Inference on Risk Neutral Measures for Incomplete Markets*

Hiroaki Kaido †
Department of Economics
University of California, San Diego

Halbert White
Department of Economics
University of California, San Diego

This version: August 20, 2008

Abstract

This paper proposes an econometric framework to estimate market risk prices associated with risk neutral measures Q under incomplete markets. We show that, under incomplete markets, the market price of risk is not point-identified but is instead identified as a bounded subset of an affine subspace. On the other hand, a structural assumption fully identifies diffusion coefficients for the data generating probability measure P. We apply Kaido and White's (2008) two-stage extension of Romano and Shaikh's (2006) and Chernuzhukov, Hong, and Tamer's (2007) *partial identification* framework to construct a set estimator and confidence regions for the identified set of market risk prices and to test hypotheses. We apply our results to study international risk sharing and risk premia for market cap range indexes.

*We thank Yacine Aït-Sahalia, Caio Almeida, Peter Carr, Richard Carson, Graham Elliott, Stephen Figlewski, Wolfgang Härdle, Mark Henry, Han Hong, Ivana Komunjer, Nour Meddahi, Andres Santos, Ross Starr, Yixiao Sun, Allan Timmermann, Rossen Valkanov, seminar participants at UCSD's GSRS, Econometrics reading group, and Rady School of Management, and participants at the 2008 SoFiE Conference for their helpful comments and suggestions.

†E-mail: hkaido@ucsd.edu. Kaido gratefully acknowledges financial support from the Nakajima Foundation.

1 Introduction

For a continuous time, continuous state model, Harrison and Kreps (1979) have shown the equivalence between the absence of arbitrage and the existence of Q, the risk neutral probability measure (henceforth RNP). Subsequently, Delbaen and Schachermayer (1994) have established the same result, by introducing a more precise notion of absence of arbitrage called "no free lunch with vanishing risk." This result is known as the first fundamental theorem of financial economics (1st FTFE). In general, in the absence of arbitrage, the price of a financial asset can be computed simply as the expected value of its payoffs under the risk neutral probability, discounted by the risk-free rate. By comparing the RNP Q to the actual data-generating probability measure (DGP) P, one can recover agents' attitude toward risk.

Another object closely related to the RNP is the stochastic discount factor (SDF), also known as the pricing kernel. Harrison and Kreps (1979) and Harrison and Pliska (1981) show that the existence of the SDF is also equivalent to the absence of arbitrage. Further, they show that the uniqueness of the RNP (equivalently SDF) is equivalent to market completeness. This is known as the second fundamental theorem of financial economics (2nd FTFE).

There is a rich literature on the estimation of the SDF. The SDF depends generally on the state variables driving asset prices. Financial economists and macroeconomists have shown that a specific functional form for the SDF can be derived from the equilibrium prices generated by rational economic agents for assets with given payoff streams. Well known examples are the CAPM studied by Sharpe (1964) and Lintner (1965) and the consumption CAPM studied by Lucas (1978) and Breeden (1979). The state variables determining the SDF in these examples are the tangent portfolio return and aggregate consumption. The standard approach is to estimate parameters associated with this function and test whether the estimated SDF can price assets correctly or not. One drawback of this approach is the requirement of observable state variables. If the state variables are measured only poorly, this directly affects the bias and precision of estimators and the level and power of tests. Further, the functional form implied by the economic model need not be correctly specified; misspecification has similar adverse effects.

Recent studies (e.g., Aït-Sahalia and Lo, 1998; Chernov and Ghysels, 2000; and Rosenberg and Engle, 2002) show that one can estimate the RNP using only asset prices. These are usually measured very precisely. Further, high frequency data are often available. These rich data sets make possible the use of nonparametric techniques that can avoid the potential misspecification problem. So far, the literature has focused on estimating on a single risk neutral probability measure. One way to justify this is to assume market completeness. In this case, the risk neutral measure is unique; one can estimate it by relying on Girsanov's theorem. Aït-Sahalia and Lo (1998) assume market completeness and nonparametrically estimate the RNP density. Chernov and Ghysels (2000) also assume completeness and propose a method to estimate parameters associated with the RNP and the actual DGP jointly, using a time series of asset returns and option prices. A different approach is taken by Rosenberg and Engle (2002). They do not assume completeness, but estimate a unique RNP closest to the DGP in terms of a certain metric.

When markets are incomplete, there exists a set $\mathscr{Q}_I$ of RNPs identified by the distribution of observed asset prices. $\mathscr{Q}_I$ is identified in the sense that any of its elements generates the same distribution of observed asset prices. That is, there are multiple observationally equivalent economic structures Q. In this case, the economic structure is only *partially identified* by the observed data. The study of partial identification was pioneered by Charles Manski; see, e.g., Manski (2003). In this paper, we contribute to the finance literature by applying the techniques of partial identification to develop methods of estimation and inference for the set of RNPs $\mathscr{Q}_I$ identified by a given vector of asset prices without assuming market completeness.

