An Axiomatic and Data Driven View on the EPK Paradox

Maria Grith, Wolfgang K. Härdle, Volker Krätschmer

Abstract

Supported by several recent investigations the empirical pricing kernel (EPK) paradox might be considered as a stylized fact. Some authors suggest that this paradox might be caused by regime switching in financial markets. Based on an economic model with state dependent utilities for the financial investors we want to emphasize a microeconomic view that succeeds in explaining the paradox via state dependent preferences. We shall also develop and investigate calibration problems in terms of data fits for basic values of the pricing kernel.

Keywords: Pricing kernel, representative agent, empirical pricing kernel, epk paradox, state dependent utilities, switching points. JEL Classification: D01, D53, C02, G13
AMS Classification: 15A29, 62G07, 62G35

1 Introduction

The empirical pricing kernel paradox emerged as an empirical phenomenon in the financial markets, that cannot be explained by means of traditional expected utility assumptions. Several recent studies support
the conjecture that this deviation evolves into a stylized fact. All the investigations are settled within
similar economic models assuming a representative agent in financial markets whose preferences have
classical expected utility representation. Additionally, the risk neutral valuation principle is supposed
to be valid for the financial markets by means of pricing kernels. If the pricing kernels represent state
contingent equilibrium prices they might be identified with the v. Neumann-Morgenstern utility indices
of the representative agent. As a consequence the pricing kernels should be nonincreasing.

In the studies of Ait-Sahalia and Lo (2000), Jackwerth (2000), Constantinidis et al. (2009), Detlefsen et al.
(2010), Beare (2011) different econometric methods have been applied to estimate pricing kernels with
varying underlying models for the financial markets. It turned out as a common result, that the estimates,
the so called empirical pricing kernels (EPK), have non-monotonic shape regardless of the used data sets.
Typically, we find a shape of the empirical pricing kernel as visualized in Figure 1 which is borrowed from
Grith et al. (2010). It shows empirical pricing kernels estimated at different maturities, left panel and
various dates, right panel. The underlying data are the DAX index returns and European options in 2006.
Significantly, all curves are quite similar: They have the same form, the same characteristic features like
e.g. the hump, and they differ in absolute terms slightly. In particular the empirical kernels fail to be
monotone, contrasting the classical theory within the expected utility framework. This is what we shall
call the EPK paradox. Further confirmations of it are provided by Golubev et al. (2011) who construct a
monotonicity test for the local concavity of the utility function and Härdle et al. (2010) who build uniform
confidence bands for the empirical pricing kernel. Typically, the null hypothesis that the pricing kernel is
nonincreasing was rejected.

The aim of this paper is to provide an extended economic model that allows non-monotone pricing kernels.
We retain the expected utility framework in a one period model and endow the financial investors with
preferences that might be state sensitive. In particularly, here, we define the state with respect to possible
realizations of the prices in the financial markets. More technically, investors switch between two utility
indexes - defined over consumption sets - at a point that we call ‘reference point’ that lies in the price
space. As a consequence, while the utility indexes are concave in consumption, it has a jump in the stock
index space. In equilibrium, this renders pricing kernel non-monotonic.

The paper is organized as follows. In section 2 we review the literature on the empirical pricing kernel
paradox. In section 3 we shall introduce our model for the financial market. The classical relationship between the utilities of representative agent and the pricing kernel will be reviewed. Afterwards we shall point out the empirical pricing kernel paradox. A simple consumption model based on state dependent utilities for the investors will be introduced in section 4. Within this framework we shall retain the relationship between preferences of investors and pricing in the market by an analogous result. In particular, our state dependent specification of preferences allows for non-monotone pricing kernels. Investigations into the switching behavior of the empirical pricing kernel involve solving inverse/calibration problems which will be developed and investigated throughout section 5. Section 6 concludes. Several mathematical results and proofs of quite technical nature have been delegated to appendices A, B, C, D respectively.

2 Literature Review

A large body of literature that investigates the mechanisms through which a locally increasing region in the pricing kernel - as a function of aggregated wealth - can be obtained have as a reference point deviations from the classical assumptions of the expected utility maximization assumptions. Hens and Reichlin (2011) conduct a systematic analysis of the EPK paradox by relaxing in turn the assumptions embedded in the standard models: complete markets, risk-averse investors and correct beliefs. They find

Figure 1: Examples of empirical pricing kernels for various maturities (left) and monthly pricing kernels for the first 6 months in 2006 for 1M maturity (right).
that incomplete markets can alone explain the puzzle; local risk-proclivity is non-robust at the aggregate level while erroneous beliefs seem to work only under certain scenarios. With both homogeneous and heterogeneous beliefs - see also Shefrin (2008), the degree of pessimism necessary to generate the bump is unreasonably high. This is in line with the finding of Ziegler (2007). However, adding distortions - a behavioral explanation originated with Kahneman and Tversky (1979) - to misestimated physical probabilities yield realistic results.

Another stream of literature rationalizes the EPK kernel by considering state dependency. In Brown and Jackwerth (2000) the pricing kernel bump can be reconstructed in models with a representative agent with state dependent utility. For example, they weigh two pricing kernels associated with two states (e.g. high-low volatility), whose likelihood of occurrence is linked to the wealth levels. However, their explanation is of very general nature and does not allow to clearly identifying the sources of the puzzle.

State dependence has been further used to explain the EPK puzzle in equilibrium models mainly based on two utility classes: habit formation, see Constantinides (1990), Campbell and Cochrane (1999), or recursive utilities, see Epstein and Zin (2001). In these papers, one typically assumes a Markov switching process for the evolution of states and derive asset related characteristics in a consumption based model. Garcia et al. (2003) investigate recursive utility functions with state dependency in the fundamentals. Melino and Yang (2003) disentangle the roles played by state dependent intertemporal substitution or time preference in explaining the risk aversion puzzle in a model with state dependent recursive preferences. Veronesi (2004) extends the state dependent utility by assuming that the agents possess a probability distribution over their state and introduces the concept of 'belief-dependent preferences'. Chabi-Yo et al. (2008) generalize the setup of Melino and Yang (2003) and investigate what type of conditioning (e.g. in preferences, economic fundamentals or beliefs) are more likely to explain the data.

In our approach we consider bivariate state dependent preferences defined with respect to the state variable, the stock price index under the assumptions that the agents in the market have true beliefs. The framework is suited in effect to homogeneous and heterogeneous agents framework with peculiarities in the resulting PK shape that reside in the implied assumptions.
Let \([0, T]\) be the time interval of investment in the financial market, where \(t = 0\) denotes the present time and \(t = T \in [0, \infty[\) the time of maturity.

It is assumed that a riskless bond and a risky asset are traded in the financial market as basic underlyings. The price process \((B_t)_{t \in [0, T]}\) of the riskless bond is defined by

\[
\frac{dB_t}{B_t} = r_t \ dt,
\]

via a deterministic Riemannian-integrable interest process \((r_t)_{t \in [0, T]}\). The price process \((S_t)_{t \in [0, T]}\) of the risky asset is taken to be a nonnegative semimartingale with constant \(S_0\) and continuously distributed marginals \(S_t\). Examples for such financial markets are the Black-Scholes model, non-parametric diffusion models as in Ait-Sahalia and Lo (2000) and GARCH models. Notice that also discrete time models may be subsumed under this setting.

Furthermore let us suppose that the financial market is arbitrage free in the sense that there exists an equivalent martingale measure. We further assume that the risk neutral valuation principle is valid for nonnegative payoffs \(\psi(S_T)\). Hence there is an unknown Radon-Nikodym density \(\pi\) of a martingale measure such that the price of any \(\psi(S_T)\) is characterized by

\[
\mathbb{E}\left[ D_T \psi(S_T) \pi \right] = \mathbb{E}\left[ D_T \psi(S_T) \mathbb{E}[\pi | S_T] \right].
\]

(1)

with the discount factor \(D_T = \exp(-\int_0^T r_x \ dx)\). By factorization we find some Borel-measurable \(\mathcal{K}_\pi\) with \(\mathbb{E}[\pi | S_T] = \mathcal{K}_\pi(S_T)\), so that

\[
\mathbb{E}\left[ D_T \psi(S_T) \pi \right] = \int_0^\infty D_T \psi(x) \mathcal{K}_\pi(x) \ p_{S_T}(x) \ dx,
\]

(2)

where \(p_{S_T}\) denotes a density function of the distribution of \(S_T\). Equation (2) gives reason to call \(\mathcal{K}_\pi\) the pricing kernel (w.r.t. \(\pi\)).

Let us now embed the financial market into an economic model where the investors of the financial market are consumers whose consumptions rely on the price \(S_T\) of the stock at maturity only. Within the classical framework, where the investor preferences may be represented by expected utilities, there...
exists a link between the risk attitude of the investors and the pricing rule of the financial markets. It is built upon the assumption of a representative agent whose indirect utility $U[\vec{e}(S_T)]$, depending on the aggregated market endowment $\vec{e}(S_T)$, has expected utility representation $U[\vec{e}(S_T)] = E[u(S_T)]$ with concave v. Neumann-Morgenstern utility index $u$. Under some further technical conditions on the investor preferences, there is some positive $\beta$ such that

$$\frac{du}{dx} \bigg|_{x=\vec{e}(s_T)} = \beta K_\pi(s_T)$$

for every realization $s_T$ of $S_T$. This relationship might be obtained by using methods as in the sections 6.1, 6.2 of Karatzas and Shreve (1998). For a rigorous formulation and derivation see Corollary A.2 in Appendix A. It should be emphasized that within the classical expected utility framework the pricing kernel has to be nonincreasing due to concavity of the utility index $u$.

