and Statistics, University of Canterbury, Private Bag 4800 Christchurch 8140, New Zealand. E-mail: rua.murray@canterbury.ac.nz

Notice

1 Introduction

This paper is motivated by a financial application: the robust calculation of an underlying price distribution of a financial asset, given market observations on the price of options based on that asset. This is the so-called **risk-neutral density** on prices.

Breeden and Litzenberger ([1], 1978) observed that the risk-neutral density is proportional to the second derivative of the market price of the corresponding option with respect to its strike price. Of course, in practice one has only partial information (in fact, a discrete sample) of this function. Nevertheless, there have been a number of investigations attempting to use these samples of market data to estimate the second derivative, and hence to compute approximations to the risk-neutral density (see Shimko ([2], 1993) for example). Clearly, such an approach is highly sensitive to the type of interpolation scheme employed.

Buchen and Kelley ([3], 1996) investigated the use of the Maximum Entropy Principle (MEP) as a natural, non-parametric method for extrapolating from observed market data so-as to estimate the underlying price density. Avellaneda *et. al.* ([4], 1997) and Avellaneda ([5], 1998) also discuss this method. The key motivation for using MEP is that the entropy objective leads to a density that is maximally noncommittal with respect to missing or unknown data (in a sense that can be made precise in many cases; see Jaynes ([6], 1982)).

Borwein *et. al.* ([7], 2003) look carefully at the foundations for MEP calculations in this context, showing that the mathematical theory of convex duality can be used to justify these computations rigorously. The duality theory also provides efficient computational approaches for the MEP optimization based on a finite-dimensional global optimization, provided the assumption of *strict convexity* is satisfied by the observed market data. Conversely, their analysis shows that constructing MEP densities given non-convex, or even non-strictly-convex price data is impossible since no traditional MEP solution exists in the presence of such constraints.

An important consequence of the traditional MEP approach is that the resulting risk-neutral density is expressed as the exponential of some real-valued function over the given price domain (typically, the nonnegative real line) implying that arbitrarily high asset prices can arise. Another contribution of Borwein *et. al.* ([7], 2003) was to provide a rigorous extension of MEP to the case of an *a priori* compactly supported price interval.

The purpose of this paper is to start where Borwein *et. al.* left off, developing a rigorous MEP approach that can accommodate more complex price domains than simply unbounded or bounded intervals. We explain precisely how this phenomenon is linked to the convexity property in the market price data and how to derive consistent MEP densities for data that lies at the **boundary**: convex, but not-strictly-convex market data. These boundary cases arise when dealing with real (noisy) data where the theoretical strict convexity condition may be violated: see Remark 3.3 below, Buchen and Kelley ([3], 1996 – for numerical examples of this phenomenon), and Guo ([8], 2001 – where convex smoothing methods are proposed to address this issue). Specific regularisation strategies for maximum entropy approaches to S&P500 data are employed by He *et. al* ([9], 2008). Recent articles by Neri and Schneider ([10], 2012) and Rodriguez and Santosa ([11], 2012), for example, show that the goal of obtaining robust maximum entropy estimates from realistic financial data continues to be an active, productive and challenging area of research in financial mathematics.

At first glance, the loss of strict-convexity appears at odds with traditional pricing theory in the sense that arbitrage free pricing is deemed to be equivalent to the existence of a strictly positive risk-neutral density. However, in the absence of strict convexity in the market data, our analysis details precisely the way in which arbitrage free pricing must be achieved by risk neutral measures. Further discussion may be found in Remark 3.4.

In the next section we set up the problem and introduce the MEP approach in detail. The MEP density is the solution to an (infinite-dimensional) constrained optimization problem given a (finite) set of market observations. We introduce the notion of *impossible price intervals* given a set of market observations in Section 3; the complement of the union of these intervals contains the support of every consistent price density. We call this the *reduced domain* (see Definition 3.1). Restricting our optimization problem to functions supported on the reduced domain, we obtain necessary and sufficient conditions for existence of a fully supported risk-neutral density (constraint qualification). In Section 4 we use convex duality to reformulate our problem as a finite-dimensional optimization problem in the case of a bounded reduced domain. Because the dual (maximization) problem is *concave* and smooth, it suffices to solve (finitely many) equations for the vanishing of the gradient of the objective. These equations also imply strong duality, provided the constraint qualification derived in Section 3 holds. In Section 4.2 we explain the connection between our geometric approach and the topologically motivated constraint qualification given in Borwein *et. al.* ([7], 2003).

In Section 4.3 we adapt the arguments above for the case of an unbounded reduced domain, obtaining strong duality in the context of suitable paired

Köthe spaces (see Dieudonne ([12], 1951) for background). We also derive the important *conjugation through the integral* property as a consequence of results of Maréchal ([13], 2001). Finally, in Section 5 we summarize our findings and suggest directions for future work.

2 Consistent Market Data and Impossible Prices

2.1 Notation and Definitions

Let us begin by defining the problem. Let m prices $0 \leq k_1 < k_2 < \cdots < k_m$ be given in the interval $I = [0, K[$ with observed market prices $d_1, d_2, \ldots d_m$ for European call options at corresponding strike price k_i and expiration time $T > 0$ (fixed for this problem). We allow either $K < \infty$ or $K = \infty$. For simplicity we assume the real rate of interest over the time interval $[0, T]$ is zero, so no discounting is necessary in the calculation of the prices d_i from the k_i and T. Henceforth, we regard T as fixed, and suppress it in the notation for market data. Let $p = p(x)$ be a probability density on prices $x \in I$ which is consistent with these market observations. More precisely, for $i = 1, 2, \ldots m$, let

$$c_i(x) := \begin{cases} 0 & \text{for } x < k_i, \\ (x - k_i) & \text{for } x \geq k_i,\ x \in I. \end{cases} \tag{1}$$

Then, to be consistent, the distribution p, given the data, must satisfy

$$\begin{aligned} d_0 &:= 1 = \int_I p(x)\, dx, \\ d_i &:= \int_I c_i(x) p(x)\, dx, \quad \text{for } i = 1, 2, \ldots m. \end{aligned} \tag{2}$$

We say that $\{k_i; d_i\}$ is a set of **consistent market data** if and only if there exists a density p on I such that Equations (2) hold. Given consistent data, a non-negative $p \in L^1(I)$ is called a **consistent price density given the market data** if and only if it satisfies Equations (2). In this article we will use the MEP to systematically choose such a consistent price density, thereby approximating the presumed underlying risk-neutral measure.

The mathematical problem is formulated as follows:

Define

$$\phi(t) := \begin{cases} t \ln t & \text{for } t > 0, \\ 0 & \text{for } t = 0, \\ +\infty & \text{for } t < 0. \end{cases} \tag{3}$$

Assume first that $K < \infty$ so that $I = [0, K[$ is a bounded interval. Then the functions $c_i \in L^\infty$ and for each $p \in L^1(I)$ we can define $\mathbb{A}p \in \mathbb{R}^{m+1}$ by

$$\begin{aligned}(\mathbb{A}p)_0 &= \int_0^K p(x)\,dx,\\ (\mathbb{A}p)_i &= \int_0^K c_i(x)p(x)\,dx, \quad \text{for } i = 1, 2, \dots m.\end{aligned} \tag{4}$$

The optimization problem to be solved is

$$(P) \qquad \begin{cases} \text{Minimize} & \int_0^K \phi(p(x))\,dx \\ \text{Subject to} & p \in L^1(I),\ \mathbb{A}p = (d_0, d_1, \dots d_m)^T = (1, d_1, \dots d_m)^T. \end{cases}$$

Note we have chosen to **minimize** the negative of the usual entropy functional. This is simply a mathematical convenience, and of no consequence since we are mainly interested in p rather than the value of the problem.

In terms of the optimization problem (P), the market data will be consistent if and only if the feasibility set is non-empty, and every element of the feasibility set will be a consistent probability density, given the market data. By choosing p satisfying (P) we obtain a **MEP approximate risk-neutral density consistent with market data** $\{k_i; d_i\}$. In the next sections we will solve this problem by way of Lagrangian duality with respect to the the spaces L^1 and $\mathbb{R}^{m+1}$. If *strong duality* can be established, one expects the optimal density to take the form

$$p(x) \sim \exp \varphi(x) > 0 \tag{5}$$

(where φ is a real-valued function over the specified price interval I). The positivity of p is inextricably linked with the strict convexity of market data.

When $K = \infty$ the setup above needs to be modified since the linear mapping $\mathbb{A}$ in (4) is well defined only on a proper subspace of L^1. Indeed, for $p \in L^1[0, \infty[$, only when $\int_0^\infty |p(x)|\, x\,dx < \infty$ is each product $c_i(x)p(x)$ integrable. The set of such p forms a dense proper subspace of L^1, so the formal primal-dual optimization must be carried out with respect to a set of paired vector spaces in the sense of Rockafellar ([14], 1968). Details required for this setting are presented in the last part of Section 3 and in Section 4.3.

2.2 On Impossible Prices

Suppose $\{k_i; d_i\}$ is a set of consistent market data. A measurable set $B \subseteq I$ is **a set of impossible prices** (given the market data $\{k_i; d_i\}$) if and only if Equation (2) implies $\int_B p(x)dx = 0$. We say that B is **the** set of impossible prices if and only if it is a maximal set of impossible prices up a set of measure zero[1].

[1] B is a set of impossible prices such that if B' is another set of impossible prices then $\nu(B' \setminus B) = 0$, where ν denotes Lebesgue measure on I.

Let us review an illustrative example suggested by Borwein *et. al.* ([7], 2003).