Our specific focus here is on the vector of time t market prices of risk, λ_t, a key element of the Girsanov transformation. In the absence of arbitrage, λ_t exists but is not uniquely identified by the asset price process when markets are incomplete. Instead, λ_t belongs to an identified set $\Lambda_{I,t}$ associated with $\mathscr{Q}_I$. By further imposing a bound on $\|\lambda_t\| := (\lambda_t'\lambda_t)^{1/2}$, we obtain an identified set denoted $\Lambda_{I,t}^M$. We then show that $\Lambda_{I,t}^M$ can be represented in terms of a set of minimizers of a certain criterion function. This enables us to apply the extremum set-estimation approach of Romano and Shaikh (2006) and Chernozhukov, Hong, and Tamer (2007) (henceforth the RS-CHT framework) to construct a set estimator and a confidence region for $\Lambda_{I,t}^M$ and to conduct hypothesis tests. In this application, we apply a two-stage procedure introduced by Kaido and White (2008) that helps to reduce the dimension of the associated set-valued estimators. To the best of our knowledge, this is the first application of such a procedure.

For concreteness, we pay particular attention to the case in which a standard multivariate geometric Brownian motion determines the evolution of asset prices. In this case, $\lambda_t = \lambda_0$, a nonrandom and time-invariant vector.

The paper is organized as follows. Section 2 specifies the asset price data generating process. In Section 3, we discuss the identification of the market price of risk. Section 4 sets forth our econometric framework for estimation and inference. In Section 5, we apply our results to study international risk sharing and risk premia associated with market capitalization range indexes. Section 6 discusses extensions of our framework to more general multivariate asset price processes. Section 7 concludes with a summary and a discussion of directions for future research.

2 The Asset Price Process

For given positive finite T, let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in[0,T]}, P)$ be a complete filtered probability space. The filtration $\{\mathscr{F}_t\} = \{\mathscr{F}_t\}_{t\in[0,T]}$ is assumed to satisfy the usual properties (e.g., Protter, 2005). Unless otherwise noted, $t \in [0,T]$ throughout. As is common, we take $\mathscr{F} = \mathscr{F}_T$. Suppose that there are $d \in \mathbb{N}$ risky assets and that the $\mathbb{R}^d-$ valued asset price process $\{S_t\}$ solves the stochastic differential equation

$$dS_t = \mu_{0t}dt + \sigma_{0t}dW_t, \quad t \in [0,T],$$

where $\{W_t\}$ is a vector of $n \in \mathbb{N}$ independent standard Brownian motions under P adapted to the filtration $\{\mathscr{F}_t\}$, $\{\mu_{0t}\}$ is an $\mathbb{R}^d-$ valued adapted drift process, and $\{\sigma_{0t}\}$ is an $\mathbb{R}^{d\times n}-$ valued adapted diffusion coefficient process. We assume without loss of generality that S_t^0 is the price of the risk-free bond with known rate of return r. Let the discounted asset prices be $S_t^{*i} = S_t^i/S_t^0, i = 1,...,d$. This setup is essentially that of Williams (2006).

Given an $\mathbb{R}^n-$ valued adapted process $\{\lambda_t\}$ such that $\int_0^T \|\lambda_t\|^2 dt < \infty$, $a.s.-P$, the Girsanov transformation defines a new adapted process $\{\tilde{W}_t\}$ by adjusting the drift of the original Brownian motion:

$$\tilde{W}_t = W_t + \int_0^t \lambda_s ds.$$

The absence of arbitrage (equivalently, the existence of the risk neutral measure) holds only for λ_t such that

$$\sigma_{0t}\lambda_t = \mu_{0t} - rS_t, \quad t \in [0,T], \quad a.s.-P. \tag{1}$$

Such a vector λ_t is called a *market price of risk*. Without further assumptions, and specifically without assuming market completeness, the market prices of risk form a set

$$\Lambda_{I,t} := \{\lambda_t : \sigma_{0t}\lambda_t = \mu_{0t} - rS_t\}.$$

We let $\mathbf{\Lambda}_I$ denote the set-valued process $\{\Lambda_{I,t}, t \in [0,T]\}$.

For our purposes here, it suffices to define market completeness in terms of $\Lambda_{I,t}$. We say that markets are *complete at t* when $\Lambda_{I,t}$ has a unique element; otherwise, we say markets are *incomplete at t*.

Under a risk neutral measure Q, $\tilde{W}_t$ follows a standard Brownian motion. After the change of measure from P to Q, the discounted asset return process can be represented by linear combinations of Brownian motions under Q:

$$\frac{dS_t^{*i}}{S_t^{*i}} = \sigma_{0t}^i \cdot d\tilde{W}_t \quad i = 1,...,d,$$

where σ_{0t}^i is the $1 \times n$ ith row of σ_{0t}. That is, under Q, any asset is expected to earn a return equal to the risk-free rate. Using this result, the prices of redundant securities can be computed by taking the expectation under Q. (See, for example, Duffie, 2001, and Williams, 2006.)

In order to study identification in a simple but important special case in what follows, we consider a running example in which $\mu_{0t}^i = \mu_0^i S_t^i$ and $\sigma_{0t}^i = \sigma_0^i S_t^i$ for $i = 1, \ldots, d$. We call this specification a *multivariate Black-Scholes economy*. We formalize this as follows.