4 Financial Investors within the Expected Utility Framework and the EPK Paradox

We shall provide a simple economic model where the pricing kernel need not to be nonincreasing. The key idea is to consider the investors as consumers whose preferences are representable by utilities dependent on the prices of the stock. An axiomatic justification for this concept of state dependent preferences is provided by Karni et al. (1983).

Let us assume that we have $m$ consumers who choose among contingent consumption plans dependent on the only state variables $S_T$, i.e. nonnegative consumption random variables $c(S_T)$. They have exogeneous endowments by initial capitals $w_0^1, ..., w_0^m > 0$ and state dependent endowment in form of nonnegative random variables $e_1(S_T), ..., e_m(S_T)$. Their individual budget constraints for $c(S_T)$ is therefore:

$$\int_0^\infty D_Tc(x)K_\pi(x)p_{S_T}(x) \, dx \leq w_0^i + \int_0^\infty D_Te_i(x)K_\pi(x)p_{S_T}(x) \, dx, \ i = 1, \ldots, m. \quad (3)$$

The consumers are assumed to have state dependent utilities in terms of extended expected utility preferences within the terminology of Mas-Colell et al. (1995): consumer $i$ has a representation of her/his preferences as:

$$U^i[c(S_T)] = E \left[ D_Tu^i[S_T, c(S_T)] \right],$$
where \( u^i : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{-\infty\} \) denotes a state dependent v. Neumann-Morgenstern utility index satisfying:

\[
\begin{align*}
  u^i(x, y) &\in \mathbb{R} \quad \text{for} \quad x \geq 0, \quad y > 0, \\
  u^i(x, \cdot) &\text{ is strictly increasing and strictly concave for any} \quad x \geq 0, \\
  u^i(\cdot, y) &\text{ is Borel-measurable for every} \quad y \geq 0.
\end{align*}
\]

At time of maturity \( T \) the market has an aggregated endowment \( \bar{e}(S_T) \) defined as \( \bar{e}(S_T) = \sum_{i=1}^m \{ w_i^0 + e_i(S_T) \} \). It is assumed that simultaneous consumption is allowed for consumption vectors \( (c_1(S_T), ..., c_m(S_T)) \) which satisfy the individual budget constraints and obey the aggregated endowment in the sense of \( \sum_{i=1}^m c_i(S_T) \leq \bar{e}(S_T) \). Such vectors are admissible, and they are gathered in set say \( A \) (Mas-Colell et al. (1995)). The consumers have chosen their consumptions \( (\bar{c}_1(S_T), ..., \bar{c}_m(S_T)) \) such that the following natural properties are fulfilled.

(i) **market clearing:**

\[
\sum_{i=1}^m \bar{c}_i(S_T) = \bar{e}(S_T).
\]

The conditions (8) and (8) describe a weak version of a contingent Arrow Debreu equilibrium (Dana and Jeanblanc (2003), sect. 7.1). As a by product \( (\bar{c}_1(S_T), ..., \bar{c}_m(S_T)) \) is a Pareto optimum too, i.e. there is no \( (c_1(S_T), ..., c_m(S_T)) \in A \) with \( U^i\{c_i(S_T)\} \geq U^i\{\bar{c}_i(S_T)\} \) for every \( i \) and such that \( U^{i_0}\{c_{i_0}(S_T)\} > U^{i_0}\{\bar{c}_{i_0}(S_T)\} \) for at least one \( i_0 \). Therefore, by the so called Negeishi method we find nonnegative weights \( \alpha_1, ..., \alpha_m \) such that

\[
\sum_{i=1}^m \alpha_iU^i\{\bar{c}(S_T)\} = \max \left\{ \sum_{i=1}^m \alpha_iU^i\{c(S_T)\} \mid \sum_{i=1}^m c_i(S_T) \leq \bar{e}(S_T) \right\}
\]

Let \( u_\alpha : \mathbb{R}_+^2 \rightarrow \mathbb{R} \cup \{-\infty, \infty\} \) be defined by

\[
u_\alpha(x, y) = \sup \left\{ D_T \sum_{i=1}^m \alpha_i u^i(x, y_i) \mid y_1, ..., y_m \geq 0, \sum_{i=1}^m y_i \leq y \right\}.
\]
Obviously, \( u_\alpha(x, \cdot) \) is strictly increasing as well as strictly concave for \( x \geq 0 \), and \( u_\alpha(\cdot, y) \) is Borel-measurable for every \( y \geq 0 \). This leads us to the extended expected utility representation

\[
U_\alpha\{\bar{e}(S_T)\} = E[D_T u_\alpha \{S_T, \bar{e}(S_T)\}],
\]

see Lemmata B.1, B.2 in Appendix B. In the next step we establish the relationship between the indirect utility \( u_\alpha \) of the representative agent and the pricing kernel \( K_\pi \). The usual Inada conditions on the state dependent utility indices \( u^1(S_T, \cdot), \ldots, u^m(S_T, \cdot) \), i.e. \( u^1(x, \cdot)[0, \infty[, \ldots, u^m(x, \cdot)][0, \infty[ \) are assumed to hold; \( u^i(x, \cdot) \) are continuously differentiable with

\[
\lim_{\varepsilon \to 0} \frac{du^i(x, \cdot)}{dy} \bigg|_{y = \varepsilon} = \infty, \quad \lim_{\varepsilon \to \infty} \frac{du^i(x, \cdot)}{dy} \bigg|_{y = \varepsilon} = 0 \quad (i = 1, \ldots, m)
\]

for \( x \geq 0 \).

The Inada conditions together with (5) imply that for any \( i \in \{1, \ldots, m\} \) and every \( x \geq 0 \) the mapping \( \frac{du^i(x, \cdot)}{dy} [0, \infty[ \) is injective onto \( ]0, \infty[ \) with continuously differentiable, strictly decreasing inverse say \( I_i(x, \cdot) \). In accordance with the consumption optimization for expected utility maximizing financial investors we assume as Dana and Jeanblanc (2003), Duffie (1996), Karatzas and Shreve (1998) that

\[
E[I_1 \{ (S_T, yK_\pi(S_T)) \}], \ldots, E[I_m \{ (S_T, yK_\pi(S_T)) \}] < \infty \quad \text{for any } y > 0.
\]

**Theorem 4.1** In addition to (4) – (11) let \( u^1(x, \cdot)[0, \infty[, \ldots, u^m(x, \cdot)][0, \infty[ \) be twice continuously differentiable for \( x \geq 0 \). Then \( u_\alpha(s_T, \cdot)[0, \infty[ \) is continuously differentiable for every realization \( s_T \) of \( S_T \).

Furthermore for any \( \alpha_i > 0 \) there exists some \( \beta_i > 0 \) such that

\[
\frac{du_\alpha(s_T, \cdot)}{dy} \bigg|_{y = \varepsilon(s_T)} = \alpha_i \frac{du^i(s_T, \cdot)}{dy} \bigg|_{y = \varepsilon_i(s_T)} = \alpha_i \beta_i K_{\pi}(s_T)
\]

for every realization \( s_T \).

The proof of Theorem 4.1 is delegated to the end of Appendix A.

**Remark:**

The relationship between pricing kernels and marginal utilities of the individual investors as stated in Theorem 4.1 occurs as a motif in the literature of financial mathematics to characterize solutions of different optimization problems. Concerning the optimal utility based investment Kramkov and Schachermayer
(1999) introduced a new condition on the asymptotic elasticity of the utilities which replaced the analogue of condition (11). Within our framework it reads as follows

$$\limsup_{y \to \infty} \frac{d u_i(x, \cdot)}{dy} \bigg|_{y < 1} \quad \text{for any } x \geq 0 \text{ and every } i \in \{1, \ldots, m\}. \quad (12)$$

Kramkov and Schachermayer restrict themselves to individual investors with classical state independent expected utility preferences. They achieved to show that the new introduced condition is a minimal requirement to describe the optimal investment in terms of the marginal utilities and a pricing kernel.

Their methods had been adapted by Karatzas and Zitkovic (2003) to characterize the optimal consumption in incomplete financial market similar to Theorem 4.1. The difference, and the mathematically more challenging point is, that in Karatzas and Zitkovic (2003) the martingale measure is not fixed in advance. Within our setting the guidelines of Kramkov and Schachermayer (1999) might be followed more directly to establish Theorem 4.1 with condition (12) instead of (11). However, a rigorous derivation would lie beyond the scope of this paper, and we use condition (12) which simplifies the argumentation in the proof of Theorem 4.1.

Let $R_T = \frac{S_T}{S_0}$ be the return at maturity. It has a continuous distribution, say $P_{R_T}$. If the market endowment specializes to $\bar{e}(S_T) = \frac{S_T}{S_0}$, Theorem 4.1 reads as follows.

**Corollary 4.2** Let $\bar{e}(S_T) = \frac{S_T}{S_0}$ and let $u^1(x, \cdot)[0, \infty], \ldots, u^m(x, \cdot)[0, \infty]$ be twice continuously differentiable for $x \geq 0$. Then under (4) – (11), $u_i(s_t, \cdot)[0, \infty]$ is continuously differentiable for every realization $r_T$ of $R_T$ and for any $\alpha_i > 0$ there exists some $\beta_i > 0$ such that

$$\frac{d u_i(r_T, \cdot)}{dy} \bigg|_{y = r_T} = \alpha_i \frac{d u^i(r_T, \cdot)}{dy} \bigg|_{y = \bar{e}_i(S_T)} = \alpha_i \beta_i K_{\pi}(S_0 r_T) = \tilde{K}_\pi(r_T).$$

Corollary 4.2 is the cornerstone for linking aggregated individual preferences to the market pricing kernel with its potential nonmonotonicities. The framework of state dependent utilities of the investors allows us to describe a switching behavior of them when facing a threshold for the price of the stock at maturity.