Example 2.1 Let $m = 2$. Suppose $0 = k_1 < k_2 < K$ and $d_1 - d_2 = k_2$. Then any consistent probability density p must satisfy

$$k_2 = d_1 - d_2 = \int_I (c_1 - c_2)(x) p(x)\, dx.$$

Since $0 \leq (c_1 - c_2)(x) \leq k_2 - k_1 = k_2$ with $(c_1 - c_2)(x) = k_2$ if and only if $x \geq k_2$ we conclude that $\int_0^{k_2} p(x)\, dx = 0$. Hence $B = [0, k_2]$ is a set of impossible prices given this data. Provided d_1 lies in the interval $k_2 < d_1 < K$, one can find a probability density p on $[k_2, K[$ such that $d_1 = \int_I x p(x)\, dx$. Such a choice of p will automatically satisfy Equation (2) since

$$\int_I c_2(x) p(x)\, dx = \int_{k_2}^{K} (x - k_2) p(x)\, dx = d_1 - k_2 = d_2.$$

Hence the market data will be consistent. However, every consistent density on the price x of the asset at time T must *a priori* assign the value zero to all prices lower than k_2. It follows that MEP solution(s) to the optimization (P) cannot be obtained through application of the theory of convex duality and expression (5).

Once again, provided $k_2 < d_1 < K$, one can in fact construct a consistent probability density fully supported on $[k_2, K[$, hence $[0, k_2]$ is **the** set of impossible prices[2]. As a consequence of the computations to follow, we will show that all MEP optimal solutions to this problem will be fully supported on $[k_2, K[$. While apparently unusual, this situation is compatible with conventional arbitrage pricing theory: a trading strategy which sells an option at strike k_1 for price d_1 and purchases an option at strike k_2 for price d_2 generates a claim depending on final price x_T of the underlying asset of

$$\begin{cases} d_1 - d_2 & \text{if} \quad x_T < k_1, \\ d_1 - d_2 - (x_T - k_1) & \text{if} \quad k_1 \leq x_T < k_2, \\ 0 & \text{if} \quad x_T \geq k_2. \end{cases}$$

Since $d_1 - d_2 = k_2 > 0$ in the first case and $d_1 - d_2 - (x_T - k_1) = k_2 - x_T > 0$ in the second, if both options are consistently priced, the absence of arbitrage opportunities dictates that there must be no net profit from this strategy. Hence $x_T \geq k_2$ almost surely. We remark that this situation is not unrealistic: for example, if the underlying asset was a *preferred share*, due to pay a dividend of k_2 on day $T + 1$, the market price x_T can be reasonably assumed to exceed k_2.

[2] It is also easy to see that if $d_1 = k_2$ or $d_1 = K$ there can be no price density p on $[0, K[$ consistent with this data.

Example 2.2 (Derive MEP solutions for Example 2.1) Suppose that $k_1 = 0$, $k_2 = 1$, $K = 2$ and, given $d_1 \in (1,2)$ let $d_2 = d_1 - 1$. Sample optimal solutions are:

$$\begin{array}{lll} d_1 = 1.1 & \Rightarrow & p(x) = \exp(2.30217 - 9.99544\,(x-1))\,\mathbf{1}_{[1,2[}(x), \\ d_1 = 1.5 & \Rightarrow & p(x) = \exp(0)\,\mathbf{1}_{[1,2[}(x), \\ d_1 = 1.9 & \Rightarrow & p(x) = \exp(-7.69327 + 9.99544\,(x-1))\,\mathbf{1}_{[1,2[}(x), \end{array}$$

(all constants to six significant figures). The computations leading to these results will be will be made in the sections to follow.

3 Consistent Data, Feasibility and Impossible Price Intervals

3.1 The Bounded Case ($K < \infty$)

We now determine impossible price sets of the form $[0, k_1[$, $[k_i, k_{i+1}[$ or $[k_m, K[$ for a given set of consistent market data. Indeed, we will show there exists a probability density p satisfying Equation (2) and which is strictly positive on the complement of the union of these intervals; as in the simple example just discussed, the union of the intervals we find will be **the** set of impossible prices.

Consistency of market data can be reformulated in terms of strict convexity of the graph of the market price d against strike price k. To use this characterization, some additional notation is convenient.

Notation. For a given sequence

$$k_0 := 0 \le k_1 < k_2 < \cdots < k_m < k_{m+1} := K < \infty,$$

define a family of **hat functions** $\{h_1, \ldots, h_{m+1}\}$ on $[0, K[$ by linear interpolation of the following values:

$$h_1(k_j) = \begin{cases} 1 & \text{if } j = 0, 1, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and for } i > 1 \quad h_i(k_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that

$$h_1 = \mathbf{1} - \tfrac{c_1 - c_2}{k_2 - k_1}, \quad h_m = \tfrac{c_{m-1} - c_m}{k_m - k_{m-1}} - \tfrac{c_m}{K - k_m}, \quad h_{m+1} = \tfrac{c_m}{K - k_m},$$
$$\text{and} \qquad h_i = \tfrac{c_{i-1} - c_i}{k_i - k_{i-1}} - \tfrac{c_i - c_{i+1}}{k_{i+1} - k_i}$$

for the remaining i. Let $\mathbf{B}$ be the $(m+1) \times (m+1)$ matrix such that

$$\mathbf{h}(x) := (h_1(x), \ldots, h_{m+1}(x))^T = \mathbf{B}\,\mathbf{c}(x) \tag{6}$$

where $\mathbf{c}(x) = (\mathbf{1}_{[0,K[}, c_1(x), \ldots, c_m(x))^T$. Clearly,

$$\operatorname{supp}(h_i) =]k_{i-1}, k_{i+1}[\cap [0, K[$$

where $\operatorname{supp}(f) := \{x \in [0,K[\ : \ f(x) > 0\}$ denotes the **support** of a non-negative function f relative to $[0,K[$. Note that

$$\sum_{i=1}^{m+1} h_i(x) = \mathbf{1}_{[0,K[}, \tag{7}$$

and that the matrix $\mathbf{B}$ depends only on the strike prices k_i and K. □

We begin with a simple lemma.

Lemma 3.1 *Let $I = [0,K[$ be bounded. Suppose that $\{k_i; d_i\}_{i=1}^m$ are consistent market data and p is any consistent price density. Let $\mathbf{B}$ be given by* (6)*, let $\mathbf{d} = (1, d_1, \dots, d_m)^T$ and put*

$$\boldsymbol{\eta} = (\eta_1, \dots, \eta_{m+1})^T := \mathbf{B}\,\mathbf{d}.$$

(a) *For each $i = 1, \dots, (m+1)$, $\eta_i = \int_0^K h_i(x)\,p(x)\,dx$.*
(b) *For each i, $\eta_i \geq 0$ and*

$$\eta_i = 0 \qquad \Leftrightarrow \qquad (k_{i-1}, k_{i+1}) \text{ is an impossible price interval}$$

(interpret $k_0 = 0$, $k_{m+1} = K$, $k_{m+2} = \infty$). Moreover, $\sum_{i=1}^{m+1} \eta_i = 1$ and $\eta_{m+1} < 1$.

Proof (a) Since p is a consistent price density, $1 = \int \mathbf{1}_{[0,K[}(x)\,p(x)\,dx$ and $d_i = \int_0^K c_i(x)\,p(x)\,dx$ for $i = 1, \dots, m$. Thus $\mathbf{d} = \int_0^K \mathbf{c}(x)\,p(x)\,dx$. Hence

$$\boldsymbol{\eta} = \mathbf{B}\,\mathbf{d} = \mathbf{B} \int_0^K \mathbf{c}(x)\,p(x)\,dx = \int_0^K \mathbf{B}\,\mathbf{c}(x)\,p(x)\,dx = \int_0^K \mathbf{h}(x)\,p(x)\,dx.$$

(b) Since $0 \leq h_i(x)$ on $[0,K[$ and $p(x) \geq 0$, we have $(h_i\,p)(x) \geq 0$. Hence, by part (a), $\eta_i \geq 0$. Furthermore, $\eta_i = 0$ if and only if $\operatorname{supp}(h_i) \cap \operatorname{supp}(p) = \emptyset$ (modulo Lebesgue measure). In particular, $]k_{i-1}, k_{i+1}[= \operatorname{supp}(h_i)$ is an impossible price interval and $h_i\,p = 0$ a.e. Next,

$$\sum_{i=1}^{m+1} \eta_i = \int_0^K \sum_{i=1}^{m+1} h_i(x)\,p(x)\,dx = \int \mathbf{1}_{[0,K[}(x)\,p(x)\,dx = 1,$$

by (7) and the fact that p is a probability density supported on $[0,K[$. Finally, suppose that $\eta_{m+1} = 1$. Then $\sum_{i=1}^m \eta_i = 1 - \eta_{m+1} = 0$ so $\eta_1 = \cdots = \eta_m = 0$. Hence

$$[0, k_2[,\ ]k_1, k_3[,\ \dots,\ ]k_m, K[$$

are all impossible price intervals, so $\operatorname{supp}(p)$ is disjoint from

$$\cup_{i=1}^m]k_{i-1}, k_{i+1}[=]0, K[.$$

To avoid this contradiction, $\eta_{m+1} < 1$.

Remark 3.1 Lemma 3.1 has a simple geometric interpretation. Consider the strike-to-market-price function $D(x)$ over $[k_1, k_m]$ obtained by piecewise linear interpolation of the market data (k_i, d_i). For $2 \leq i \leq m-1$ the value of η_i equals the slope of D over $[k_i, k_{i+1}]$ minus its slope over $[k_{i-1}, k_i]$ so the specification $\eta_i \geq 0$ implies that D is convex over the interval $[k_{i-1}, k_{i+1}]$. Extend D continuously to the interval $[0, \infty[$ giving it slope -1 on $[0, k_1]$ and setting $D \equiv 0$ on $[K, \infty[$. The conditions in Lemma 3.1 b) require that D be a convex function on $[0, \infty[$ with slope starting at -1 near zero and rising to zero on $[K, \infty[$ (this is the sum condition). Moreover, the same arguments show this convexity constraint on market data must hold in the presence of any risk neutral probability and not just for absolutely continuous probabilities.