Assumption 1 (Multivariate Black-Scholes) *Let $\{W_t\}$ be a vector of $n \in \mathbb{N}$ independent standard Brownian motions under P adapted to the filtration $\{\mathscr{F}_t\}$. Let $\{S_t\}$ be a vector of $d \in \mathbb{N}$ asset prices such that $S_0^i = 1$ and solving the stochastic differential equations*

$$dS_t^i = \mu_0^i S_t^i dt + \sigma_0^i S_t^i dW_t, \quad t \in [0, T], \quad i = 1, \ldots, d,$$

where $\mu_0 \in \mathbb{R}^d$ has elements μ_0^i, $i = 1, \ldots, d$, and $\sigma_0 \in \mathbb{R}^{d \times n}$ has $1 \times n$ rows σ_0^i, $i = 1, \ldots, d$. Further, $\{S_t\}$ does not admit arbitrage.

For this process, the market prices of risk always lie in the non-random time-invariant set

$$\Lambda_{I,0} = \{\lambda : \sigma_0 \lambda = \mu_0 - r\iota\},$$

where ι is a d-dimensional vector of ones. Any process $\{\lambda_t\}$ such that $\lambda_t \in \Lambda_{I,0}$, $t \in [0, T]$, is an admissible market price of risk process in this economy. As we know that the true Black-Scholes market price of risk is a constant, say λ_0, we consider only non-random, time-invariant processes $\{\lambda_t\}$ such that for all $t \in [0, T]$, $\lambda_t = \lambda$ for some fixed $\lambda \in \Lambda_{I,0}$.

3 Identifying the Market Price of Risk

3.1 The market price of risk and the RNP

Under the change of measure from the objective measure P to the risk neutral measure Q, the risk adjustment is fully determined by the Radon-Nikodym derivative[1] dQ/dP. In the continuous-time setting, one can define a density process of Radon-Nikodym derivatives $\xi := \{\xi_t\}$ where $\xi_t := E_t[dQ/dP]$, with $E_t(\cdot) := E(\cdot \mid \mathscr{F}_t)$. As dQ/dP is $\mathscr{F} = \mathscr{F}_T-$ measurable, we have $\xi_T = dQ/dP$. The history $\lambda^t := \{\lambda_\tau, \tau \in [0, t]\}$ uniquely indexes the density ξ_t and therefore characterizes the

[1]The Radon-Nikodym derivative $\mathscr{D}$ is defined as an $\mathscr{F}$-measurable strictly positive random variable such that for any $A \in \mathscr{F}$, $Q(A) = \int_A \mathscr{D} dP$.

risk adjustment. In general, given an adapted process $\{\lambda_t\}$, the corresponding densities can be written

$$\xi_t = \exp\left(-\int_0^t \lambda_s \cdot dW_s - \frac{1}{2}\int_0^t \|\lambda_s\|^2 ds\right) \quad t \in [0,T]. \tag{2}$$

Accordingly, ξ is also known as the *stochastic exponential* of $\{-\lambda_t\}$. In the multivariate Black-Scholes economy, ξ_t simplifies to

$$\xi_t = \exp\left(-\lambda_0 \cdot W_t - \frac{1}{2}\|\lambda_0\|^2 t\right), \tag{3}$$

where λ_0 is the true market price of risk in the Black-Scholes economy.

In general, there are multiple processes $\{\lambda_t\}$ consistent with the no-arbitrage requirement. This implies that there are multiple ways to change the measure from P to Q. Therefore, even if P is identified by the observed data, Q cannot be uniquely identified under incomplete markets.

The role of the Radon-Nikodym derivative is best understood in terms of the pricing equation. Consider a "European-type" asset paying zero for $t < T$ and $\varphi(W_T)$ in period T, where φ is a Borel measurable real-valued function. Let $\varphi_\lambda : \mathbb{R}^n \to \mathbb{R}$ be a measurable function such that

$$\varphi_\lambda(\tilde{W}_T) = \varphi_\lambda\left(W_T + \int_0^T \lambda_s ds\right) = \varphi(W_T).$$

For example, a contingent claim that pays 1 monetary unit if W_T is in a measurable set A and zero otherwise has a payoff $\varphi(W_T) = 1_{\{W_T \in A\}}$, where $1_{\{\cdot\}}$ is the indicator function taking the value 1 if the condition in brackets $\{\cdot\}$ is true and 0 otherwise. Then the payoff function in terms of $\tilde{W}_T$ is $\varphi_\lambda(\tilde{W}_T) = 1_{\{\tilde{W}_T \in A_\lambda\}}$, where A_λ is a translation of A by $\int_0^T \lambda_s ds$.