In more detail, let us assume that each consumer $i$ is disposed of two basic continuous, strictly increasing and strictly concave utility indices $u_i^0, u_i^1 : [0, \infty] \to \mathbb{R} \cup \{-\infty\}$ with $u_i^0(y), u_i^1(y) \in \mathbb{R}$ for $y > 0$. He or she is changing between these indices dependent on a threshold $x_i > 0$ for the return $R_T$, i.e.

$$u^i(r_T, \cdot) = 1_{[0, x_i]}(r_T) u_i^0 + 1_{[x_i, \infty]}(r_T) u_i^1$$

(13)
for every realization $r_T$ of $R_T$. Here $1_A$ denotes the indicator function of subset $A$. The reader might think of $u_i^0, u_i^1$ as utility indices representing bearish and bullish risk attitudes of consumer $i$, and that her or his revealed attitudes are adapted to the prices of the financial market.

In order to simplify notations, let us assume that the thresholds are ordered by $x_1 \leq \ldots \leq x_m$. There exist different competing potential representative agent groups in the market with respective representations $U^1_\alpha \{ \bar{e}(S_T) \}, \ldots, U^{m+1}_\alpha \{ \bar{e}(S_T) \}$ of indirect utilities defined by

$$U^i_\alpha \{ \bar{e}(S_T) \} = \sup \left\{ \sum_{k=1}^{i-1} \mathbb{E} \left[ D_T u_k^0 (c_k(S_T)) \right] + \sum_{k=i}^{m+1} \mathbb{E} \left[ D_T u_k^1 (c_k(S_T)) \right] \mid \sum_{k=1}^{m+1} c_k(S_T) \leq \bar{e}(S_T) \right\}.$$  

In view of Lemmata B.1, B.2 they have expected utility representations

$$U^i_\alpha \{ \bar{e}(S_T) \} = \mathbb{E} \left[ D_T u^i_\alpha \{ \bar{e}(s_T) \} \right],$$

where

$$u^i_\alpha (y) = \sup \left\{ D_T \left( \sum_{k=1}^{\alpha} \alpha_k u^0_k (y) 1_{k \geq i} + \sum_{k=1}^{m} \alpha_k u^1_k (y) 1_{k < i} \right) \mid y_1, \ldots, y_k \geq 0, \sum_{k=1}^{m} y_k \leq y \right\}$$

for $y \geq 0$, $i \in \{ 1, \ldots, m+1 \}$. It is now a routine exercise to verify that

$$u_\alpha (x, y) = 1_{[0, x_1]} (x) u^1_\alpha (y) + \sum_{i=1}^{m-1} 1_{[x_i, x_{i+1}]} (x) u^{i+1}_\alpha (y) + 1_{[x_m, \infty]} (x) u^{m+1}_\alpha (y) \text{ for } x, y \geq 0.$$  

As a consequence the indirect utility $U^i_\alpha \{ \bar{e}(S_T) \}$ might be interpreted as expressing the hegemony of the different potential representative agents. Moreover, via Corollary 4.2 we obtain for some $\beta > 0$ and any realisation $r_T$ of $R_T$

$$1_{[0, x_1]} (r_T) \frac{d u^1_\alpha (y)}{dy} \bigg|_{y=r_T} + \sum_{i=1}^{m-1} 1_{[x_i, x_{i+1}]} (r_T) \frac{d u^{i+1}_\alpha (y)}{dy} \bigg|_{y=r_T} + 1_{[x_m, \infty]} (r_T) \frac{d u^{m+1}_\alpha (y)}{dy} \bigg|_{y=r_T} = \tilde{\kappa}_\pi (r_T) \quad (14)$$

From this observation it becomes clear that the pricing kernel is nonincreasing separately on the intervals $[0, x_1], [x_1, x_2], \ldots, [x_m, \infty]$, but it might fail to be monotone just at the switching points $x_1, \ldots, x_m$.

**Example 1.** For illustration let us assume that the distribution of $R_T$ has $[0, \infty]$ as support, and that the investors have an identical switching point say $x_0$.

$$1_{[0, x_0]} (r_T) \frac{d u^1_\alpha (r_T, \cdot)}{dy} \bigg|_{y=r_T} + 1_{[x_0, \infty]} (r_T) \frac{d u^{m+1}_\alpha (r_T, \cdot)}{dy} \bigg|_{y=r_T} = \tilde{\kappa}_\pi (r_T)$$
for every realization $r_T$ of $R_T$.

Furthermore, let us suppose that each investor $i$ switches between CRRA utilities $u^j_i(y) = y^{\gamma_j^i}/\gamma_j^i$ ($j = 0, 1$) with $0 < \gamma_j^i < \gamma_j^0 < 1$, inducing the Arrow-Pratt coefficients of absolute risk aversion $\rho_j^i(y) = \frac{1-\gamma_j^i}{y}$ ($j = 0, 1; y > 0$). It follows $\rho_j^i(y) > \rho_j^0(y)$ for $i \in \{1, ..., m\}$ and $y > 0$, which means that $u^1_1, ..., u^1_m$ represent a more risk averse attitude than $u^0_1, ..., u^0_m$ (Mas-Colell et al. (1995), p. 191). In particular for stock returns lower or equal $x_0$ we have a bullish market, whereas we obtain a bearish market when stock returns exceed $x_0$.

The application of Lemma B.1 and Proposition B.3 in Appendix B yields

$$r_T = F^0\left(\frac{du^1_1(y)}{dy} \bigg|_{y=r_T}\right) = F^1\left(\frac{du^m_m(y)}{dy} \bigg|_{y=r_T}\right)$$

for any positive realization $r_T$, where

$$F^j : ]0, \infty[ \to ]0, \infty[, z \mapsto \sum_{i=1}^m \left( \frac{z}{\alpha_i} \right)^{\frac{1}{\gamma_j^i}} \frac{1}{\gamma_j^i - 1} \quad (j = 0, 1)$$

are decreasing bijective mappings. If $x_0$ is larger than the intersection of $F^0$ and $F^1$

$$F^0\left(\frac{du^1_0(y)}{dy} \bigg|_{y=x_0}\right) = x_0 = F^1\left(\frac{du^m_m(y)}{dy} \bigg|_{y=x_0}\right) > F^0\left(\frac{du^m_m(y)}{dy} \bigg|_{y=x_0}\right).$$
for any realization $r_T \geq x_0$. Therefore
\[
\frac{d u^{m+1}_\alpha(y)}{dy} \bigg|_{y=x_0} > \frac{d u^1_\alpha(y)}{dy} \bigg|_{y=x_0}
\]
That means that $\tilde{K}_\pi$ is not monotone at $x_0$. A graphical illustration for this example is in Figure 2 for $\gamma^0 = 0.75$ and $\gamma^1 = 0.5$.

---

**Figure 3:** The relationship between the shape of the pricing kernel and the weight function $w(r_T) = d\omega(r_T)/dr_T$ for linear, exponential, logistic, constant and bell shaped specifications for $\gamma^0 = \gamma^1 = 0.5$, $b_0 = 1$ and $b_1 = 1.2$

---

**Example 2.**

Next, we illustrate the case of heterogeneous reference points $x_i$ in 13 by assuming that all agents switch between the same two utility indexes $u^j_i(y) = u^j(y)$, $(j = 0, 1)$ for all $i = 1, \ldots, m$. Let us denote
\[
\omega(r_T) = 1 - \frac{1}{m} \sum_{i=1}^{m} 1_{[0,x_i]}(r_T)
\]
the cumulative distribution function of the switching points; it can be also be interpreted as the share of agents that have preferences described by $u^1$ for the realization $r_T$ of $R_T$. By further assuming that $\omega$ is
<table>
<thead>
<tr>
<th>Weighting function</th>
<th>$f(x)$</th>
<th>$x$ Parameter</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$x$</td>
<td>1:19</td>
<td></td>
</tr>
<tr>
<td>Exponential</td>
<td>$\frac{1}{1+e^{-x}}$</td>
<td>-9:9</td>
<td></td>
</tr>
<tr>
<td>Logistic</td>
<td>$(1-\alpha)^x$</td>
<td>0:18</td>
<td>$\alpha = 0.25$</td>
</tr>
<tr>
<td>Bell-shaped</td>
<td>$\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$</td>
<td>[-3 3]</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Weighting function and associated parameters

nonincreasing we can define the positive weights $w(r_T) = d\omega(r_T)/dr_T$, with $\int_0^\infty w(r_T)dr_T = 1$, designating the proportion of market participants that switch from $u^0$ to $u^1$ on $[r_T, r_T + dr_T]$. We exemplify with the utility functions $u^0(y) = b_0 y^\gamma$ and $u^1(y) = b_1 y^\gamma$, for some positive constants $b_0 < b_1$, that retain the monotonicity relationship between the Arrow-Pratt coefficients of absolute risk aversion established in the previous example. Without loss of generality we assume that the agents are equally important, that is $\alpha_1 = \alpha_2 = \cdots = \alpha_m = \alpha$. We illustrate the results in Figure 3 for $\gamma^0 = \gamma^1 = 0.5$, $b_0 = 1$ and $b_1 = 1.2$ and the weight functions computed using the formula: $w(r_T) = \frac{f(x)}{c}$, with $c$ a normalizing constant, that insures that $w$ is a density function on the interval $[0.9 1.1]$ and $f$ defined in Table 1. Given our parametric specifications for the utility indexes and $\omega$ we can rewrite equation 14 as

$$
\tilde{K}_\pi(r_T) = \left[ \frac{r_T}{\{1 - \omega(r_T)\}b_0^\gamma + \omega(r_T)b_1^\gamma} \right]^{-\gamma},
$$

for every possible realization $r_T$ of $R_T$.