Remark 3.2 In the notation of Borwein *et. al.* ([7, Prop. 2] ,2003), $\xi_i = \eta_{i+1}$ for $0 \leq i < m$. The easiest way to see this is to observe that $\mathbf{M}^T = \mathbf{B}^{-1}$, where $\mathbf{M}$ is in [7, Lemma 1]. In Borwein *et. al.* ([7], 2003), a geometrically derived constraint qualification condition is used to write down the necessary conditions for feasible market data ($\xi_i \geq 0$); the sufficient conditions $\eta_i > 0$ for $0 \leq i \leq m$ allow for the solution of problem (P) via convex duality. The condition in Borwein *et. al.* ([7], 2003) immediately following Proposition 2 for feasibility and strong duality in the case $K < \infty$ should read

$$\langle \mathbf{N}^{-1}\mathbf{B}(d_1, \ldots, d_m)^T, \mathbf{u} \rangle < 1 - (K - k_m)^{-1}\, d_m.$$

Remark 3.3 Buchen and Kelley ([3], 1996) consider the numerical solution of the MEP problem with simulated data. They comment that noise in the data can lead (easily) to violations of convexity of the strike-price curve, and this causes numerical difficulties, leading to trapping at a "$\cdots$ local maximum, far from its global maximum" [3, p154]. In fact, the vector $\boldsymbol{\eta}$ measures convexity, with negative values reflecting convexity violation. The numerical maximization procedure involves solving $\partial_i Q = 0$ (see Equation (19) below), and when $\eta_i \leq 0$ the absence of solutions will cause inevitable numerical difficulties.

Remark 3.4 The value of η_i should be interpreted as the fair price for a *butterfly spread* portfolio comprising a_1 options at strike k_{i-1}, a_2 options at strike k_i, a_3 options at strike k_{i+1} with $a_1\, c_{i-1}(x) + a_2\, c_i(x) + a_3\, c_{i+1}(x) = h_i(x)$ (*cf.* Equation (6)) ($a_1, a_3 > 0$ and $a_2 < 0$). Since this portfolio generates a positive pay-off for an underlying price $x_T \in]k_{i-1}, k_{i+1}[$, when $\eta_i = 0$, fair pricing of the individual options implies that $x_T \in]k_{i-1}, k_{i+1}[$ occurs with probability 0. Taking this reasoning further, when $\eta_i < 0$ there is no risk-neutral probability measure (by Theorem 3.1), and an arbitrage opportunity exists: one simply constructs the above portfolio for a guaranteed upfront profit of $|\eta_i|$. Moreover, since the observed market data are inconsistent, there is no risk-neutral measure, and potentially $x_T \in]k_{i-1}, k_{i+1}[$ creating a possibility of further profit from exercising the appropriate combination of options.

Theorem 3.1 (Necessary conditions for consistent price data) *Let $\{k_i; d_i\}_{i=1}^m$, $(d_i \geq 0)$ be a set of consistent market data with $k_i < k_{i+1}$ for each $i = 1, \ldots, m-1$ and $k_m < K < \infty$. Let $\boldsymbol{\eta} = \mathbf{B}\,(1, d_1, \ldots, d_m)^T$ (where $\mathbf{B}$ is as in (6)). Then the appropriate one of the following conditions holds:*

- (C) *if $k_1 > 0$ then $\eta_i > 0$ implies at least one of η_{i-1} or η_{i+1} is also positive for $i = 2, \ldots, m$ and $\eta_{m+1} > 0 \Rightarrow \eta_m > 0$;*
- (C') *if $k_1 = 0$, then condition (C) is augmented by $\eta_1 > 0 \Rightarrow \eta_2 > 0$.*

Proof Let p be a consistent price density. Suppose first that $1 < i < m+1$. By Lemma 3.1(b), if $\eta_{i-1} = 0 = \eta_{i+1}$ then both (k_{i-2}, k_i) and (k_i, k_{i+2}) are impossible price intervals. In particular, $]k_{i-1}, k_{i+1}[\cap \operatorname{supp}(p) \subset \{k_i\}$, so $\eta_i = 0$ by Lemma 3.1(a). Similarly, if $\eta_m = 0$ then $\eta_{m+1} = 0$. This establishes (C). In case $k_1 = 0$, $\operatorname{supp}(h_1) =]0, k_2[\subset]0, k_3[=]k_1, k_3[$ so if $\eta_2 = 0$ then $\operatorname{supp}(h_1)$ is contained in an impossible price interval (by Lemma 3.1(b)), so $\eta_1 = \int_0^K h_1(x)\, p(x)\, dx = 0$.

According to Lemma 3.1, consistent price data may *a priori* exclude certain prices in the interval I. These intervals are characterized by $\eta_i = 0$, and (C) (or (C')) provides some additional clarification. It turns out that these conditions are also sufficient for the consistency of market data and the existence of a solution to problem (P). In establishing these facts it is helpful to replace (P) with an equivalent optimization problem; some additional notation is required for this.

Definition 3.1 Given market data $\{k_i; d_i\}$ let $\boldsymbol{\eta} = \mathbf{B}\,(1, d_1, \ldots, d_m)^T$ (where $\mathbf{B}$ is as in Equation (6)). If each $\eta_i \geq 0$ and (C) (or (C')) holds define[3]

$$I_0 := I \setminus \cup_{\{i\,:\,\eta_i = 0\}}]k_{i-1}, k_{i+1}[. \tag{8}$$

The set I_0 is the *reduced domain* (as described in Section 1).

Next, define $\mathbb{B} : L^1[0, K[\to \mathbb{R}^{m+1}$ by

$$(\mathbb{B}p)_i := \int_0^K h_i(x)\, p(x)\, dx = \int_0^K (\mathbf{B}\mathbf{c}(x))_i\, p(x)\, dx = [\mathbf{B}\,\mathbb{A}p]_i. \tag{9}$$

In this notation, the constraints $\mathbb{A}p = \mathbf{d}$ are simply

$$\mathbb{B}p = \mathbf{B}\mathbb{A}\,p = \mathbf{B}\,\mathbf{d} = \boldsymbol{\eta}.$$

When p is a consistent price density, Lemma 3.1 implies that $\operatorname{supp}(p) \subseteq I_0$ (see Lemma 3.2 below) so the integrals defining $\mathbb{B}$ can be restricted to I_0 without effecting the value of $\mathbb{B}p$. It is fruitful to make this domain reduction explicit

[3] In case $\eta_1 = 0$, remove the interval $[0, k_2[$ instead of $]0, k_2[$ so that 0 does not become an isolated point of I_0.

in the definition of the constraints. Let $\boldsymbol{\eta}$ and I_0 be as in Definition 3.1 (we do not need *a priori* that the price data are consistent) and define the *domain reduced operator* $\mathbb{B}_0 : L^1(I_0) \to \mathbb{R}^{m+1}$ by

$$(\mathbb{B}_0 p)_i = \int_{I_0} h_i(x)\, p(x)\, dx. \tag{10}$$

We now define:

$$(P_0) \qquad \begin{cases} \text{Minimize} & \int_{I_0} \phi(p(x))\, dx \\ \text{Subject to} & p \in L^1(I_0),\ \mathbb{B}_0 p = \boldsymbol{\eta}. \end{cases} \tag{11}$$

Example 3.1 (Example 2.1 revisited) Recall that $0 = k_1 < k_2 < K < \infty$ and $d_1 = d_2 + k_2 \in]k_2, K[$. Then

$$\mathbf{B} = \begin{bmatrix} 1 & -1/k_2 & 1/k_2 \\ 0 & 1/k_2 & -(1/k_2 + 1/(K-k_2)) \\ 0 & 0 & 1/(K-k_2) \end{bmatrix}$$

and $\boldsymbol{\eta} = \mathbf{B}\,(1, d_1, d_2)^T = \frac{1}{K-k_2}\,(0, K-d_1, d_1-k_2)^T$. If $d_1 = K$ or $d_1 = k_2$ then the price data is inconsistent by virtue of Lemma 3.1. If $k_2 < d_1 < K$ then $\eta_1 = 0$ while $\eta_2, \eta_3 > 0$. Hence (C') is satisfied,

$$I_0 = [0, K[\ \setminus\ ]0, k_2[$$

and (P_0) is formulated as

$$\begin{cases} \text{Minimize} & \int_{k_2}^{K} p(x)\, \log p(x)\, dx \\ \text{Subject to} & p \in L^1[k_2, K[,\ \ \int_{k_2}^{K} (x-k_2)\, p(x)\, dx = d_2,\ \ \int_{k_2}^{K} p(x)\, dx = 1. \end{cases}$$

Theorem 3.2 (Sufficient conditions for consistency of market data) *Let $\{k_i, d_i\}_{i=1}^m$ be market data, and let $\boldsymbol{\eta} = \mathbf{B}\,(1, d_1, \cdots, d_m)^T$ be such that the conditions in Definition 3.1 hold ($\mathbf{B}$ is defined in (6)); let I_0 be defined as in (8). If condition (C) (or (C')) holds then the market data are consistent, and any p which is feasible for (P) has $\operatorname{supp}(p) \subset I_0$ and satisfies $\mathbb{B}_0 p = \boldsymbol{\eta}$. Moreover, $[0, K[\ \setminus I_0$ is* **the** *set of impossible prices associated to the market data.*

The proof of Theorem 3.2 is deferred until the end of Section 4.1 where it is a simple consequence of the duality arguments required for a rigorous derivation of the MEP solution to (P_0). For now, we give two more lemmas.