Generally, there are two equivalent ways to compute the asset price p_0 at $t = 0$ for such an asset[2]. We have

$$p_0 = E^P\left[m_T\ \varphi(W_T)\right] = e^{-rT} E^Q[\varphi_\lambda(\tilde{W}_T)].$$

The first equality uses the DGP P and the $\mathscr{F}_T$-measurable *stochastic discount factor* (SDF) m_T. The second equality uses the RNP Q and the risk-free rate to price the payoff $\varphi_\lambda(\tilde{W}_T)$. To represent the SDF, we write

$$\begin{aligned}
E^P\left[m_T\ \varphi(W_T)\right] &= e^{-rT} E^Q[\varphi_\lambda(\tilde{W}_T)] = e^{-rT}\int \varphi_\lambda(\tilde{W}_T) dQ \\
&= \int e^{-rT} \varphi(W_T) \frac{dQ}{dP} dP = \int e^{-rT} \xi_T\ \varphi(W_T) dP \\
&= E^P\left[e^{-rT} \xi_T\ \varphi(W_T)\right].
\end{aligned}$$

[2] See Aït-Sahalia and Lo (2000) or chapter 6 of Duffie (2001), for example.

Thus, $m_T = e^{-rT}\xi_T$, the discounted Radon-Nikodym derivative. The SDF discounts the future payoff by e^{-rT} and adjusts its risk by ξ_T. If $\lambda_t = 0$ for all t, then $\xi_T = 1$ and no risk adjustment takes place. This is the case of risk neutrality. For the multivariate Black-Scholes economy, ξ_T is a log-normal random variable with mean 1 and variance $e^{\|\lambda_0\|^2 T} - 1$; in this case, risk neutrality is equivalent to $\lambda_0 = 0$.

The density process ξ of Radon-Nikodym derivatives is a stochastic process defined by the stochastic integral in (2). For what follows, we will take the variance of ξ_T to be finite. This condition is known as the *L^2-reducibility of* $\{\lambda_t\}$ (see, e.g., Duffie, 2001). Further, this finiteness assumption has a portfolio interpretation and a link to the option pricing bound studied in Cochrane and Saá-Requejo (2000). To ensure that ξ_T has finite variance in the Black-Scholes economy, we simply bound[3] λ_0 :

Assumption 2 (Bounded Risk Price) *For the Black-Scholes economy, there exists* $0 < M < \infty$ *such that* $\|\lambda_0\| \leq M$.

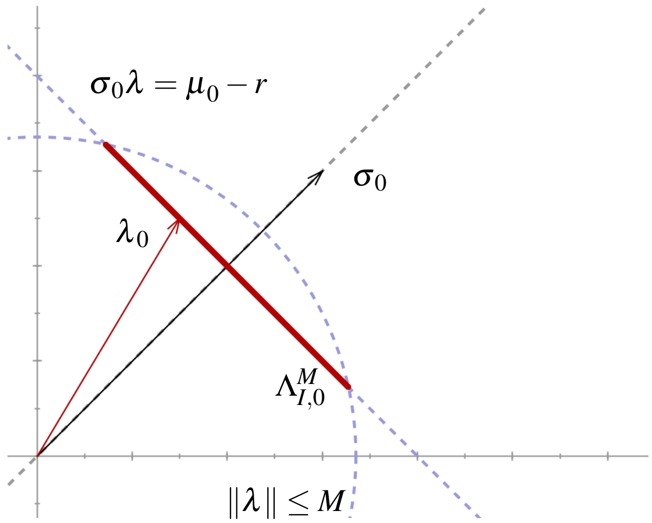
For the Black-Scholes economy, the identified set for the market price of risk is thus

$$\Lambda_{I,0}^M := \{\lambda : \sigma_0 \lambda = \mu_0 - r\iota, \quad \|\lambda\| \leq M\}.$$

An illustration of the identified set $\Lambda_{I,0}^M$ with $d = 1$ and $n = 2$ is given by Figure 1. In this example, the risk exposure of the single traded asset is determined by a vector $\sigma_0 \in \mathbb{R}^2$ such that both elements of σ_0 are non-zero; and $\lambda_0 \in \mathbb{R}^2$ is the true market price of risk. As there are two fundamental sources of risk in this economy, the traded security does not reveal λ_0. Instead it reveals all λ's that lie on the iso-risk premium line perpendicular to σ_0, as λ_0 is observationally equivalent to any other λ on the iso-risk premium line. To see this, fix σ_0, and let $N(\sigma_0)$ be the null space of σ_0. Consider $\lambda := \lambda_0 + b\eta$, where $b \in \mathbb{R}$ and $\eta \in N(\sigma_0)$. Then, λ_0 and λ give the same value of the risk premium by construction. As the joint distribution of the $d = 1$ discounted asset prices is fully characterized by the drift (risk premium) and the variance-covariance structure, λ_0 and λ are observationally equivalent. Thus, one cannot identify λ_0 by simply examining the distribution of asset prices. Instead, this distribution only reveals the iso-risk premium line. Together with L^2-boundedness, the identified set becomes a finite line segment. In more general cases, the identified set is a finite subset of an affine subspace orthogonal to the row space of σ_0.

[3]This boundedness also implies that the Radon-Nikodym derivative satisfies the Novikov condition. See Duffie (2001) and Williams (2006). We discuss a more general condition in Section 6.