5 The Calibration Problem

In section 4 we developed a microeconomic framework, where in equilibrium the pricing kernel $\tilde{K}_\pi(R_T)$ may be expressed by a mixture of different marginal utilities

$$
\tilde{K}_\pi(R_T) = \sum_{i=1}^{m} 1_{[x_{i-1},x_i]}(R_T) \frac{du_i}{dy} \bigg|_{y=R_T} + 1_{[x_m,\infty]}(R_T) \frac{du_{m+1}}{dy} \bigg|_{y=R_T}
$$

13
for some \( m \in \mathbb{N} \), \( 0 = x_0 < \ldots < x_m < \infty \) and mappings \( u_1, \ldots, u_{m+1} : [0, \infty] \to \mathbb{R} \cup \{-\infty\} \) whose restrictions to \([0, \infty[\) are strictly increasing, strictly convex twice continuously differentiable real-valued functions satisfying the Inada conditions. The values \( x_1, \ldots, x_m \) will be called the switching points of the pricing kernel \( \widetilde{K}_\pi(R_T) \). They are the amounts of returns where the investors are feeling themselves compelled to change their risk attitudes. Note that neither the number and the location of the switching points nor the marginal utilities of the \( u_1, \ldots, u_{m+1} \) are known. In order to calibrate them we suggest to use basis functions. For that purpose fix a set \( \mathcal{V} \) of strictly decreasing continuously differentiable mappings \( v : ]0, \infty[ \to \mathbb{R} \) satisfying \( \lim_{x \to \infty} v(x) = 0 \). As test functions for the data fitting use mixtures \( \sum_{i=1}^{N} v_i(x) 1_{[x_{i-1}, x_i)}(r_T) \) of functions from \( \mathcal{V} \). The quality of the approximation may be expressed by quadratic deviations w.r.t. a fixed continuous distribution \( \hat{F} \) with density function say \( \hat{p} \) and compact support enclosed in \([0, \infty[\). The mapping \( \hat{p} \) might be a kernel estimation of a density function for the distribution of \( R_T \). Henceforth we shall denote the minimum and maximum of the support by \( z \) and \( \bar{z} \) respectively.

We shall focus on the choice of appropriate basic values of the pricing kernel and the distribution of the switching points for a given partition of the support points. We might be also interested in the location of the switching points for some \( N \) which follow from the first step, for some preference parameters. The second objective is also possible to reach by approximating \( v(x) \) by sieves with interval varying parameters. Technically, for known approximating functions, the inverse problem translates in finding the sieve coefficients and the support points of the piecewise non-increasing pricing kernels. However, without some assumptions about the market composition the economic interpretation of the estimates is difficult. An advantage of the approach, however, is that it can easily incorporate economic restrictions (e.g. as integrability of the RND to one). Furthermore, the approximating functions can be chosen so that it reflects the preferences of a group of agents with 'identical' switching point, in which situations, parameters of the approximating functions are free as well and there is a link between the number of the switching points and number of the approximating functions.

Since our focus herein is on the aggregated market behavior we will follow the first approach. It is based on the discretization of \( \tilde{K}_\pi([z, \bar{z}]) \) on some basic grid. More precisely, we choose a partition \( z \overset{\text{def}}{=} x_0 \leq x_1 < \ldots < x_N \overset{\text{def}}{=} \bar{z} \) of \([z, \bar{z}]\), and use \( \tilde{K}_\pi(x_1, \ldots, x_N) \overset{\text{def}}{=} \sum_{i=1}^{N} \tilde{K}_\pi(x_i) 1_{[x_{i-1}, x_i)} \) as an approximation of \( \tilde{K}_\pi([z, \bar{z}]) \). This idea suggests itself by the following observation.
Proposition 5.1 The mapping \( \tilde{K}_\pi \sum_{i=1}^{N} \tilde{K}_\pi(a + \frac{i}{N}(b-a))1_{[a+(i-1)(b-a),a+n(b-a)]} \) converges to \( \tilde{K}_\pi[a, b] \) uniformly on compacta of continuity points of \( \tilde{K}_\pi[a, b] \) for any nondegenerated interval \( [a, b] \subseteq [0, \infty] \).

Proposition 5.1 may be derived immediately from the well known Korovkin like approximation result for mappings on \( [0, 1] \) (cf. e.g. Witting and Müller-Funk (1995), Satz B5.2).

Here, we shall not care about the choice of \( N \), it will be assumed to be exogenously given - in the following we will evaluate \( v \) at the observation points and then find the switching points for different \( N \). Denoting by \( Z_N \) the set of all \( (x_1, ..., x_N) \in \mathbb{R}^N \) satisfying \( x_0 \overset{\text{def}}{=} z \leq x_1 \leq ... \leq x_N = \bar{z} \), the choice of the basic points might be described in the following terms.

**Inverse problem:**

Find \( (x_1^*, ..., x_N^*) \in Z_N \) such that

\[
\inf_{v_1, ..., v_N \in \mathcal{V}} \int \sum_{i=1}^{N} \left[ \tilde{K}_\pi(x_i^*) - v_i(x) \right]^2 1_{[x_{(i-1)}, x_i^*]}(x) \hat{p}(x) \, dx = \inf_{(x_0, ..., x_N) \in Z_N} \inf_{v_1, ..., v_N \in \mathcal{V}} \int \sum_{i=1}^{N} \left[ \tilde{K}_\pi(x_i) - v_i(x) \right]^2 1_{[x_{(i-1)}, x_i]}(x) \hat{p}(x) \, dx
\]

As a convention we tacitly set \( \tilde{K}_\pi(0) \overset{\text{def}}{=} \infty \) and \( \infty \cdot 0 \overset{\text{def}}{=} 0 \). We shall discuss the solvability of the inverse problem in Appendix D.

### 5.1 Real Data Application

Section 5 exposes the calibration methodology for known or observable pricing kernels as a function of the index payoffs. In practice however, they are not directly observable and one can retrieve them from call prices/strike prices data using nonparametric techniques, see Ait-Sahalia and Lo (2000), Jackwerth (2000), Rosenberg and Engle (2002), Chernov (2003). In our real data example, we use a sample of EPK estimates from April 2003 to June 2006 for one month investment horizon (21 trading days) in DAX index as also used in Grith et al. (2010) - the detailed explanation of the estimation methodology can be found in the reference hereof.

The problem with fitting basic pricing kernel to the observed functionals, is that they will be rather flat - constant - for the regions with increased values and there results are not compatible with the assumption of no switching point on such interval. Therefore, we consider the case of the pricing kernel
Figure 4: Upper panel left: $\hat{K}_\pi$ (dashed) and $v^*$ (rombs); right: the implied absolute risk aversion. Lower panel left weight function $w$ associate with $\omega^*$, right: $\hat{p}$ (solid), $\hat{q}$ (rombs)

$v(s_j; \omega(s_j), b_0, b_1, \gamma)$ specified by 15; that is, we maintain the assumption of heterogeneous agents with possible distinct switching points. We will show that such a specification can approximate well the observed pricing kernel. Based on an estimate $\hat{K}_\pi$ of the pricing kernel discretized at points $j = 0, \ldots, n$, the operationalization of the inverse problem described in 16 implies solving for $\omega^*, b_0^*, b_1^*, \gamma^* = \arg \min F$, with

$$F = \sum_{j=1}^{n} \left\{ \hat{K}_\pi(s_j) - v(s_j; \omega(s_j), b_0, b_1, \gamma) \right\}^2 \hat{p}(s_j) \Delta_j$$

where $\omega = \{\omega^*(s_j)\}_{j=1}^{n}$, $n$ gridpoints and $\Delta_j = s_j - s_{j-1}$, under the restrictions that $1 - \omega$ is a cumulative distribution function and $\hat{q} = \hat{v}^* \hat{p}$, $v^* = v(\cdot; \omega(s_j), b_0^*, b_1^*, \gamma^*)$ is a density.

Notice, that in the problem above we do not necessitate $N$ and the problem of finding the optimal location
of the switching points can be restated in terms of finding the optimal $\omega$ at the grid points. We display the estimation results in figure 4.

Given preference parameters, we can estimate the switching points $x^*_1, ..., x^*_N = \arg \min F_N$, with

$$F_N = \sum_{j=1}^{n} \left\{ \hat{K}_\pi(s_j) - \sum_{i=1}^{N+1} v_i(s_j) I[x_{i-1}, x_i](s_j) \right\}^2 \hat{p}(s_j) \Delta_j$$

for given $N$ and $v_i(s) = v(s; i/N, b_i^0, b_i^*, \gamma^*)$. Results are showed in figure 5.

The 'optimal' number of switching points is to be found by comparing the marginal rate of improvement of the function $F_N$ with a threshold. In figure 5 we observe that the minimum of the objective function $F_N$ becomes relatively flat with $N = 4$. In choosing an optimal number of switching points one should balance parsimony with the precision of estimation; therefore, an additional function that penalizes the number of switching points can be used on deciding on the optimal $N$. 

Figure 5: $\hat{K}_\pi$ (dashed) and $v^*_i$ (solid) for $N = 1, 2, 3, 4$
Figure 6: Upper panel: The relationship between the minimum of the objective function $F_N$ and the number of switching points $N$. Lower panel: cdf of the switching points implied by 17 (rombs) and by 17 for $N = 2$ (dash-dotted), $N = 3$ (dashed), $N = 4$ (solid).

6 Conclusions

Based on our specification for the marginal investors’ preferences it turns out that dependent on the stock prices the v. Neumann-Morgenstern utility index for the representative agent might switch between different types of utilities, meaning possible changes of the risk attitudes. The reference points and the types of utilities describe local risk attitudes of the investors. We empirically investigate the switching behavior of the pricing kernel in a simulation study and real data example. While our analysis is conducted for a fixed investment horizon, since we are only taking a snapshot of the market and try to explain the observed shape in the pricing kernel, it can also be implemented in a dynamic context. In the last figure 7, for instance, we plot the estimated preference parameters estimated independently for each day. These,
Figure 7: Preference parameters estimated between April 2003 and June 2006

together with the distribution of the switching points define the evolution of the market pricing kernel. In addition, we only considered the market return as a state variable in the utility function. An extension of the model should include additional state variables or alternatively try to explain what determines the switching points.