Lemma 3.2 *Let $\{k_i; d_i\}_{i=1}^m$ be price data, and let $\boldsymbol{\eta} = \mathbf{B}\,(1, d_1, \cdots, d_m)^T$ be such that the conditions in Definition 3.1 hold ($\mathbf{B}$ is defined in (6)); let I_0 be defined as in (8). Then p is feasible for (P) if and only if p is supported on I_0 and is feasible for (P_0). Moreover, if the market data is consistent, the values of optimization problems (P) and (P_0) are identical.*

Proof First, note that for $p \in L^1[0,K[$, $\mathbb{B}p = \boldsymbol{\eta}$ if and only $\mathbb{A}p = \mathbf{d}$. If p is feasible for (P) then the latter condition holds, p is a consistent price density, and by Lemma 3.1(b), $(I \setminus I_0) = \cup_{\{i\,:\,\eta_i=0\}}]k_{i-1}, k_{i+1}[$ is essentially disjoint from $\mathrm{supp}(p)$. Hence $p = \mathbf{1}_{I_0}\, p$ almost everywhere, and

$$\mathbb{B}_0 p = \int_0^K \mathbf{h}(x)\, \mathbf{1}_{I_0}(x)\, p(x)\, dx = \int_0^K \mathbf{h}(x)\, p(x)\, dx = \mathbb{B}p = \boldsymbol{\eta}.$$

Moreover, since $\mathrm{supp}(p) \subseteq I_0$, we have $\phi(p(x)) = \phi(p(x))|_{I_0}$ and

$$\int_0^K \phi(p(x))\, dx = \int_{I_0} \phi(p(x))\, dx.$$

Therefore, the value of (P_0) is less than or equal to the value of (P). Conversely, if $\mathrm{supp}(p) \subseteq I_0$ and $\mathbb{B}_0 p = \boldsymbol{\eta}$ then $\mathbb{B}\, p = \mathbb{B}(\mathbf{1}_{I_0} p) = \mathbb{B}_0 p = \boldsymbol{\eta}$. Hence p is feasible for (P) and

$$\int_{I_0} \phi(p(x))\, dx = \int_0^K \phi(p(x))\, dx.$$

The value of (P) is therefore less than or equal to the value of (P_0), and the proof is complete.

Lemma 3.3 *Let $\{k_i; d_i\}_{i=1}^m$ be price data satisfying $k_i < k_{i+1}$ for $1 \le i < m$ with $\boldsymbol{\eta}$ and I_0 as in Definition 3.1. If condition* (C) *(or* (C')*) is satisfied then:*

(a) *I_0 is a non-empty union of closed intervals of the form $[k_l, k_r]$ $(l < r)$, (and possibly $[k_L, K[$);*
(b) *Given $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$ let $\varphi(x) = \sum_{i=1}^{m+1} \lambda_i\, h_i(x)$. Then $\varphi\, \mathbf{1}_{I_0} \ge 0$ if and only if $\lambda_i \eta_i \ge 0$ for $i = 1, \dots, m+1$.*
(c) *Let φ be as in part* (b)*. If $\varphi\, \mathbf{1}_{I_0} \ge 0$ and $\langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle = 0$ then $\varphi \mathbf{1}_{I_0} = 0$.*

Proof (a) Since I_0 is obtained from $[0,K[$ by removing open intervals with endpoints from $\{k_i\}$, I_0 is a union of closed intervals with endpoints from $\{k_i\}$ (and possibly K). These intervals have non-empty interior (if $x \in]k_i, k_{i+1}[\setminus I_0$, then either $]k_{i-1}, k_{i+1}[$ or $]k_i, k_{i+2}[$ has been removed from $[0,K[$ in the construction of I_0 so at least one of $\{k_i, k_{i+1}\}$ does not belong to I_0; thus if $k_i, k_{i+1} \in I_0$ then $[k_i, k_{i+1}] \subset I_0$). Finally, the condition $\sum_{i=1}^m \eta_i > 0$ from Lemma 3.1(b) taken together with $C^{(\prime)}$ guarantees at least one pair $\{k_i, k_{i+1}\}$ $(i = 1, \dots, m)$ such that $[k_i, k_{i+1}[\subset I_0$ (interpreting $k_{m+1} = K$).
(b) First, we show that $\lambda_i\, \eta_i \ge 0$ for $i = 1, \dots, m+1$. To this end, note that φ is piecewise linear with $\varphi(k_i) = \lambda_i$ $(i = 1, \dots, m)$. Suppose that $\lambda_i < 0$. Then $\varphi(k_i) < 0$ whereas $\varphi(k_i) \mathbf{1}_{I_0}(k_i) \ge 0$. Thus $k_i \notin I_0$, so $\eta_i = 0$ and $\lambda_i \eta_i = 0$. Otherwise, $\lambda_i \ge 0$ and hence $\lambda_i\, \eta_i \ge 0$. The case $i = m+1$ is left as an exercise (use $\eta_{m+1} > 0 \Rightarrow \eta_m > 0$). The other direction is easy.
(c) Since $\varphi\, \mathbf{1}_{I_0} \ge 0$ and $\langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle = 0$, part (b) implies that $\lambda_i \eta_i = 0$ for all i. If $x \in]k_i, k_{i+1}[\subset I_0$ then $\eta_i, \eta_{i+1} > 0$ (by construction of I_0) so that $\lambda_i = \lambda_{i+1} = 0 = \varphi(k_i) = \varphi(k_{i+1})$. Hence $\varphi \mathbf{1}_{[k_i, k_{i+1}]} = 0$. It follows that $\varphi\, \mathbf{1}_{I_0} = 0$.

Remark 3.5 (Following Remark 3.1) While the condition $\eta_i \geq 0$ implies that the piecewise linear interpolation D of the market data must be convex, when convexity is not strict (*ie* two or more consecutive segments of the interpolating graph have the same slope), values $\eta_i = 0$ result, and corresponding intervals must be disjoint from the support of every consistent price density. The situation where at least one $\eta_i = 0$ is the boundary between consistent and inconsistent market data.

Remark 3.6 As we shall see, the identification of I_0 is also critical for success of the convex duality techniques that we apply in the next section. Without such a reduction, the dual functional, although easily defined (and concave) will fail to attain a global maximum over $\mathbb{R}^{m+1}$ and strong duality will no longer hold. Domain reduction 'flattens out' this dual objective in its non-coercive directions and thus leads to strong duality over a slightly modified function space for the primal problem. This technique has been useful in other contexts, for example, in approximating the invariant density for certain expanding dynamical systems through variational analysis; see Bose and Murray ([15], 2007). In that context, impossible price intervals correspond to transient regions for the dynamics; such regions must be disjoint from the support of any absolutely continuous dynamically invariant measure, an appealing parallel with the current analysis.

Remark 3.7 In Section 4.1, the kernel of $\mathbb{B}_0^T$ is explicitly associated with data where $\eta_i = 0$. Clearly, the vector $\boldsymbol{\eta}$ can be used to verify consistency of price data, form any required domain reduction and to obtain explicit control of the kernel of $\mathbb{B}_0^T$; this is an essential step for the numerical solution of (P).

3.2 Consistency, Feasibility and Impossible Price Intervals on an Unbounded Domain

When $K = \infty$ the statements in Theorems 3.1 and 3.2 are essentially unchanged. However, minor modifications are needed to notation and (some) proofs.

Definition 3.2 When $K = \infty$ the hat functions $\{h_i\}_{i=1}^{m-1}$ are defined as at the beginning of Section 3, however

$$h'_m := \frac{c_{m-1} - c_m}{k_m - k_{m-1}} \qquad \text{and} \qquad h'_{m+1} := c_m.$$

Let $\mathbf{B}'$ be the matrix such that $(h_1, \ldots, h'_m, h'_{m+1})^T = \mathbf{B}'(\mathbf{1}_{[0,\infty[}, c_1, \ldots, c_m)^T$ and let $\boldsymbol{\eta}' = \mathbf{B}'(1, d_1, \ldots, d_m)^T$. If the components η'_i of $\boldsymbol{\eta}'$ satisfy $\eta'_i \geq 0$ for each i, put

$$I_0 = [0, \infty[\setminus \cup_{\{i \,:\, \eta'_i = 0\}}]k_{i-1}, k_{i+1}[$$

(where $]k_m, k_{m+2}[$ is interpreted as $]k_m, \infty[$).

The basic mappings $\mathbb{A}$ and $\mathbb{B}$ are modified as follows. Since $\text{span}\{c_0, c_1, \dots c_m\}$ is no longer a subspace of L^∞, we consider $\mathbb{A}'$, the action of $\mathbb{A}$ restricted to

$$L := \left\{ p \in L^1([0,\infty[) \ : \ \int_0^\infty x\,|p(x)|dx < \infty \right\}, \tag{12}$$

a dense subspace of $L^1(I)$. For $p \in L$ we have both $\mathbb{A}'p \in \mathbb{R}^{m+1}$ and

$$\mathbb{B}'p := \mathbf{B}'\mathbb{A}'p \in \mathbb{R}^{m+1}.$$

The conditions (C) and (C') are based on $\{\eta_i'\}$ and the map $\mathbb{B}_0'$ is defined in the same way as in the $K < \infty$ case.

With these modifications in place, the relevant optimization problems will be denoted (P') and (P_0') respectively. In particular,

$$(P_0') \qquad \begin{cases} \text{Minimize} & \int_{I_0} \phi(p(x))\,dx \\ \text{Subject to} & p \in L,\ \mathbb{B}_0'p = \boldsymbol{\eta}'. \end{cases} \tag{13}$$

Observe that, for any consistent price density p,

$$\int_0^\infty x\,p(x)\,dx \le \int_0^{k_m} k_m\,p(x)\,dx + \int_{k_m}^\infty (k_m + c_m(x))\,p(x)\,dx = k_m + d_m < \infty.$$

Hence the value and solution of any primed problem will be the same as if it were optimized over all of L^1, subject to the constraints.

Further modifications related to the two previous theorems are

- The condition in Equation (7) is replaced with $\sum_{i=1}^m h_i'(x) = 1$ for all $x \in [0,\infty[$.
- The last sentence in Lemma 3.1(b) should be replaced with "Moreover, $\sum_{i=1}^m \eta_i' = 1$." Note it is now possible to have $\eta_m' = 1$ (if and only if $\text{supp}(p)$ is contained in $[k_m, \infty[$). The modification to the proof is obvious.
- Theorems 3.1 and 3.2 still hold, except references to $\mathbf{B}$ are replaced with $\mathbf{B}'$, and I_0 is possibly unbounded (as in Definition 3.2 above).
- Lemma 3.2 refers to Definition 3.2 (instead of Definition 3.1).
- Lemma 3.3 is unchanged (except that the meaning of $\boldsymbol{\eta}$ has changed).