Figure 1: The identified set ($d = 1$ and $n = 2$)



3.2 A common factor structure

When σ_0 satisfies m a priori restrictions $\rho(\sigma_0) = 0$, these can facilitate estimation of the market price of risk and may even suffice to identify σ_0. A leading case of such restrictions is that of common factors. For the Black-Scholes case, we impose this formally as follows.

Assumption 3 (Common Factors) *Each asset return depends on its unique idiosyncratic risk and $n-d$ common factors.*

Assumption 3 implies that the returns are correlated with each other only through common risk factors. The $n-d$ common risk factors are represented by $n-d$ independent Brownian motions. An example with $d = 3$ and $n = 4$ is

$$\sigma_0 = \begin{bmatrix} \sigma_{11} & 0 & 0 & \sigma_{14} \\ 0 & \sigma_{22} & 0 & \sigma_{24} \\ 0 & 0 & \sigma_{33} & \sigma_{34} \end{bmatrix}.$$

With common factors, each asset return depends only on $(n-d+1)$ Brownian motions. This reduces the number of nonzero parameters in σ_0 from nd in the general case to $(n-d+1)d$.

In general, the $d \times d$ asset returns covariance matrix $\Sigma_0 := \sigma_0\sigma_0'$ provides $d(d+1)/2$ restrictions on σ_0. The zero restrictions of Assumption 3 impose $(d-1)d$ further restrictions. This reduces the dimension of the unidentified aspects of σ_0. Specifically, when $n \leq (3d-1)/2$, Assumption 3 ensures that one can fully identify σ_0 from elements of Σ_0. Our examples in Section 5 take $d = 3$ and $n = 4$, a case in which σ_0 is fully identified.

We emphasize that this is inherently a *structural* restriction; that is, the data are generated by a process obeying this condition. Although alternative representations of the asset price process may exist that do not obey this restriction, these have only a stochastic and not a structural interpretation. Assumption 3 thus specifies an economic interpretation for the vector of Brownian motions. We interpret the first d elements as idiosyncratic risks and the last $n-d$ elements as common risks. As Section 5 illustrates, the meaning of W_t may vary depending on the given application.

Above, we represented the restrictions on σ_0 as $\rho(\sigma_0) = 0$, an $m \times 1$ zero vector. The common factors assumption has a simple representation of this form. For these cases, we have

$$\rho(\sigma_0) = \rho_0 vec(\sigma_0),$$

where $vec(\sigma_0)$ stacks the columns to yield an $nd \times 1$ column vector, and ρ_0 is an $m \times nd$ matrix. The matrix ρ_0 has rows whose elements are zero, except for a one in the position that identifies an element of σ_0 that is to take the value zero. For the example above with $n = 4$ and $d = 3$, $m = 6$. Further, the third row of the 6×12 matrix ρ_0 contains a one in the fourth position (corresponding to σ_{12}, which is set to zero), with the remaining row elements zero.

4 Econometric Framework

In this section, we propose estimation and hypothesis testing procedures for the market price of risk in the Black-Scholes economy following the set-estimation and hypothesis testing frameworks of Romano and Shaikh (2006) and Chernozhukov, Hong, and Tamer (2007). These authors study extremum estimators where the criterion functions do not have a unique minimizer. For estimation, the basic idea is to use lower contour sets of the sample criterion function as set-valued estimators or confidence regions. For hypothesis testing, Romano and Shaikh (2006) propose a subsampling procedure for a likelihood ratio statistic. In this section, we exploit these methods by showing that the identified set for the Black-Scholes economy risk price, $\Lambda^M_{I,0}$, can be characterized as a set of minimizers of a specific criterion function.

4.1 Applying the RS-CHT framework

In the multivariate Black-Scholes economy, the vector of returns of d securities over the time interval $[s,t]$ obeys a multivariate normal distribution with mean $(t-s)(\mu_0 - (\|\sigma_0^1\|^2, \cdots, \|\sigma_0^d\|^2)'/2)$ and covariance matrix $(t-s)\Sigma_0$. Eq. (1) implies that the drift μ_0 is determined, once we specify (σ_0, λ_0) and r. Therefore, for any given constant r, the joint density of asset returns depends only on σ_0 and λ_0.

Consider a partition $\pi := \{0 =: t_0, t_1, ..., t_{N-1}, t_N := T\}$ of the interval $[0,T]$. Suppose we observe a series of asset prices $\{S_{t_j}\}_{j=0}^N$ over this partition. Let R_{t_j} be the $d \times 1$ vector of asset returns from period t_j to t_{j+1}: i.e., the ith element of R_{t_j} is $R^i_{t_j} := \ln S^i_{t_j} - \ln S^i_{t_{j-1}}$, $i = 1, ..., d$. Let $f(R_{t_1}, ..., R_{t_N}; \theta)$ denote the likelihood of a sample of asset returns at $\theta := (\sigma, \lambda) \in \Theta := \mathbb{S} \times \Lambda \subseteq \mathbb{R}^{d\times n} \times \mathbb{R}^n$, where $\mathbb{S}$ is a non-empty subset of $\mathbb{R}^{d\times n}$. In the multivariate Black-Scholes economy, returns are independent over time, so that

$$f(R_{t_1}, \cdots, R_{t_N}; \theta) = \prod_{j=1}^N f(R_{t_j}; \theta),$$

where $f(R_{t_j}; \theta)$ defines the likelihood for asset returns in t_j; this is a $d-$variate normal likelihood.