A Appendix

We continue with the consumption model of section 4, retaking all assumptions and notations. The aim of this section is to provide a proof for Theorem 4.1. For this purpose let us firstly characterize the optimal
consumptions \( \bar{c}_1(S_T), ..., \bar{c}_m(S_T) \) of the individual consumer.

Assumption (3.5) enables us to apply the dominated convergence theorem to show

(A1) continuity of the mappings

\[
g^i_{s_T}: 0, \infty \to \mathbb{R}, y \mapsto I_i\{s_T, y\mathcal{K}_\pi(s_T)\}\mathcal{K}_\pi(s_T) \quad (s_T \geq 0, i \in \{1, ..., m\}).
\]

Furthermore, the Inada conditions together with monotone and dominated convergence imply

(A2) \( \lim_{y \to 0} g^i_{s_T}(y) = \infty \) and \( \lim_{y \to \infty} g^i_{s_T}(y) = 0. \)

We are now ready to extend the classical characterization of the optimal consumption to the case of extended expected utility preferences.

**Theorem A.1** Assuming (4) – (11), there exist \( y_1, ..., y_m > 0 \) such that

\[
\bar{c}_i(S_T) = I_i\{S_T, y_i\mathcal{K}_\pi(S_T)\} \quad \text{for} \quad i = 1, \ldots, m
\]

**Proof:**

Let us fix \( i \in \{1, ..., m\} \) and denote \( x_i \overset{\text{def}}{=} w_i^0 + \mathbb{E}[e_i(S_T)\mathcal{K}_\pi(S_T)] \). Since \( x_i > 0 \) we may find in view of (A1), (A2) some \( y_i > 0 \) with \( g(y_i) = x_i \).

Let \( c(S_T) \) be a nonnegative random variable with \( \mathbb{E}[c(S_T)\mathcal{K}_\pi(S_T)] \leq x_i \). Then

\[
\mathbb{E}[u\{S_T, c(S_i)\}] + y_i(x_i - \mathbb{E}[c(S_T)\mathcal{K}_\pi(S_T)]) = y_i x_i + \mathbb{E}[u\{S_T, c(S_i)\}] - y_i c(S_T)\mathcal{K}_\pi(S_T) \leq \\
y_i x_i + \sup_{x \geq 0} \mathbb{E}[u(S_T, x) - y_i x\mathcal{K}_\pi(S_T)] = \\
y_i x_i + \mathbb{E}[u(S_T, I_i\{S_T, y_i\mathcal{K}_\pi(S_T)\})] - y_i I_i\{S_T, y_i\mathcal{K}_\pi(S_T)\}\mathcal{K}_\pi(S_T)] = \mathbb{E}[u\{S_T, I_i(S_T, y_i\mathcal{K}_\pi(S_T))\}].
\]

Therefore \( I_i(S_T, y_i\mathcal{K}_\pi(S_T)) \) solves the optimization problem of consumer \( i \). Moreover, the numerical representation \( U_i \) of consumer’s \( i \) preferences is strictly concave in view of strict concavity of \( u^i(x, \cdot) \) for every \( x \geq 0 \). In particular \( I_i(S_T, y_i\mathcal{K}_\pi(S_T)) \) is the unique solution, hence being identical with \( \bar{c}_i(S_T) \). \( \square \)

Before starting with the proof of Theorem 4.1 let us consider for purposes of reference the classical case of the consumer being expected utility maximizer. Indeed as an additional corollary of Theorem 4.1, we may retain the folk result concerning the risk neutral price valuation and the v. Neumann-Morgenstern utility index of the representative agent. More precisely, let us assume that there exist mappings \( u_1, ..., u_r \) from \( \mathbb{R}_+ \) into \( \mathbb{R} \cup \{-\infty\} \) satisfying \( u^1(x, \cdot) = u_1, ..., u^m(x, \cdot) = u_m \) for \( x \geq 0 \), and
(A3) \( u_1(y), ..., u_m(y) \in \mathbb{R} \) for \( y > 0 \),

(A4) \( u_1, ..., u_m \) are continuous, strictly increasing as well as strictly concave.

Then
\[
u(y) = \sup \left\{ \sum_{i=1}^{m} \alpha_i u_i(y_i) \mid y_1, ..., y_m \geq 0, \sum_{i=1}^{m} y_i \leq y \right\} = u_\alpha(x, y) \text{ for } x, y \geq 0.
\]
The adaptions of (10), (11) read as follows. We shall impose the so called Inada conditions on the state independent utility indices \( u_1, ..., u_m \), i.e.

(A5) \( u_1|_{0,\infty}, ..., u_m|_{0,\infty} \) are assumed to be continuously differentiable satisfying
\[
\lim_{e \to 0} \frac{du_i}{dy} \bigg|_{y=e} = \infty, \quad \lim_{e \to \infty} \frac{du_i}{dy} \bigg|_{y=e} = 0 \quad (i = 1, \ldots, m).
\]

(A6) \( E[I_1(y K_\pi(S_T))], \ldots, E[I_m(y K_\pi(S_T))] < \infty \) for any \( y > 0 \), where \( I_1, \ldots, I_r \) denote the inverses of \( \frac{du_1}{dy}, \ldots, \frac{du_m}{dy} \) respectively.

We may conclude immediately from Theorem 4.1 the announced result.

**Corollary A.2** Let \( \bar{e}(S_T) = \frac{S_T}{S_0} \) and let \( u_1|_{0,\infty}, ..., u_m|_{0,\infty} \) be twice continuously differentiable. Then under (A3) - (A6), \( u|_{0,\infty} \) is continuously differentiable, and for any \( \alpha_i > 0 \) there exists some \( \beta_i > 0 \) such that
\[
\frac{du_i}{dy} \bigg|_{y=\bar{e}(S_T)} = \alpha_i \frac{du_i}{dy} \bigg|_{y=\bar{c}_i(S_T)} = \alpha_i \beta_i K_\pi(S_T) \text{ for any realization } s_T \text{ of } S_T.
\]

**Proof of Theorem 4.1:**

Without loss of generality let us set \( \{1, \ldots, r\} \overset{\text{def}}{=} \{ i \in \{1, \ldots, m\} \mid \alpha_i > 0 \} \). Then, defining \( g_i \overset{\text{def}}{=} \alpha_i u_i \), we have \( u_\alpha = \sum_{i=1}^{r} g_i \), and we may apply Lemmata B.1, B.2 and Proposition B.3 (cf. Appendix B). Then, in view of Lemmata B.1, B.2 and B.3, we obtain
\[
u_\alpha(s_T, \bar{e}[s_T]) = \sum_{i=1}^{r} \alpha_i u^i(s_T, \bar{c}[s_T])
\]
for every realization \( s_T \) of \( S_T \).

On one hand by Theorem A.1, there exist \( y_1, ..., y_m > 0 \) such that
\[
\bar{c}_i(S_T) = I_i(S_T, y_i K_\pi(S_T)) > 0 \quad \text{for } i = 1, \ldots, r.
\]
On the other hand, due to Proposition B.3, \( u_\alpha(s_T, \cdot) |_{[0, \infty]} \) is differentiable for every realization \( s_T \), satisfying
\[
\alpha_i \frac{u^i(s_T, \cdot)}{dy} \bigg|_{y \equiv \bar{e}(s_T)} = \frac{u(s_T, \cdot)}{dy} \bigg|_{y \equiv \bar{e}(s_T)}
\]
for \( i \in \{1, \ldots, r\} \) and any realization \( s_T \). Notice that by construction the random variable \( \bar{e}(S_T) \) has strictly positive outcomes only. Now, the statement of Theorem 4.1 is clear. \(\square\)

### B Appendix

Throughout this section let the mappings \( g_1, \ldots, g_r : \mathbb{R}^2_+ \to \mathbb{R} \cup \{-\infty\} \) satisfy the following conditions:

\begin{enumerate}[(B0)]
  \item \( g_1(x, y), \ldots, g_r(x, y) \in \mathbb{R} \) for \( x \geq 0, y > 0 \);
  \item \( g_1(x, \cdot), \ldots, g_r(x, \cdot) \) are continuous, strictly increasing and strictly concave for \( x \geq 0 \);
  \item \( g_1(\cdot, y), \ldots, g_r(\cdot, y) \) are Borel-measurable for \( y \geq 0 \).
\end{enumerate}

Furthermore, let \( g : \mathbb{R}^2_+ \to \mathbb{R} \cup \{-\infty, \infty\} \) be defined by
\[
g(x, y) = \sup \left\{ \sum_{i=1}^{r} g_i(x, y_i) \mid y_1, \ldots, y_r \geq 0, \sum_{i=1}^{r} y_i \leq y \right\}.
\]

Indeed \( g(x, 0) = \sum_{i=1}^{r} g_i(x, 0) \in \mathbb{R} \cup \{-\infty\} \) for \( x \geq 0 \), and
\[
-\infty < \sum_{i=1}^{r} g_i(x, \frac{y}{r}) \leq g(x, y) \leq \sum_{i=1}^{r} g_i(x, y) < \infty
\]
for \( x \geq 0, y > 0 \) due to (B0), (B1).

**Lemma B.1** For any \( x, y \geq 0 \) there is some unique \( \phi(x, y) = (\phi_1(x, y), \ldots, \phi(x, y)) \in \mathbb{R}^m_+ \) such that
\[
\sum_{i=1}^{m} \phi_i(x, y) \leq y \quad \text{and} \quad \sum_{i=1}^{r} g_i \{x, \phi_i(x, y)\} = g(x, y).
\]

Furthermore, \( \sum_{i=1}^{m} \phi_i(x, y) = y \).