Remark 3.8 The space L introduced by Equation (12) is an example of a *Köthe Space*; see, for example, Dieudonne ([12], 1951). The significance of this will become apparent when we investigate duality for the unbounded case in Section 4.3

Remark 3.9 (A possible technical simplification) Mathematically, one could deal only with the unbounded domain, and obtain the finite K case through domain reduction. One simply augments the data $\{k_i; d_i\}_{i=1}^m$ with $(k_{m+1}, d_{m+1}) = (K, 0)$ and treats the domain as $[0, \infty[$. The resulting matrix $\mathbf{B}'$ has dimension $(m+2) \times (m+2)$, and the $(m+2)$th moment condition is

$$\eta'_{m+2} = \int_K^\infty (x - K)\, p(x)\, dx = 0,$$

so that $[k_{m+1}, \infty[= [K, \infty[$ is excluded from I_0, guaranteeing finite support of the optimal p. This unified approach makes comparison with the results of Borwein *et. al.* ([7], 2003) less transparent, so we have not followed it.

4 Duality and Satisfaction of the Constraint Qualification

4.1 Duality for (P_0) when the Reduced Domain I_0 is Bounded

We proceed directly using the theory of convex duality.

First, put $\Phi(p) = \int_{I_0} \phi(p(x))\, dx$.

Standing assumptions throughout this section: Assume $\{k_i; d_i\}$ are such that $k_i < k_{i+1} < K$ for $i = 1, \ldots, m-1$, $(\eta_i)_{i=1}^{m+1} = \boldsymbol{\eta} = \mathbf{B}\,(1, d_1, \ldots, d_m)^T$ (where $\mathbf{B}$ is defined by (6)), that $\eta_i \geq 0$ for each i and condition $(C^{(\prime)})$ holds. Also, $\mathbb{B}$ and $\mathbb{B}_0$ are as defined by (9) and (10). □

For each $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$ define

$$\mathbb{B}_0{}^T \boldsymbol{\lambda} := \mathbf{1}_{I_0} \sum_{i=1}^{m+1} \lambda_i\, h_i. \tag{14}$$

Notice that for any $p \in L^1(I_0)$ and $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$,

$$\int_0^K [\mathbb{B}_0{}^T \boldsymbol{\lambda}](x)\, p(x)\, dx = \sum_{i=1}^{m+1} \lambda_i \int_0^K \mathbf{1}_{I_0} h_i(x)\, p(x)\, dx = \langle \boldsymbol{\lambda}, (\mathbb{B}_0 p) \rangle. \tag{15}$$

Thus, $\mathbb{B}_0{}^T : \mathbb{R}^{m+1} \to L^\infty(I_0)$ is the dual action of $\mathbb{B}_0 : L^1(I_0) \to \mathbb{R}^{m+1}$.

The Dual Functional Q and Feasibility of (P_0)

Assume initially that (P_0) is feasible. The Lagrangian for the primal problem (P_0) is

$$\begin{aligned} L(p, \boldsymbol{\lambda}) &= \int_{I_0} \phi(p(x))\, dx + \sum_{i=1}^{m+1} \lambda_i (\eta_i - \int_{I_0} h_i(x) p(x)\, dx) \\ &= \Phi(p) + \langle \boldsymbol{\lambda}, \boldsymbol{\eta} - \mathbb{B}_0 p \rangle \end{aligned} \tag{16}$$

where $p \in L^1(I_0)$ and $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$. Set

$$\begin{aligned} Q(\boldsymbol{\lambda}) &:= \inf_p L(p, \boldsymbol{\lambda}) \\ &= \langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle - \sup_p \left\{ \int_{I_0} [\mathbb{B}_0{}^T \boldsymbol{\lambda}](x)\, p(x)\, dx - \Phi(p) \right\} \\ &= \langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle - \Phi^*(\mathbb{B}_0{}^T \boldsymbol{\lambda}) \end{aligned} \tag{17}$$

where Φ^* is the Fenchel conjugate of the convex functional Φ; that is

$$\Phi^*(g) := \sup_{f \in L^1(I_0)} \left\{ \int_{I_0} f(x)\, g(x)\, dx - \Phi(f) \right\}.$$

When (P_0) is feasible, the principle of weak duality holds: for every $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$ and feasible $p \in L^1(I_0)$ we have $Q(\boldsymbol{\lambda}) \leq \Phi(p)$ with equality (strong duality) if and only if p is optimal in the primal problem (P_0).

Lemma 4.1 (Rockafellar [14, Corollary to Theorem 2]) *Let $\phi : \mathbb{R} \to \mathbb{R}$ be proper and convex and let $\Phi : L^1(I_0) \to \mathbb{R} \cup \{\pm\infty\}$ be defined by*

$$\Phi(f) := \int_{I_0} \phi(f(x))\, dx.$$

Then Φ is convex and its Fenchel conjugate on L^∞ is computed as

$$\Phi^*(g) = \int_{I_0} \phi^*(g(x))\, dx \qquad \forall g \in L^\infty(I_0).$$

An easy calculation shows that $\phi^*(y) = \exp(y - 1)$ and hence

$$Q(\boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle - \int_{I_0} \exp(\mathbb{B}_0{}^T \boldsymbol{\lambda} - 1)\, dx, \quad \forall \boldsymbol{\lambda} \in \mathbb{R}^{m+1}. \tag{18}$$

In particular we observe that Q is real-valued, concave and Gâteaux differentiable at every point $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$ with derivative

$$(\partial_i Q)(\boldsymbol{\lambda}) = \eta_i - \int_{I_0} h_i(x) \exp\left(\mathbb{B}_0{}^T \boldsymbol{\lambda} - 1\right)\, dx. \tag{19}$$

Note that $\partial_i Q(\boldsymbol{\lambda}) = 0$ for $1 \leq i \leq m+1$ if and only if $\mathbb{B}_0 \left(\exp\left(\mathbb{B}_0{}^T \boldsymbol{\lambda} - 1\right)\right) = \boldsymbol{\eta}$.

Example 4.1 (Example 2.1 revisited) As previously, let $0 = k_1 < k_2 < K$ and $d_1 = d_2 + k_2 \in]k_2, K[$. Then integrals in $\mathbb{B}_0$ are defined over $[k_2, K[$, $\boldsymbol{\eta} = \frac{1}{K - k_2}(0, K - d_1, d_1 - k_2)^T$ and hence

$$Q(\boldsymbol{\lambda}) = \tfrac{\lambda_2\,(K - d_1) + \lambda_3\,(d_1 - k_2)}{K - k_2} - \int_{k_2}^{K} \exp\left\{ \lambda_2 - 1 + (\lambda_3 - \lambda_2)\, \tfrac{x - k_2}{K - k_2} \right\} dx.$$

Clearly, $\ker \mathbb{B}_0^T = \{(\lambda_1, 0, 0) \ : \ \lambda_1 \in \mathbb{R}\}$, and Q can be maximized by solving $\nabla Q = 0$ (since it is a concave function).

Lemma 4.2 *Suppose that* $\bar{\boldsymbol{\lambda}}$ *is a local maximizer for the functional* Q *on* $\mathbb{R}^{m+1}$ *defined by* (18)*. Set* $\bar{p} := \exp({\mathbb{B}_0}^T\bar{\boldsymbol{\lambda}} - 1)$*. Then*

(a) $\bar{p} \in L^1(I_0)$ *and* $\mathbb{B}_0\bar{p} = \boldsymbol{\eta}$*.*
(b) $\bar{p}(x) = \{\int_{I_0} \exp(\sum_{i=1}^{m+1}(\bar{\boldsymbol{\lambda}})_i h_i(y))\,dy\}^{-1} \exp(\sum_{i=1}^{m+1}(\bar{\boldsymbol{\lambda}})_i h_i(x))$*.*
(c) $\Phi(\bar{p}) = \langle \bar{\boldsymbol{\lambda}}, \boldsymbol{\eta}\rangle - 1 < \infty$*.*
(d) $Q(\bar{\boldsymbol{\lambda}}) = \Phi(\bar{p})$*. In particular, strong duality holds and* $p = \bar{p}$ *is the unique optimal solution of the primal optimization problem* (P_0)*.*

Proof (a) Since Q is differentiable and $\bar{\boldsymbol{\lambda}}$ is finite, $\partial_i Q(\bar{\boldsymbol{\lambda}}) = 0$ for each i. Hence the function $\bar{p}$ is bounded and satisfies $\mathbb{B}_0\,\bar{p} = \boldsymbol{\eta}$.
(b) By Lemma 3.2, $\mathbb{B}\bar{p} = \boldsymbol{\eta}$. Hence $\mathbb{A}\bar{p} = (1, d_1, \ldots, d_m)^T$. In particular,

$$\int_0^K \bar{p}(x)\,dx = 1, \qquad \text{so} \qquad \bar{p} = \frac{\bar{p}}{\int_0^K \bar{p}(x)\,dx}.$$

(c) Using the formula for Φ,

$$\Phi(\bar{p}) = \int_0^K \bar{p}(x)\,[{\mathbb{B}_0}^T\bar{\boldsymbol{\lambda}} - 1](x)\,dx = \int_0^K \bar{p}(x)\,[{\mathbb{B}_0}^T\bar{\boldsymbol{\lambda}}](x)\,dx - \int_0^K \bar{p}(x)\,dx$$
$$= \langle \mathbb{B}_0\bar{p}, \bar{\boldsymbol{\lambda}}\rangle - 1$$

by Equation (15).
(d) Since $\int_0^K \bar{p}(x) = 1$, the formula (18) for Q gives $Q(\bar{\boldsymbol{\lambda}}) = \langle \boldsymbol{\eta}, \bar{\boldsymbol{\lambda}}\rangle - 1 = \Phi(\bar{p})$ as claimed. Since $\bar{p}$ is feasible for (P_0), the principle of weak duality holds, and

$$\Phi(\bar{p}) \geq \inf_{\{\mathbb{B}_0 p = \boldsymbol{\eta}\}} \Phi(p) \geq \sup_{\{\boldsymbol{\lambda}\}} Q(\boldsymbol{\lambda}) \geq Q(\bar{\boldsymbol{\lambda}}) = \Phi(\bar{p})$$

showing that $\bar{p}$ is optimal; the fact that Φ is strictly convex implies the required uniqueness.