The coefficients $\theta_0 := (\sigma_0, \lambda_0) \in \Theta$ index the true DGP measure P. Let the *criterion function* $\bar{Q}_N : \Theta \to \bar{\mathbb{R}}_+$ be the shifted expected negative average log likelihood defined by

$$\bar{Q}_N(\theta) := E^P\left[-N^{-1}\sum_{j=1}^N \ln f(R_{t_j}; \theta)\right] - q_{0,N}, \tag{4}$$

where

$$q_{0,N} := E^P\left[-N^{-1}\sum_{j=1}^N \ln f(R_{t_j}; \theta_0)\right].$$

The criterion function thus has minimum value 0 at θ_0. This minimum is not unique; letting $\Theta_0^M := \{\sigma_0\} \times \Lambda^M_{I,0}$, we also have

$$\bar{Q}_N(\Theta_0^M) = 0.$$

Further, the asset return covariances only reveal $\Sigma_0 = \sigma_0 \sigma_0'$, so σ_0 cannot be identified from observations of d asset returns without further restrictions. Specifically, let

$$\Theta^M_{I,0} := \{(\sigma, \lambda) \in \Theta : \sigma\sigma' = \Sigma_0,\ \ \sigma\lambda = \mu_0 - r\iota,\ \ \|\lambda\| \leq M\}.$$

Then $\Theta_0^M \subset \Theta^M_{I,0}$, and $\Theta^M_{I,0}$ contains all the minimizers of $\bar{Q}_N$. That is,

$$\bar{Q}_N(\Theta^M_{I,0}) = 0 \quad \text{and} \quad \bar{Q}_N(\theta) > 0 \quad \text{for } \theta \notin \Theta^M_{I,0}.$$

Working with $\bar{Q}_N$ and $\Theta^M_{I,0}$ enables us to apply the RS-CHT framework to our problem.

Accordingly, let $Q_N : \Omega \times \Theta \to \bar{\mathbb{R}}_+$ be the sample criterion function defined by

$$Q_N(\theta) = -N^{-1}\sum_{j=1}^N \ln f(R_{t_j}; \theta) - q_N, \tag{5}$$

where $q_N = \inf_\Theta -N^{-1}\sum_{j=1}^N \ln f(R_{t_j}; \theta)$. Following Romano and Shaikh (2006) and Chernozhukov, Hong, and Tamer (2007), we define an ε-level set of the sample criterion function by

$$\hat{\Theta}_N(\varepsilon) := \{\theta \in \Theta : N \cdot Q_N(\theta) \leq \varepsilon\}.$$

When we choose ε properly, the random set $\hat{\Theta}_N(\varepsilon)$ is a consistent set estimator or a confidence region for the identified set.

There are, however, several challenges to directly applying the RS-CHT framework to our problem. First, the identified set $\Theta^M_{I,0}$ has a high dimension, $nd+n$. This leads to computational difficulties and can also hamper the interpretation of results. Further, for any fixed value of σ such that $\sigma\sigma' = \Sigma_0$, every element of the set $\Lambda^M_{I,0}(\sigma) := \{\lambda \in \Lambda : \sigma\lambda = \mu_0 - r\iota,\ \|\lambda\| \le M\}$ minimizes $\bar{Q}_N$. This suggests that, as one changes the value of σ, the set $\Lambda^M_{I,0}(\sigma)$ rotates. Consequently, $\Theta^M_{I,0}$ may cover quite a large subset of Θ. Finally, $\Theta^M_{I,0}$ need not be convex. This may cause additional technical difficulties.

These difficulties can be mitigated or avoided by applying a two-stage procedure proposed by Kaido and White (2008), described next.

4.2 A two-stage procedure

In this section, we describe a two-stage procedure proposed by Kaido and White (2008) that reduces the dimension of the set estimator and the associated confidence region. With sufficient restrictions, some elements of σ_0 can even be fully identified. In such cases, we can replace identified elements of σ_0 in the sample criterion function with their consistent estimators. Even if this is not possible, restrictions on elements of σ_0 can still simplify estimation substantially.

We summarize Kaido and White's (2008) measurability and consistency results for the two-stage set estimator as follows. Let $m \in \mathbb{N}$ be the number of restrictions on σ_0 and let $\rho : \mathbb{R}^{d\times n} \to \mathbb{R}^m$ embody these restrictions as

$$\rho(\sigma_0) = 0.$$

The identified (σ, λ) values that satisfy all our restrictions are the elements of

$$\Theta^M_{I,0,\rho} := \{(\sigma,\lambda) \in \Theta : \sigma\sigma' = \Sigma_0,\ \ \rho(\sigma) = 0,\ \ \sigma\lambda = \mu_0 - r\iota,\ \ \|\lambda\| \le M\}.$$

Let $\hat{\Sigma}_N$ be a bounded consistent estimator of Σ_0, and let $K(\mathbb{S})$ be a collection of closed subsets of $\mathbb{S}$. Define a first-stage restricted set-estimator $\hat{S}_N : \Omega \to K(\mathbb{S})$ of σ_0 by

$$\hat{S}_N(\omega) = \{\sigma \in \mathbb{S} : \sigma\sigma' = \hat{\Sigma}_N(\omega),\ \ \rho(\sigma) = 0\}. \tag{6}$$

This is a random set of diffusion coefficients that are consistent with the sample covariance of the returns and that satisfy the restriction $\rho(\sigma) = 0$.