22
Proof:
Let \( x, y \geq 0 \). For \( y = 0 \) the statement of Lemma B.1 is obvious. So let \( y > 0 \), which means \( g(x, y) \in \mathbb{R} \).
Due to (B1), the mapping
\[
f : \left\{ (y_1, \ldots, y_r) \in \mathbb{R}^r_+ \mid \sum_{i=1}^r y_i \leq y, \sum_{i=1}^r g_i(x, y_i) \geq g(x, y) - 1 \right\} \to \mathbb{R}, (y_1, \ldots, y_r) \mapsto \sum_{i=1}^r g_i(x, y_i)
\]
is continuous, strictly concave, and defined on a nonvoid convex compact set. Therefore \( f \) attains its maximum at a unique \( \phi(x, y) \).
Obviously, \( \sum_{i=1}^r \phi_i(x, y) = y \) because \( f \) is strictly increasing too by (B1).
The proof is complete. \( \square \)

Lemma B.1 defines a mapping \( \phi = (\phi_1, \ldots, \phi_r) : \mathbb{R}^2_+ \to \mathbb{R}^r_+ \). It is Borel-measurable as will be shown now.

Lemma B.2 \( \phi \) is Borel-measurable.

Proof:
It suffices to show that \( \phi^{-1}\left( \bigtimes_{i=1}^r [0, a_i] \right) \) is a Borel-subset of \( \mathbb{R}^2_+ \). For this purpose define for any \( (a_1, \ldots, a_r) \in \mathbb{R}^r_+ \) the mapping \( g_{a_1 \ldots a_r} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\} \) by
\[
g_{a_1 \ldots a_r}(x, y) = \sup \left\{ \sum_{i=1}^r g_i(x, y_i) \mid (y_1, \ldots, y_r) \in \bigtimes_{i=1}^r [0, a_i], \sum_{i=1}^r y_i \leq y \right\}.
\]
Notice that \( g_{a_1 \ldots a_r}(x, y) \in \mathbb{R} \) for \( x \geq 0, y > 0 \), analogously to \( g(x, y) \in \mathbb{R} \) for \( x \geq 0, y > 0 \). Furthermore \( g_1(x, \cdot), \ldots, g_r(x, \cdot) \) are continuous for any \( x \geq 0 \). Hence, setting \( \mathcal{R}_{a_1 \ldots a_r} = \bigtimes_{i=1}^r [0, a_i] \times \mathbb{Q}^m \),
\[
g^{-1}_{a_1 \ldots a_r}(\lfloor z, \infty \rfloor) = \bigcup_{(y_1, \ldots, y_r) \in \mathcal{R}_{a_1 \ldots a_r}} \left( \sum_{i=1}^r \alpha_i g_i(\cdot, y_i) \right)^{-1}(\lfloor z, \infty \rfloor) \times \left[ \sum_{i=1}^r y_i, \infty \right] (z \in \mathbb{R}).
\]
Thus \( g^{-1}_{a_1 \ldots a_r}(\lfloor z, \infty \rfloor) \) is a Borel-subset of \( \mathbb{R}^2_+ \) for every \( z \in \mathbb{R} \) by assumption (B2). Then we may conclude that
\[
\phi^{-1}\left( \bigtimes_{i=1}^r [0, a_i] \right) = \left( \sup_{(b_1, \ldots, b_r) \in \mathbb{Q}^r_+} g_{b_1 \ldots b_r} - g_{a_1 \ldots a_r} \right)^{-1}(\{0\})
\]
is a Borel subset of \( \mathbb{R}^2_+ \) for any \( (a_1, \ldots, a_r) \in \mathbb{R}^r_+ \), which completes the proof. \( \square \)

In order to characterize the mapping \( \phi \) in terms of derivatives of the functions \( g_1(x, \cdot), \ldots, g_r(x, \cdot) \), it is customary to impose the Inada conditions, i.e.
for any \( x \geq 0 \) the mappings \( g_1(x, \cdot) | 0, \infty[, ..., g_r(x, \cdot) | 0, \infty[ \) are assumed to be continuously differentiable satisfying
\[
\lim_{\epsilon \to 0} \frac{\partial g_i(x, \cdot)}{\partial y} \bigg|_{y=\epsilon} = \infty, \quad \lim_{\epsilon \to \infty} \frac{\partial g_i(x, \cdot)}{\partial y} \bigg|_{y=\epsilon} = 0, \quad i = 1, \ldots, r.
\]

The Inada conditions together with condition (B1) imply that for any \( i \in \{1, \ldots, r\} \) and every \( x \geq 0 \) the mapping \( \frac{\partial g_i(x, \cdot)}{\partial y} | 0, \infty[ \) is injective onto \( ]0, \infty[ \) with continuously differentiable, strictly decreasing inverse say \( I_i(x, \cdot) \).

**Proposition B.3** Let the assumptions (B0) - (B3) be fulfilled, and let \( g_1(x, \cdot) | 0, \infty[, ..., g_r(x, \cdot) | 0, \infty[ \) be twice continuously differentiable. Then for any \( x \geq 0 \) the mapping \( g(x, \cdot) | 0, \infty[ \) is differentiable satisfying
\[
\phi(x, y) = \left( I_1 \left[ x, \frac{\partial g_1(x, \cdot)}{\partial y} \big| y \right], \ldots, I_r \left[ x, \frac{\partial g_r(x, \cdot)}{\partial y} \big| y \right] \right) \text{ for } y > 0.
\]

**Proof:**

Let for \( x \geq 0 \) the mapping \( F_x : ]0, \infty[ \times ]0, \infty[ \to \mathbb{R} \) be defined by \( F_x(y, z) = \sum_{i=1}^{r} I_i(x, z) - y \).

Since the mappings \( g_1(x, \cdot) | 0, \infty[, ..., g_r(x, \cdot) | 0, \infty[ \) are assumed to be strictly concave and twice continuously differentiable, their second derivatives are strictly negative. Then by local inverse theorem the mappings \( I_1(x, \cdot), ..., I_r(x, \cdot) \) are continuously differentiable, having strictly negative derivatives. In particular \( F_x \) is continuously differentiable, satisfying
\[
\frac{\partial F_x}{\partial z} \bigg|_{(y,z)} \neq 0 \text{ for } y, z > 0.
\]

Furthermore, since \( I_1(x, \cdot), ..., I_r(x, \cdot) \) are continuous and strictly decreasing mappings onto \( ]0, \infty[ \), we may find for any \( y > 0 \) a unique \( \varphi(y) > 0 \) with \( F(y, \varphi(y)) = 0 \). Drawing on the implicit function theorem, \( y \mapsto \varphi(y) \) defines a differentiable mapping \( \varphi : ]0, \infty[ \to ]0, \infty[ \).
Moreover, for \( y > 0 \) and \( y_1, \ldots, y_r \geq 0 \) with \( \sum_{i=1}^r y_i \leq y \), we may conclude

\[
\sum_{i=1}^r g_i(x, y_i) + \varphi(y)(y - \sum_{i=1}^r y_i) = \varphi(y)y + \sum_{i=1}^r \{g_i(x, y_i) + \varphi(y)y_i\} \leq \\
\varphi(y)y + \sum_{i=1}^r \sup_{z \geq 0} \{g_i(x, z) + \varphi(y)z\} = \\
\varphi(y)y + \sum_{i=1}^r [g_i(x, I_i(x, \varphi(y))) + \varphi(y)I_i(x, \varphi(y))] = \\
\sum_{i=1}^r g_i[x, I_i\{x, \varphi(y)\}] - F_x\{y, \varphi(y)\} = \sum_{i=1}^r g_i[x, I_i\{x, \varphi(y)\}].
\]

This means

\[ g(x, y) = \sum_{i=1}^r g_i[x, I_i\{x, \varphi(y)\}], \]

and hence by Lemma B.1

\(*) \quad \varphi(x, y) = (I_1[x, \varphi(y)], \ldots, I_r[x, \varphi(y)]). \]

As a further consequence \( g(x, \cdot)||0, \infty[ \) is differentiable satisfying

\[
\frac{dg(x, \cdot)}{dy} \bigg|_y = \sum_{i=1}^r \varphi(y) \frac{dI_i(x, \cdot)}{dy} \bigg|_y = \varphi(y) \frac{d}{dy} \left( \sum_{i=1}^r I_i(x, \cdot) \circ \varphi \right) \bigg|_y = \varphi(y).
\]

For the last equation notice that \( \sum_{i=1}^r I_i(x, \cdot) \circ \varphi \) is just the identity on \( ]0, \infty[ \). In view of \( (*) \) the proof is complete. \( \square \)

### C Appendix

**Proposition C.1** Let \( g : ]0, \infty[ \to \mathbb{R} \) be a mapping whose set of continuity points \( C(g) \) is an open subset of \( \mathbb{R} \). Furthermore, let \( F \) be an atomless distribution with satisfying \( F([z, \infty[) = 1 \) for some \( z \geq 0 \). Additionally, let \( (X_N)_N \) denote an \( i.i.d. \) sequence of random variables over some probability space \( (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}) \) with common distribution \( F \), inducing for any \( N \) the order statistics \( X_{1:N}, \ldots, X_{N:N} \) with \( z \overset{\text{def}}{=} X_{0:N} \leq X_{1:N} \leq \ldots \leq X_{N:N} \). Then for any continuity point \( x \) of \( g \) which belongs to the topological interior \( \text{int}(\text{supp}(F)) \) of
the support of \( F \) we obtain
\[
\sum_{i=1}^{N} g(X_{i:N}) 1_{[X_{(i-1):N},X_{i:N}]} \rightarrow g(x) \text{ a.s.}
\]

In particular, setting \( i(F) := \inf \{ x : F(]-\infty, x]) > 0 \} \) and \( s(F) := \sup \{ x : F(]-\infty, x]) < 1 \} \), we may even achieve this convergence for any \( x \in C(g) \cap \{ i(F) \}, s(F) \} \text{ if } supp(F) \text{ is convex.}

**Proof:**

Let \( x \in C(g) \cap \text{int}(supp(F)) \), let \( q^-_F, q^+_F \) denote the lower and upper quantile functions of \( F \) respectively, and let us define \( \alpha_0 := F(]-\infty, x]) \). Since \( F \) is atomless, and since \( x \in \text{int}(supp(F)) \) we have \( \alpha_0 \in ]0, 1[ \) and \( q^-_F(\alpha_0) = q^+_F(\alpha_0) = x \).