It remains to show that Q must attain a local maximum. In fact we will show that Q attains a global maximum, although the presence of a kernel for ${\mathbb{B}_0}^T$ cannot be excluded (as illustrated by previous numerical example); hence, the extrema may not be unique. This adds a slight complication to the usual duality computations.

Lemma 4.3 *Assume* $\{k_i; d_i\}$ *satisfy the standing assumptions, that* ${\mathbb{B}_0}^T$ *is defined by* (14) *and let* Q *be given by* (18)*. Then*

$$\mathbb{R}^{m+1} = \operatorname{Ker}{\mathbb{B}_0}^T \oplus \operatorname{Range}(\mathbb{B}_0) \qquad \textit{and} \qquad \boldsymbol{\eta} \in \operatorname{Range}(\mathbb{B}_0).$$

Furthermore, Q *is constant along hyperplanes parallel to* $\operatorname{Ker}{\mathbb{B}_0}^T$*.*

Proof By Equation (15),

$$\mathbf{y} \in (\mathrm{Range}\,(\mathbb{B}_0))^{\perp} \qquad \Leftrightarrow \qquad \int_0^K [{\mathbb{B}_0}^T \mathbf{y}](x)\, p(x)\, dx = 0 \text{ for every } p \in L^1(I_0);$$

this occurs if and only if $[{\mathbb{B}_0}^T \mathbf{y}] = 0$ almost everywhere, so the decomposition is established. Next, let $\boldsymbol{\lambda} \in \mathrm{Ker}{\mathbb{B}_0}^T$ so $\varphi := {\mathbb{B}_0}^T \boldsymbol{\lambda} = 0$. By Lemma 3.3(b), $\eta_i \lambda_i \geq 0$ and $\eta_i \lambda_i \leq 0$ for $1 \leq 1 \leq m+1$. Hence $\boldsymbol{\lambda}^T \boldsymbol{\eta} = 0$; that is,

$$\boldsymbol{\eta} \in \left(\mathrm{Ker}{\mathbb{B}_0}^T\right)^{\perp} = \mathrm{Range}(\mathbb{B}_0).$$

The last part follows from the formula for Q.

Remark 4.1 Note that an effect of reducing the domain from I to I_0 is to decrease the range of $\mathbb{B}$ to $\mathrm{Range}\{\mathbb{B}_0\}$ (while retaining $\boldsymbol{\eta}$ in the range); a corresponding enlargement of $\mathrm{Ker}{\mathbb{B}_0}^T$ thus occurs.

Since hyperplanes parallel to $\mathrm{Ker}\,\mathbb{B}_0^T$ are level sets for Q, finding a maximum for Q is equivalent to finding a maximum for Q restricted to the subspace $\mathrm{Range}\{\mathbb{B}_0\}$. Indeed, once the maximizing $\bar{\boldsymbol{\lambda}}$ has been found, the maximum value for Q on $\mathbb{R}^{m+1}$ will occur at all points along the hyperplane $\bar{\boldsymbol{\lambda}} + \mathrm{Ker}{\mathbb{B}_0}^T$.

Lemma 4.4 *Suppose* $0 \neq \boldsymbol{\lambda} \in \mathrm{Range}\{\mathbb{B}_0\}$ *and* $\boldsymbol{\lambda} \perp \boldsymbol{\eta}$. *Then* $[{\mathbb{B}_0}^T \boldsymbol{\lambda}]^+ \neq 0$ *and* $[{\mathbb{B}_0}^T \boldsymbol{\lambda}]^- \neq 0$.

Proof By linearity, it is enough to prove the second inequality. Suppose that $[{\mathbb{B}_0}^T \boldsymbol{\lambda}]^- = 0$. Then ${\mathbb{B}_0}^T \boldsymbol{\lambda} \geq 0$ so that ${\mathbb{B}_0}^T \boldsymbol{\lambda} = 0$ by Lemma 3.3(c). Thus $\boldsymbol{\lambda} \in \mathrm{Ker}{\mathbb{B}_0}^T \cap \mathrm{Range}\{\mathbb{B}_0\}$, a contradiction to Lemma 4.3.

Lemma 4.5 *The functional* Q*, restricted to* $\mathrm{Range}\{\mathbb{B}_0\}$ *is coercive. That is,* $\lim_{r\to\infty} Q(r\,\boldsymbol{\lambda}) = -\infty$ *if* $0 \neq \boldsymbol{\lambda} \in \mathrm{Range}\{\mathbb{B}_0\}$.

Proof Let $0 \neq \boldsymbol{\lambda} \in \mathrm{Range}\{\mathbb{B}_0\}$. If $\boldsymbol{\lambda}^T \boldsymbol{\eta} < 0$ then

$$Q(r\,\boldsymbol{\lambda}) = r\,\boldsymbol{\lambda}^T\,\boldsymbol{\eta} - \Phi^*(r\,\mathbb{B}_0^T \boldsymbol{\lambda}) \leq r\,\boldsymbol{\lambda}^T\,\boldsymbol{\eta} \to -\infty$$

as $r \to \infty$. Otherwise, decompose

$$\boldsymbol{\lambda} = \alpha\,\boldsymbol{\eta} + \boldsymbol{\lambda}_0$$

where $\alpha \geq 0$ and $\boldsymbol{\lambda}_0 \perp \boldsymbol{\eta}$ (recall that $\boldsymbol{\eta} \in \mathrm{Range}\{\mathbb{B}_0\}$ by Lemma 4.3). Suppose that $\boldsymbol{\lambda}_0 \neq 0$. By Lemma 4.4, $[{\mathbb{B}_0}^T \boldsymbol{\lambda}_0]^+ \neq 0$ so there is a constant $\delta_0 > 0$ and a measurable set $J \subset I_0$ with $m(J) > 0$ such that

$${\mathbb{B}_0}^T \boldsymbol{\lambda}_0 \geq \delta_0\,\mathbf{1}_J.$$

Since $\boldsymbol{\eta}$ has non-negative components, when $r > 0$,

$$\mathbb{B}_0{}^T(r\,\alpha\,\boldsymbol{\eta})(x) = r\,\alpha \sum_{i=1}^{m+1} \mathbf{1}_{I_0}(x)\,\eta_i\,h_i(x) \geq 0$$

so that

$$\mathbb{B}_0{}^T(r\,(\alpha\,\boldsymbol{\eta} + \boldsymbol{\lambda}_0)) \geq 0 + r\,\delta_0\,\mathbf{1}_J.$$

Thus,

$$\Phi^*(\mathbb{B}_0{}^T(r\,\boldsymbol{\lambda})) = \int_{I_0} \exp\left(\mathbb{B}_0{}^T(r\,\boldsymbol{\lambda})(x) - 1\right)\,dx \geq \int_J e^{r\,\delta_0 - 1}\,dx = e^{r\,\delta_0-1}\,m(J).$$

Thus,

$$Q(r\,\boldsymbol{\lambda}) = r\,\boldsymbol{\lambda}^T\,\boldsymbol{\eta} - \Phi^*(\mathbb{B}_0{}^T(r\,\boldsymbol{\lambda})) \leq r\,\alpha\,\boldsymbol{\eta}^T\boldsymbol{\eta} - e^{r\,\delta_0-1}\,m(J) \to -\infty$$

as $r \to \infty$. The remaining case is when $\boldsymbol{\lambda}_0 = 0$ and $\alpha > 0$. Let $[k_i, k_{i+1}] \subset I_0$. Put $J = [k_i, k_{i+1}]$ and $\delta_0 = \min\{\eta_i, \eta_{i+1}\}$. Then $[\mathbb{B}_0{}^T\boldsymbol{\eta}] \geq \delta_0\,\mathbf{1}_J$ and the same argument as above completes the proof.

Proof of Theorem 3.2 when I_0 is bounded: The conditions in the theorem are the standing assumptions of this section. By Lemma 4.5, $Q|_{\mathrm{Range}(\mathbb{B}_0)}$ is coercive. Since Q is also continuous and concave, it attains a global maximum at some point $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^{m+1}$. By Lemma 4.2, (P_0) is feasible and $\bar{p}$ solves (P_0). By Lemma 3.2, (P) is feasible and is solved by $\bar{p}$. □

4.2 A Topological Formulation of the Constraint Qualification

It is interesting to note that the direct calculations of the previous section can be subsumed in general arguments based on a suitable topological constraint qualification. Having made the reduction to (P_0) (using Lemma 3.2), the essential step is to show that if

$$\eta_i \geq 0 \text{ for } i = 1, \ldots, (m+1) \qquad \text{and} \qquad (C) \text{ (or } (C')) \text{ is satisfied}$$

then $\mathbb{B}_0 p = \boldsymbol{\eta}$ for at least one p with $\int_0^K \phi(p(x))\,dx < \infty$. Let

$$\mathcal{B} = \left\{\mathbb{B}_0 p \;:\; p \in L^1(I_0) \text{ and } \int_0^K \phi(p(x))\,dx < \infty\right\}.$$

By arguments similar to Borwein *et. al.* ([7], 2003), (P_0) is feasible if the following holds:

$$(CQ) \qquad\qquad \boldsymbol{\eta} \in \mathrm{ri}\mathcal{B}.$$

The "standing assumptions" of the previous section are needed to define I_0 (and therefore formulate (P_0) and (CQ)). It turns out that (CQ) is automatically satisfied.

Proof that (CQ) *is satisfied:* Let $\mathbf{H}$ be the $(m+1)\times(m+1)$ matrix whose entries are 0 except for $H_{ii}=1$ when $\eta_i>0$. Clearly, $\mathbf{H}$ is diagonal and $\mathbf{H}\boldsymbol{\eta}=\boldsymbol{\eta}$.