Using this first-stage set estimator, let the second-stage set-estimator for $\Theta^M_{I,0,\rho}$ be defined by

$$\hat{\Theta}_N(\omega) := \{(\sigma,\lambda) \in \Theta : NQ_N(\omega,\sigma,\lambda) \le \hat{\varepsilon}(\omega),\ \ \sigma \in \hat{S}_N(\omega)\}, \tag{7}$$

where $\hat{\varepsilon}$ is now permitted to be random.

An important special case occurs when the restrictions suffice to identify σ_0. When $\hat{\sigma}_N$ is a consistent estimator of σ_0, the first-stage set estimator becomes a singleton, i.e., $\hat{S}_N = \{\hat{\sigma}_N\}$. The second-stage set estimator is then $\hat{\Theta}_N = \{\hat{\sigma}_N\} \times \hat{\Lambda}_N$, where

$$\hat{\Lambda}_N(\omega) := \{\lambda \in \Lambda : NQ_N(\omega, \hat{\sigma}_N(\omega)), \lambda) \leq \hat{\varepsilon}(\omega)\}. \tag{8}$$

4.3 Effros-measurability

The first step in analyzing the two-stage set estimator is to establish its measurability. A useful measurability concept for set-valued functions is *Effros*-measurability. Effros-measurability ensures that many functionals of interest, such as the distance between random sets, become random variables; it is also flexible, handling as many random elements as one typically requires. See Molchanov (2005) for details.

Definition 1 (Effros-Measurability) *Let $(\Omega, \mathscr{F})$ be a measurable space. Let $l \in \mathbb{N}$, and let $\mathscr{G}$ be a topology on $\mathbb{R}^l$. Let $K(\mathbb{R}^l)$ be a collection of closed subsets of $\mathbb{R}^l$. A map $X : \Omega \to K(\mathbb{R}^l)$ is* Effros-measurable *with respect to $\mathscr{F}$ if, for each open set $G \in \mathscr{G}$,*

$$X^-(G) := \{\omega : X(\omega) \cap G \neq \emptyset\} \in \mathscr{F}.$$

The next result establishes Effros-measurability for general two-stage estimators.

Theorem 2 *Let $(\Omega, \mathscr{F}, P)$ be a complete probability space, and let $\Theta = \mathbb{S} \times \Lambda$, where $\mathbb{S}$ and Λ are compact subsets of finite dimensional Euclidian spaces.*

Let $Q : \Omega \times \Theta \to \bar{\mathbb{R}}_+$ be such that for each $\theta \in \Theta$, $Q(\,\cdot\,, \theta)$ is measurable-$\mathscr{F}$ and for $F \in \mathscr{F}$ with $P(F) = 1$, $Q(\omega, \cdot\,)$ is continuous on Θ for each $\omega \in F$.

Let $\hat{S} : \Omega \to K(\mathbb{S})$ be Effros-measurable with respect to $\mathscr{F}$.

Then for any measurable $\hat{\varepsilon} : \Omega \to \mathbb{R}^+$, the $\hat{\varepsilon}$- level set $\hat{\Theta}_{\hat{\varepsilon}} : \Omega \to K(\mathbb{S} \times \Lambda)$, defined by

$$\hat{\Theta}(\omega, \hat{\varepsilon}(\omega)) = \{(\sigma, \lambda) \in \Theta : Q(\omega, \sigma, \lambda) \leq \hat{\varepsilon}(\omega), \;\; \sigma \in \hat{S}(\omega)\},$$

is Effros-measurable with respect to $\mathscr{F}$.

The proof of this and other formal results can be found in Kaido and White (2008).

For the special case where $\hat{S}$ is a singleton, e.g., when the diffusion coefficient is point identified, we have the following result.

Corollary 3 *Let the conditions of Theorem (2) hold, and suppose $\hat{S}$ is a singleton such that $\hat{S} = \{\hat{\sigma}\}$, where $\hat{\sigma} : \Omega \to \mathbb{S}$ is measurable-$\mathscr{F}$. Then*

(i) For each $\lambda \in \Lambda$, $\tilde{Q}(\,\cdot\,,\lambda) := Q(\,\cdot\,,\hat{\sigma}(\cdot),\lambda)$ is a measurable function on Ω and for $\tilde{F} \in \mathscr{F}$ with $P(\tilde{F}) = 1$, $\tilde{Q}(\omega,\,\cdot\,)$ is continuous on Λ for each $\omega \in \tilde{F}$.