Let \( \varepsilon > 0 \) be fixed. By assumption \( C(g) \) is open so that there is some \( \delta > 0 \) such that \( g[|x - \delta, x + \delta[ \) is continuous, and \( |g(y) - g(x)| < \varepsilon \) for any \( y \in ]x - \delta, x + \delta[ \). Furthermore, the lower and upper quantile functions of \( F \) differ on at most countable points only, and the distribution function of \( F \) is continuous.

Hence we may find an antitone sequence \( (\alpha_n)_{n \in \mathbb{N}} \) in \( ]\alpha_0, 1[ \) which converges to \( \alpha_0 \) such that \( q^-_F(\alpha_n) := q^-_F(\alpha_n) = q^+_F(\alpha_n) \) for every \( n \in \mathbb{N} \).

We may draw on the asymptotic theory of order statistics to find some \( A \in \tilde{F}, P(A) = 1 \), such that
\[
\lim_{N \to \infty} X_{1:N}(\omega) = i(F) \text{ for } \omega \in A, \text{ and } \lim_{N \to \infty} X_{[N\alpha_j+1]:N}(\omega) = q^-_F(\alpha_j)
\]
holds for every \( \omega \in A \) and any \( j \in \mathbb{N} \), where \( [N\alpha_j + 1] \) denotes the largest \( l \in \mathbb{N} \) with \( l \leq N\alpha_j + 1 \) (cf. Witting and Müller-Funk (1995), Satz 7.108, Satz 7.120).

Now let us fix \( \omega \in A \). Putting things together, we have \( q^-_F(\alpha_j) \in ]x, x + \delta[ \), and thus
\[
X_{1:N}(\omega) < x < X_{[N\alpha_j+1]:N}(\omega) < x + \delta
\]
for sufficiently large \( j \) and \( N \). Hence, setting \( i_N := \min \{ i \in \{1, \ldots, N\} : x \leq X_{i:N}(\omega) \} \), we obtain the inequalities \( X_{i:N}(\omega) \leq X_{[N\alpha_j+1]:N}(\omega) < x + \delta \), implying
\[
\left| \sum_{i=1}^{N} g(X_{i:N}) 1_{[X_{(i-1):N},X_{i:N}]}(\omega) - g(x) \right| = |g(X_{i:N}(\omega)) - g(x)| < \varepsilon.
\]

The remaining part of follows immediately from the observation that \( \text{int}(supp(F)) = \{ i(F) \}, s(F) \} \text{ is valid for convex } supp(F) \).

\[ \square \]

26
D Appendix

Throughout this section we shall assume

$$\sup_{v \in V} \int v(x)^{2+\delta} \hat{p}(x) \, dx < \infty$$

for some $\delta > 0$.

**Remark D.1** In the following situations the family $V$ fulfills the assumed integrability condition:

1. $\int_0^1 \hat{p}(x) \, dx < \infty$ for some $\delta > 0$, and $V \overset{\text{def}}{=} \{ v : [0, \infty] \to \mathbb{R}, x \mapsto (x+a)^{b-1} \mid a \geq 0, b \in ]0,1[ \}$.

2. $z > 0$, and $\sup_{v \in V} v(z) < \infty$, e.g. if $V \overset{\text{def}}{=} \{ v : [0, \infty] \to \mathbb{R}, x \mapsto (x+a)^{b-1} \mid a \geq 0, b \in ]0,1[ \}$.

The integrability condition means that $V$ is $L_2$–norm bounded. In particular $V$ as well as $\{ v^2 \mid v \in V \}$ are uniformly $\hat{F}$–integrable, and the weak closure $cl^w(V)$ of $V$ is weakly compact because $L_2(\hat{F})$, equipped with the $L_2$–norm, is a reflexive Banach space. Moreover, since the standard Borel $\sigma$–algebra on $\mathbb{R}$ is countably generated, the $L_2$–norm topology on $L_2(\hat{F})$ is separable. Hence the relative topology of the weak topology to $cl^w(V)$ is metrizable (cf. Dunford and Schwarz (1958), Theorem V.6.3), and thus as a compact topology also separable. So we obtain the following result.

**Lemma D.2** Under the integrability condition on $V$ the weak closure $cl^w(V)$ of $V$ is weakly compact, and the relative weak topology to $cl^w(V)$ is separably metrizable.

For abbreviation let us define $g : \mathbb{Z}_N \times L_2(\hat{F})^N \to \mathbb{R} \cup \{ \infty \}$ by

$$g(x_1, \ldots, x_N, v_1, \ldots, v_N) = \begin{cases} \sum_{i=1}^N (K(x_i) - v_i(x))^2 1_{[x_{i-1}, x_i]}(x) \hat{p}(x) \, dx : (v_1, \ldots, v_N) \in V^N \\ \infty : \text{ otherwise} \end{cases}$$

Let us gather some basic properties of $g$.

**Lemma D.3** Let $\tau_1$ be the relative topology of the standard topology on $\mathbb{R}^N$ to $\mathbb{Z}_N$, let $\tau_2$ denote the product weak topology on $L_2(\hat{F})^N$, and let $\tau_1 \times \tau_2$ stand for the product topology of $\tau_1$ and $\tau_2$. Furthermore let us denote the set of continuity points of $\tilde{K}_\pi$ by $\mathcal{C}(\tilde{K}_\pi)$. Then under the integrability condition on $V$ the mapping $g$ satisfies the following properties:

1. $g(x_1, \ldots, x_N, \cdot) \mid V^N$ is lower continuous w.r.t. $\tau_2$ for every $(x_1, \ldots, x_N) \in \mathbb{Z}_N$;

27
2. \( g(\cdot, v_1, ..., v_N) \) is measurable w.r.t. the Borel \( \sigma \)-algebra generated by \( \tau_1 \) for any \((v_1, ..., v_N) \in \mathcal{V}^N \), and the restrictions of these mappings to \( Z_N \cap C(\widetilde{K}_\pi)^N \) are even continuous w.r.t. \( \tau_1 \);

3. The family \( \left\{ g(\cdot, v) | Z_N \cap C(\widetilde{K}_\pi)^N \mid v \in \mathcal{V}^N \right\} \) is equicontinuous w.r.t. \( \tau_1 \), in particular the restriction of \( g \) to \( (Z_N \cap C(\widetilde{K}_\pi)^N) \times \mathcal{V}^N \) is lower semicontinuous w.r.t. \( \tau_1 \times \tau_2 \).

Proof:
The statement 1. may be verified by routine procedures. For that purpose it should be observed that we may restrict ourselves to sequential weak lower semicontinuity by Lemma D.2 and that the squared \( L^2 \)-norm is weakly lower semicontinuous as a convex \( L^2 \)-norm continuous mapping.

Statement 2. follows easily from measurability of \( \widetilde{K}_\pi \) and the following observation

\((*)\) For every \( \varepsilon_1 > 0 \), there is some \( \varepsilon_2 > 0 \) such that

\[
\sup_{w \in \mathcal{W}} \int |w(x)| \left| 1_{[y_1, y_2]}(x) - 1_{[z_1, z_2]}(x) \right| \hat{p}(x) \, dx < \varepsilon_1
\]

whenever \((y_1, y_2), (z_1, z_2) \in [0, \infty]^2 \) with \( y_1 \leq y_2, z_1 \leq z_2 \) and \((y_1 - z_1)^2 + (y_2 - z_2)^2 < \varepsilon_2 \). Here \( \mathcal{W} \) denotes an arbitrary family of uniformly \( \hat{F} \)-integrable real-valued mappings on \([0, \infty[\).

In order to see \((*)\) observe \( \int \left| 1_{[y_1, y_2]} - 1_{[z_1, z_2]} \right| \hat{p}(x) \, dx = \int \left| 1_{[y_1, y_2]} - 1_{[z_1, z_2]} \right| \hat{p}(x) \, dx \to 0 \) as \((z_1, z_2) \to (y_1, y_2)\), and that \( \left| 1_{[y_1, y_2]} - 1_{[z_1, z_2]} \right| \) is just indicator function of the symmetric difference of the involved intervals.

For the proof of statement 3. it suffices to show \( \tau_1 \)-equicontinuity of \( \left\{ g(\cdot, v) | Z_N \cap C(\widetilde{K}_\pi)^N \mid v \in \mathcal{V}^N \right\} \) because the second part of statement 3. follows then in view of statement 1.. So let us fix some \((x_1, ..., x_n) \in Z_N \cap C(\widetilde{K}_\pi)^N \). Since \( Z_N \cap C(\widetilde{K}_\pi)^N \in \tau_1 \), we may find a bounded neighborhood \( U \in \tau_1 \) of \((x_1, ..., x_n)\) such that \( U \subseteq C(\widetilde{K}_\pi)^N \cap [\rho, \tilde{z}]^N \) for some \( \rho > 0 \). Then we obtain for any \((\tilde{x}_1, ..., \tilde{x}_N) \in U \) and any \((v_1, ..., v_N) \in \mathcal{V}^N \)
Next let us introduce the lower semicontinuous envelope \( \text{lsc}(g) \) of \( g \), defined to be the largest \( \tau_1 \times \tau_2 \) -- lower semicontinuous \( \mathbb{R} \cup \{ \infty \} \) - valued mapping dominated by \( g \). Notice that \( g \) and \( \text{lsc}(g) \) coincide on the set \( \left( Z_N \cap C(\tilde{K}_\pi)^N \right) \times N^N \) by Proposition 3.6 in DalMaso (1993).