Since $\int_0^K [\mathbb{B}_0{}^T\boldsymbol{\lambda}](x)\,p(x)\,dx = \langle \boldsymbol{\lambda}, (\mathbb{B}_0 p)\rangle$, it follows that $\langle \boldsymbol{\lambda}, \mathbf{y}\rangle \geq 0$ for all $\mathbf{y}\in\mathcal{B}$ if and only if $(\mathbb{B}_0{}^T\boldsymbol{\lambda})(x)\geq 0$. By Lemma 3.3(b) and the definition of $\mathbf{H}$, this latter condition occurs if and only if $(\mathbf{H}\boldsymbol{\lambda})_i \geq 0$ for all $i=1,\ldots,(m+1)$. The same argument as in Borwein *et. al.* ([7, Proposition 2], 2003) gives

$$\mathrm{ri}\mathcal{B} = \mathbf{H}\{]0,\infty[^{m+1}\} = \left\{\mathbf{x}\in\mathbb{R}^{m+1} \;:\; x_i>0 \text{ if } \eta_i>0, x_i=0 \text{ if } \eta_i=0\right\}.$$

It follows immediately that (CQ) is satisfied. □

4.3 Constraint Qualification for an Unbounded Domain I_0

The primal problem to be solved is now (P_0') from Equation (13) in Section 3. The existence of a primal-dual optimal risk neutral density for unbounded price intervals $(K=\infty)$ is established in Theorem 4.1. The finiteness of the problem (P_0') is not *a priori* obvious, and is a consequence of direct arguments (below) which establish the existence of a unique optimizer for $Q|_{\mathrm{Range}(\mathbb{B}_0')}$.

Naturally, we must begin with $K=\infty$. In the derivation of necessary conditions $(C^{(\prime)})$ one follows the arguments in the last part of Section 3 for the identification of $I_0\subseteq[0,\infty[$ (possibly removing $]k_m,\infty[$ if $\eta_{m+1}=0$). If the reduced domain I_0 is bounded, then the arguments of the previous section apply. We therefore consider only the case where $\eta_{m+1}>0$. Note in particular that Theorem 3.1 forces $\eta_m>0$, so $[k_{m-1},\infty[\subset I_0$.

Recall the reduced functional domain defined by Equation (12) in Section 3 when $K=\infty$. We reformulate this in the spirit of Köthe space as:

$$\Gamma := \{\mathbf{1}_{[0,\infty[},\ x\},$$
$$L := \left\{p \text{ measurable on } [0,\infty[\;:\; v\cdot p\in L^1[0,\infty[\text{ for all } v\in\Gamma\right\}.$$

The dual space is now naturally defined as

$$L^* := \left\{u \text{ measurable on } [0,\infty[\;:\; v\cdot u\in L^1[0,\infty[\text{ for all } v\in L\right\} \tag{20}$$

with a bilinear functional forming the pairing between L and L^*:

$$\langle f,g\rangle := \int_{I_0} f(x)g(x)\,dx, \quad f\in L,\ g\in L^*.$$

To summarize, L and L^* are **paired Köthe spaces** in the sense of Dieudonne ([12], 1951) or Maréchal ([13], 2001). Unfortunately L is not decomposable (meaning, it fails to be closed under the action $u\to \mathbf{1}_T g + (\mathbf{1}-\mathbf{1}_T)u$ where $g\in L^\infty$ and T is a measurable set of finite measure). Consequently, the key step in Section 4 that allowed *conjugation through the integral*, namely

Lemma 4.1, is not applicable. However, for the specific case of the entropy objective under consideration, the conjugation step is justified by the following result:

Proposition 4.1 (Maréchal [13, Proposition 2], 2001) *Suppose L and L^* are paired Köthe spaces as above. If the function $u^* \in L^*$ is such that the function* $\exp(u^*(x) - 1)$ *belongs to L, then $\Phi^*(u^*) = \int \phi^*(u^*(x))\,dx$.*

We now recommence the derivation from Section 4, in particular Equation (17) for the the dual functional Q in this setting:

$$\begin{aligned} Q(\boldsymbol{\lambda}) &= \langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle - \Phi^*({\mathbb{B}_0'}^T \boldsymbol{\lambda}), \\ \Phi^*(u^*) &:= \sup_{u \in L} \left\{ \int_{I_0} u(x)\, u^*(x)\, dx - \Phi(u) \right\}. \end{aligned}$$

The dual operator ${\mathbb{B}_0'}^T : \mathbb{R}^{m+1} \to L^*$ is derived as in Equation (15) but now on the unbounded domain I_0:

$$\int_0^\infty [{\mathbb{B}_0'}^T \boldsymbol{\lambda}](x)\, u(x)\, dx = \sum_{i=1}^{m+1} \boldsymbol{\lambda}_i \int_0^\infty \mathbf{1}_{I_0} h_i(x)\, u(x)\, dx = \langle \boldsymbol{\lambda}, (\mathbb{B}_0' u) \rangle$$

for all $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$ and $u \in L$.

In order to simplify notation, we henceforth omit the prime from the operator $\mathbb{B}_0'$ (but keep the subscript, indicating reduced domain). Thus

$${\mathbb{B}_0}^T : \mathbb{R}^{m+1} \to L^*.$$

Note that when $\lambda_{m+1} > 0$,

$${\mathbb{B}_0}^T \boldsymbol{\lambda} = {\mathbb{B}_0}^T (\lambda_1, \ldots, \lambda_m, 0)^T + \lambda_{m+1} \mathbf{1}_{I_0}\, c_m,$$

so ${\mathbb{B}_0}^T \boldsymbol{\lambda}(x)$ diverges linearly to infinity as $x \to \infty$ on I_0. In this case an elementary estimate shows that $\Phi^*\left({\mathbb{B}_0}^T \boldsymbol{\lambda}\right) = \infty$. If $\lambda_{m+1} = 0$ then ${\mathbb{B}_0}^T \boldsymbol{\lambda}$ is compactly supported. Then it is an easy matter to choose $u \in L$ (for example, also compactly supported, with support disjoint from ${\mathbb{B}_0}^T \boldsymbol{\lambda}$) leading to arbitrarily large positive values for the integral

$$\int_{I_0} {\mathbb{B}_0}^T \boldsymbol{\lambda} u(x) - \phi(u(x))\, dx = -\int_{I_0} \phi(u(x))\, dx.$$

Consequently, $Q(\boldsymbol{\lambda}) = -\infty$ on closed half-plane $\lambda_{m+1} \geq 0$, Q is proper and concave but only upper-semicontinuous and not differentiable on all of $\mathbb{R}^{m+1}$.

On the other hand, on $\mathbb{R}^{m+1}_{\ominus} := \{\lambda_{m+1} < 0\}$, ${\mathbb{B}_0}^T \boldsymbol{\lambda}(x)$ diverges linearly to minus infinity as $x \to \infty$ on I_0. It follows that $\exp({\mathbb{B}_0}^T \boldsymbol{\lambda}(x) - 1)$ is an element of the Köthe space L and Proposition 4.1 shows that for all such λ

$$Q(\boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle - \int_{I_0} \exp({\mathbb{B}_0}^T \boldsymbol{\lambda} - 1)\, dx.$$

Now, using the same arguments as in the bounded domain case, on $\mathbb{R}^{m+1}_{\ominus}$, Q is seen to be real-valued concave and Gâteaux differentiable. The analogue of formula (19) holds for $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}_{\ominus}$ and appropriate versions of Lemmas 4.2–4.4 hold with at most minor modifications to statements and proofs (replacing $L^1 - L^\infty$ with $L - L^*$ and using primed objects where appropriate, for example).

Theorem 4.1 *In the notation so far, under condition* $(C^{(\prime)})$, *the problem* (P') *with* $[0, K[= [0, \infty[$ *is equivalent to finding*

$$\inf_{\{p \in L \,:\, \mathbb{A}' p = \mathbf{d}\}} \Phi(p) = \sup_{\boldsymbol{\lambda} \in \mathbb{R}^{m+1}} \langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle - \int_{I_0} \exp\left([{\mathbb{B}_0}^T \boldsymbol{\lambda}](x) - 1 \right) dx.$$

The dual (right-hand) problem has a unique solution $\bar{\boldsymbol{\lambda}}$ *when restricted to* $\mathrm{Range}(\mathbb{B}_0) \cap \mathbb{R}^{m+1}_{\ominus}$. *Moreover,* $\bar{\boldsymbol{\lambda}}$ *satisfies* $\partial_i Q(\bar{\boldsymbol{\lambda}}) = 0$ *for each* $i = 1, \ldots, m+1$ *(see Equation (19)) and the unique solution to* (P') *is* $p = \mathbf{1}_{I_0} \exp({\mathbb{B}_0}^T \bar{\boldsymbol{\lambda}} - 1)$.

Proof The crucial step is to establish a version of Lemma 4.5 for the case of unbounded I_0 and $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$. Since $Q(\boldsymbol{\lambda})$ is upper-semicontinuous and bounded above by the linear function $\langle \boldsymbol{\lambda}, \boldsymbol{\eta} \rangle$, it suffices to establish that Q restricted to $\mathrm{Range}(\mathbb{B}_0)$ has a bounded and non-empty upper level set. By upper-semicontinuity, such a set will be compact, so that $\sup Q|_{\mathrm{Range}(\mathbb{B}_0)}$ is attained at a finite $\bar{\boldsymbol{\lambda}}$. Since $Q(\bar{\boldsymbol{\lambda}}) > -\infty$, $\lambda_{m+1} < 0$ so Q is differentiable at $\bar{\boldsymbol{\lambda}}$ and satisfies $\nabla Q(\bar{\boldsymbol{\lambda}}) = \mathbf{0}$.

Claim: For $\mathbf{0} \neq \boldsymbol{\lambda} \in \mathrm{Range}(\mathbb{B}_0)$, $\lim_{r \to \infty} Q((0, \ldots, 0, -2)^T + r\,\boldsymbol{\lambda}) = -\infty$.

Proof of theorem given the claim: Note that since I_0 is unbounded, $\eta_{m+1} > 0$. By Lemma 3.3(b), if $\boldsymbol{\lambda} \in \mathrm{Ker}{\mathbb{B}_0}^T$ then $\lambda_i\, \eta_i = 0$ for each i. Hence

$$(0, \ldots, 0, -2)\boldsymbol{\lambda} = 0,$$

so $(0, \ldots, 0, -2)^T \in (\mathrm{Ker}{\mathbb{B}_0}^T)^\perp = \mathrm{Range}(\mathbb{B}_0)$. Let $Q_0 = Q((0, \ldots, -2)^T)$. By the claim, for each $\hat{\boldsymbol{\lambda}} \in \mathrm{Range}(\mathbb{B}_0) \cap \{|\hat{\boldsymbol{\lambda}}| = 1\}$ there is an $R > 0$ such that $Q((0, \ldots, 0, -2)^T + R\hat{\boldsymbol{\lambda}}) < Q_0 - 1$. Since Q is upper-semicontinuous, there is an $\epsilon > 0$ such that if $|\hat{\boldsymbol{\lambda}}_1 - \hat{\boldsymbol{\lambda}}| < \epsilon$ and $|\hat{\boldsymbol{\lambda}}_1| = 1$ then

$$Q((0, \ldots, 0, -2)^T + R\,\hat{\boldsymbol{\lambda}}_1) < Q_0 - 1.$$

Since Q is concave, for each such $\hat{\boldsymbol{\lambda}}_1$, $Q((0, \ldots, 0, -2)^T + r\hat{\boldsymbol{\lambda}}_1) < Q_0 - 1$ for all $r > R$. The unit sphere in $\mathrm{Range}(\mathbb{B}_0)$ can be covered by finitely many such ϵ-neighbourhoods; thus there is an $R_{\max}$ such that

$$Q(\boldsymbol{\lambda}) \geq Q_0 - 1 \Rightarrow |\boldsymbol{\lambda} - (0, \ldots, 0, -2)^T| \leq R_{\max}.$$

The theorem now follows.

Proof of claim: There are several cases to consider. First, suppose that $\boldsymbol{\lambda}^T \boldsymbol{\eta} < 0$. Since $\Phi^*({\mathbb{B}_0}^T(\cdot)) \geq 0$,

$$Q((0, \ldots, 0, -2)^T + r\,\boldsymbol{\lambda}) \leq \langle (0, \ldots, 0, -2)^T + r\,\boldsymbol{\lambda}, \boldsymbol{\eta} \rangle = -2\,\eta_{m+1} + r\boldsymbol{\lambda}^T \boldsymbol{\eta} \to -\infty$$

as $r \to \infty$. Second, suppose that $\boldsymbol{\lambda}^T \boldsymbol{\eta} \geq 0$ and $\lambda_{m+1} > 0$. Then, by the remarks just before the statement of the theorem, whenever $r\lambda_{m+1} \geq 2$

$$Q((0, \dots, 0, -2)^T + r\boldsymbol{\lambda}) = -\infty.$$

Finally, suppose $\boldsymbol{\lambda}^T \boldsymbol{\eta} \geq 0$ and $\lambda_{m+1} \leq 0$. Similar to the proof of Lemma 4.5, decompose $\boldsymbol{\lambda} = \alpha\, \boldsymbol{\eta} + \boldsymbol{\lambda}_0$ where ${\boldsymbol{\lambda}_0}^T \boldsymbol{\eta} = 0$ and $\alpha \geq 0$. If $\boldsymbol{\lambda}_0 = \mathbf{0}$ then $\alpha > 0$ (since $\boldsymbol{\lambda} \neq 0$), so $0 \geq \lambda_{m+1} = (\alpha\, \boldsymbol{\eta} + \mathbf{0})_{m+1} = \alpha\, \eta_{m+1} > 0$. To avoid this contradiction, $\boldsymbol{\lambda}_0 \neq \mathbf{0}$, and by Lemma 4.4, $[{\mathbb{B}_0}^T \boldsymbol{\lambda}_0]^+ \neq 0$. In fact,

$$\mathbf{1}_{[0,k_m]}\, [{\mathbb{B}_0}^T \boldsymbol{\lambda}_0]^+ = [{\mathbb{B}_0}^T \boldsymbol{\lambda}_0]^+ \neq 0.$$

(To see this, note that $\mathbf{1}_{[0,k_m]}\, [{\mathbb{B}_0}^T \boldsymbol{\lambda}_0]^+ = [{\mathbb{B}_0}^T \boldsymbol{\lambda}_0]^+ - [(\boldsymbol{\lambda}_0)_{m+1}\, c_m]^+$. But $(\boldsymbol{\lambda}_0)_{m+1} = \lambda_{m+1} - \alpha\, \eta_{m+1} \leq 0$ so equality follows.) Hence, there is $J \subseteq [0, k_m]$ such that $m(J) > 0$ and $\mathbf{1}_J\, {\mathbb{B}_0}^T \boldsymbol{\lambda}_0 \geq \delta_0$, for a $\delta_0 > 0$. Then

$${\mathbb{B}_0}^T ((0, \dots, 0, -2)^T + r\, \boldsymbol{\lambda}) \geq r\, \delta_0\, \mathbf{1}_J,$$

and, since $(0, \dots, 0, -2)^T + r\, \boldsymbol{\lambda} \in \mathbb{R}^{m+1}_{\ominus}$,

$$Q((0, \dots, 0, -2)^T + r\, \boldsymbol{\lambda}) \leq \alpha\, r\, \boldsymbol{\eta}^T \boldsymbol{\eta} - e^{r\, \delta_0 - 1}\, m(J) \to -\infty \qquad \text{as } r \to \infty.$$

5 Conclusions

The maximum entropy principle (MEP) was proposed by Buchen and Kelly ([3], 1996) as a way of approximating a risk neutral probability measure for the price of a financial asset at a future time, inferred from market prices of simple options on that asset. Moment conditions are obtained from the no-arbitrage assumption that all options have been priced consistently. On the one hand, the absence of arbitrage opportunities implies that the options' strike-price curve must be convex; on the other hand, the usual process of solution of the MEP via Lagrangian duality relies on strict convexity of this curve. Our work elucidates the second observation, making clear the connection between existence of a risk-neutral measure, and the applicability of the MEP. In particular, we build on work of Borwein *et. al.* ([7], 2003), and show how to use the MEP when strict convexity of the data is lost.

In reality, market price data is not always convex, since it is prone to several potential sources of inconsistency: transcription errors, incorrect or underdetermined pricing models, mistakes by a trader, imbalances between market players in the amount of information available to be fed into the models. Any of these mechanisms can be regarded a source of noise on otherwise consistent data, leading to possible convexity violation and consequent arbitrage opportunities.

Recent interest in applying the MEP setup to real market data (see Rodriguez and Santosa ([11], 2012) and Neri and Schneider ([10], 2012)) has considered a range of strategies for regularising, projecting or otherwise perturbing non-convex data. Our results show that at the point of crossing between consistent to inconsistent data, the data *in fact remain consistent.* This insight is important for understanding the performance of various numerical strategies, and suggests that the idea of "convexifying" data via projection onto a feasible set is worthy of further investigation.

Acknowledgements The authors thank an anonymous referee who pointed out an error in our use of Köthe spaces and conjugation past the integral in an earlier version of this paper.

The first author acknowledges the hospitality of the Department of Mathematics and Statistics, University of Canterbury, New Zealand during sabbatical leave, where much of this work was carried out.

References

1. Breeden, M., Litzenberger, R.H.: Price of state-contingent claims implicit in option prices. J. of Business 51, 621–651 (1978)
2. Shimko, D.: Bounds of probability. Risk 6, 33–37 (1993)
3. Buchen, P.W., Kelley, M.: The maximum entropy distribution of an asset inferred from option prices. J. Financial and Qual. Anal. 31, 143-159 (1996)
4. Avellaneda, M., Friedman, C., Holmes, R. and Samperi, D.: Calibrating volatility surfaces via relative entropy minimization. Appl. Math. Finance 4, 37–64 (1997)
5. Avellaneda, M.: The minimum-entropy algorithm and related methods for calibrating asset-pricing models. In: Proceedings of the International Congress of Mathematicians, Vol III. Doc. Math., pp545-563 (1998)
6. Jaynes, E.: On the rationale of maximum-entropy methods. Proc. IEEE 70, 939–952 (1982)
7. Borwein, J., Choksi, R., Maréchal, P.: Probability distributions of assets inferred from option prices via the principle of maximum entropy. SIAM J. Optim. 14, 464–478 (2003).
8. Guo, D.: Maximum entropy in option pricing: a convex-spline smoothing method. J. Futures Markets 21, 819–832 (2001)
9. He, C., Coleman, T.F., Li, Y.: Computation and analysis for a constrained entropy optimization problem in finance. J. Comput. Appl. Math. 222, 159–174 (2008)
10. Neri, C., Schneider, L.: Maximum entropy distributions inferred from option portfolios on an asset. Finance Stoch. 16, 293–318 (2012)
11. Rodriguez, J., Santosa, F.: Estimation of asset distributions from option prices: Analysis and regularization. SIAM J. Financial Math. 3, 374–401 (2012)
12. Dieudonne, J.: Sur les espaces de Köthe. J. Anal. Math. 1, 81–115 (1951)
13. Maréchal, P.: A note on entropy optimization. In: Lassonde, M (ed.): Approximation, Optimization and Mathematical Economics, pp. 205–211. Physica-Verlag (2001)
14. Rockafellar, R.T.: Integrals which are convex functionals. *Pac. J. Math.* 24, 525–539 (1968)
15. Bose, C., Murray, R.: Duality and the computation of approximate invariant densities for nonsingular transformations, SIAM J. Opt. 18, 691–709 (2007)