The lower semicontinuous envelope will turn out to be a useful tool concerning the solvability of the inverse problems posed in the subsections before. Let us begin with the one coming out of the pure numerical
approach. In terms of $g$ it reads as follows.

$$\begin{align*}
\text{minimize } & \inf_{v \in V} g(x_1, \ldots, x_N, v) \\
\text{among all } & (x_1, \ldots, x_N) \in \mathbb{Z}_N.
\end{align*}$$

Fortunately, we may apply directly Theorem 3.8 in DalMaso (1993), observing that $\mathbb{Z}_N \times cl^w(V)^N$ is compact and sequentially compact w.r.t. $\tau_1 \times \tau_2$, and by Lemma D.3

$$\lim_{m \to \infty} \left| \inf_{v \in V} g(x_m, v) - \inf_{v \in V} g(x_0, v) \right| \leq \lim_{m \to \infty} \sup_{v \in V} \left| g(x_m, v) - g(x_0, v) \right| = 0$$

for any sequence $(x_m)_{m \in \mathbb{N}_0}$ in $\mathbb{Z}_N \cap \mathcal{C}(\mathcal{K}_\pi)^N \in \tau_1$ with $x_m \to x_0$.

**Theorem D.4** Let $((x_m, v_m))_m$ be a sequence in $\mathbb{Z}_N \times V^N$ such that $\lim_{m \to \infty} g(x_m, v_m) = \inf_{v \in V} g(x_m, v)$. Then this sequence has at least one cluster point in $\mathbb{Z}_N \times cl^w(V)^N$, and every such cluster point $(x^*, v^*)$ satisfies

$$lsc(g)(x^*, v^*) = \min_{(x, v) \in \mathbb{Z}_N \times cl^w(V)^N} lsc(g)((x, v)) = \inf_{(x, v) \in \mathbb{Z}_N \times V^N} g(x, v),$$

with $\inf_{v \in V} g(x^*) = \inf_{\xi \in \mathbb{Z}_N} \inf_{v \in V} g(\xi, v)$ if the components of $x^*$ are continuity points of $\mathcal{K}_\pi$.

In view of Tonelli’s theorem the inverse problem we have formulated after introducing the Monte-Carlo approach may be described in terms of $g$ by

$$\begin{align*}
\text{minimize } & \inf_{v \in V^N} \mathbb{E}_{F(N)} \left[ g(X_{1:N}^F, \ldots, X_{N:N}^F, v) \right] \\
\text{among all } & F \in \mathcal{P}_N.
\end{align*}$$

Concerning the solvability of this problem we may observe the following result.

**Theorem D.5** Let $\tau_w$ be the topology of weak convergence on the set of distributions on $\mathbb{R}$, and let $\tau_w \times \tau_2$ denote the product topology of $\tau_w$ and the product weak topology $\tau_2$ on $L_2(\hat{F})^N$. Furthermore, let us consider a sequence $((F_m, v_m))_m$ in $\mathcal{P}_N \times V^N$ fulfilling

$$\lim_{m \to \infty} \mathbb{E}_{F_m(N)} \left[ g(X_{1:N}^{F_m}, \ldots, X_{N:N}^{F_m}, v_m) \right] = \inf_{F \in \mathcal{P}_N} \inf_{v \in V^N} \mathbb{E}_{F(N)} \left[ g(X_{1:N}^F, \ldots, X_{N:N}^F, v) \right].$$

30
Then this sequence has a cluster point \( w.r.t. \) to \( \tau_w \times \tau_2 \). Every such cluster point \((F^*, \nu^*)\) satisfies

\[
\min \left\{ \liminf_{m \to \infty} \mathbb{E}_{F_{m(N)}} \left[ g(X_{1:N}^{F_m}, ..., X_{N:N}^{F_m}, \tilde{\nu}_m) \right] \mid \tilde{\nu}_m \to \nu^* \right\}
\]

\[
= \min_{\nu \in L_2(F^1_N)} \min \left\{ \liminf_{m \to \infty} \mathbb{E}_{F_{m(N)}} \left[ g(X_{1:N}^{F_m}, ..., X_{N:N}^{F_m}, \tilde{\nu}_m) \right] \mid \tilde{\nu}_m \to \nu \right\}
\]

\[
= \inf_{F \in P_N} \inf_{\nu \in V_N} \mathbb{E}_{F(N)} \left[ g(X_{1:N}^F, ..., X_{N:N}^F, \nu) \right],
\]

where even \( \inf_{\nu \in V_N} \mathbb{E}_{F^*} \left[ g(X_{1:N}^F, ..., X_{N:N}^F, \nu) \right] = \inf_{F \in P_N} \inf_{\nu \in V_N} \mathbb{E}_{F(N)} \left[ g(X_{1:N}^F, ..., X_{N:N}^F, \nu) \right] \) holds if \( z > 0 \), and the set of discontinuity points of \( K_\pi \) is a \( F^* \)-null set (e.g. if \( F^* \) is a continuous distribution).

**Proof:**

First of all, since the topology of weak convergence for distributions on \( \mathbb{R} \) is completely as well as separably metrizable, and since the supports of the distributions \( F_m \) are uniformly enclosed in a compact subset of \( \mathbb{R} \), the sequence \((F_m)_m\) is uniformly tight, and therefore has a cluster point by Prokhorov’s theorem. Furthermore \( d^w(V)^N \) has been observed as sequentially \( \tau_2 \)-compact by Lemma D.2, so that we may conclude that \(((F_m, \tilde{\nu}_m))_m\) has a cluster point \( w.r.t. \) \( \tau_w \times \tau_2 \).

Now, let \((F^*, \nu^*)\) be a cluster point of \(((F_m, \nu_m))_m\). Without loss of generality we may assume that it is even a limit point. Then the induced sequence \((F_{m(N)})_m\) of the respective joint distributions of the order statistics \( X_{1:N}^{F_m}, ..., X_{N:N}^{F_m} \) has the joint distribution \( F_{m(N)}^* \) of \( X_{1:N}^{F^*}, ..., X_{N:N}^{F^*} \) as a limit point \( w.r.t. \) the topology of weak convergence on the set of distributions on \( \mathbb{R}^N \). Indeed, denoting the joint distribution of the order statistics \( U_{1:N}, ..., U_{N:N} \) obtained from an i.i.d sample of size \( N \) according to the uniform distribution on \( ]0,1[ \) by \( F_{U(N)} \), we have \( F_{m(N)} \left( \frac{X_{i-1}}{i} - \infty, x_i \right) = F_{U(N)} \left( \frac{X_{i-1}}{i} - \infty, F_m([- \infty, x_i]) \right) \) for any \( (x_1, ..., x_N) \in \mathbb{R}^N \), and an analogous expression for \( F_N^* \). Then the claim follows immediately from the Helly Bray theorem.

Now we may apply directly Theorem 7.8 from DalMaso (1993) to the sequence \((G_m)_m\) of functions \( G_m : L_2(\tilde{F})^N \to \mathbb{R} \cup \{ \infty \} \), defined by

\[
G_m(\nu) \overset{\text{def}}{=} \mathbb{E}_{F_{m(N)}} \left[ g(X_{1:N}^{F_m}, ..., X_{N:N}^{F_m}, \nu) \right] : \nu \in V_N
\]

\[\infty \quad : \text{otherwise} \]

This proves the first part of the Theorem D.5.
In the case of \( z > 0 \), the mapping \( \tilde{K}_\pi \mid [z, \bar{z}] \) is bounded, because as a result of section 4 we may describe \( \tilde{K}_\pi \) by \( \tilde{K}_\pi = \sum_{i=1}^{k} 1_{[y_{i-1}, y_i]} w_i \), where \( 0 = y_0 < y_1 < \ldots < y_k = \infty \), and \( w_1, \ldots, w_k \) denote decreasing real-valued mappings on \([0, \infty[\). This implies that \( g(\cdot, v) \) is bounded too for every \( v \in V \).

Moreover, \( [X_{F,m}^N = x] \) is a null set for any distribution \( F \) whose distribution function is continuous at \( x \). Hence, denoting the set of continuity points of \( \tilde{K}_\pi \) by \( C(\tilde{K}_\pi)^N \), we have \( F_N(\mathcal{Z}_N \times C(\tilde{K}_\pi)^N) = 1 \) for every \( F \in \{F_m, F^* \mid m \in \mathbb{N} \} \) because by assumption \( \tilde{K}_\pi \) is not continuous at finitely many points only. Therefore in view of Lemma D.3 we might draw on a well known result concerning uniform convergence w.r.t. the topologies of weak convergence (cf. Billingsley and Topsoe (1967), Theorem 1) to conclude

\[
\left| \inf_{v \in V^N} \mathbb{E}_{F_{m,N}} \left[ g(X_{1:N,m}^N, \ldots, X_{N:N,m}^N, v) \right] - \inf_{v \in V^N} \mathbb{E}_{F_{N}} \left[ g(X_{1:N}^{F^*}, \ldots, X_{N:N}^{F^*}, v) \right] \right| \\
\leq \sup_{v \in V^N} \left| \mathbb{E}_{F_{m,N}} \left[ g(X_{1:N,m}^N, \ldots, X_{N:N,m}^N, v) \right] - \mathbb{E}_{F_{N}^{F^*}} \left[ g(X_{1:N}^{F^*}, \ldots, X_{N:N}^{F^*}, v) \right] \right| \to 0.
\]

The proof is complete. 

\[\square\]

References