(ii) For any measurable $\hat{\varepsilon} : \Omega \to \mathbb{R}_+$, the $\hat{\varepsilon}$- level set $\hat{\Lambda}_{\hat{\varepsilon}} : \Omega \to K(\Lambda)$, defined by

$$\hat{\Lambda}(\omega,\hat{\varepsilon}(\omega)) = \{\lambda \in \Lambda : \tilde{Q}(\omega,\lambda) \leq \hat{\varepsilon}(\omega)\},$$

is Effros-measurable with respect to $\mathscr{F}$.

The following proposition establishes the Effros-measurability of our constrained first-stage estimator, enabling us to apply the above results.

Proposition 4 *Let $(\Omega,\mathscr{F})$ be a measurable space, and let $\mathbb{S}$ be a compact subset of $\mathbb{R}^{d\times n}$, where d and n are finite positive integers.*

Let Ψ be a set of bounded symmetric positive semi-definite matrices. Let $\hat{\Sigma} : \Omega \to \Psi$ be measurable-$\mathscr{F}$, and let $\rho : \mathbb{S} \to \mathbb{R}^m$ be continuous, where m is a finite positive integer.

Let $\hat{S} : \Omega \to K(\mathbb{S})$ be defined by

$$\hat{S}(\omega) = \{\sigma \in \mathbb{S} : \sigma\sigma' = \hat{\Sigma}(\omega),\ \ \rho(\sigma) = 0\}.$$

Then $\hat{S}$ is Effros-measurable with respect to $\mathscr{F}$.

4.4 Consistent set estimation

Next we provide results ensuring the consistency of the two-stage set estimator. Consistency is expressed in terms of the Hausdorff metric on the space of closed sets.

Definition 5 *Let Θ be a compact subset of a finite dimensional Euclidean space. For any two closed subsets A and B of Θ, the Hausdorff metric is*

$$d_H(A,B) = \max\left[\sup_{a\in A}\inf_{b\in B}\|b-a\|, \sup_{b\in B}\inf_{a\in A}\|b-a\|\right],$$

where $\|\cdot\|$ is the Euclidean norm, and $d_H(A,B) := \infty$ if either A or B is empty.

The Effros-measurability of the two-stage set estimator implies the measurability of the Hausdorff distance between the set estimator and the identified set[4]. This makes it possible to discuss the consistency of this set estimator. Our first result provides conditions under which the general two-stage set estimator is consistent.

[4]This follows from Theorem 2.25 (vi) p.37 in Molchanov (2005).

Theorem 6 *Let* $(\Omega, \mathscr{F}, P)$ *and* $\Theta = \mathbb{S} \times \Lambda$ *satisfy the conditions of Theorem (2), and suppose that for* $N = 1, 2, ...,$ Q_N *and* $\hat{S}_N$ *satisfy the conditions on* Q *and* $\hat{S}$ *imposed in Theorem (2).*

Suppose there exists $\bar{Q}_N : \Theta \to \bar{\mathbb{R}}_+$ *such that* $\sup_{\theta \in \Theta} |Q_N(\,\cdot\,, \theta) - \bar{Q}_N(\theta)| = o_p(1)$. *Let* $S \in K(\mathbb{S})$ *and define*

$$\Theta_I := \underset{(\sigma,\lambda)\in S\times\Lambda}{\arg\min}\ \bar{Q}_N(\theta),$$

such that $\bar{Q}_N(\Theta_I) = 0$ *for all N sufficiently large.*

Let $\hat{\varepsilon}$ *be* $\mathscr{F}$*-measurable such that* $\hat{\varepsilon}/N = o_p(1)$ *and*

$$\lim_{N\to\infty} P\left[\omega : \sup_{\theta\in\Theta_I} Q_N(\omega, \theta) \le \hat{\varepsilon}(\omega)/N\right] = 1.$$

Suppose further that $d_H(\hat{S}_N, S) = o_p(1)$, *and let*

$$\hat{\Theta}_N(\omega) := \{(\sigma, \lambda) \in \Theta : Q_N(\omega, \sigma, \lambda) \le \hat{\varepsilon}(\omega)/N,\ \ \sigma \in \hat{S}_N(\omega)\}.$$

Then $\hat{\Theta}_N$ *is Effros-measurable with respect to* $\mathscr{F}$, *and* $d_H(\hat{\Theta}_N, \Theta_I) = o_p(1)$.

The next result treats the important special case in which S is fully identified (i.e., S is a singleton). This shows that the natural second-stage set estimator $\hat{\Lambda}_N$ is a consistent estimator for the identified set Λ_I.

Corollary 7 *Let the conditions of Theorem (6) hold, and suppose that S is a singleton,* $S = \{\sigma_0\}$. *Let*

$$\Lambda_I := \underset{\lambda\in\Lambda}{\arg\min}\, \bar{Q}_N(\sigma_0, \lambda).$$

Let $\hat{\sigma}_N : \Omega \to \mathbb{S}$ *be measurable-*$\mathscr{F}$ *such that* $\hat{\sigma}_N = \sigma_0 + o_p(1)$, *and let*

$$\hat{\Lambda}_N(\omega) := \{\lambda \in \Lambda : Q_N(\omega, \hat{\sigma}_N(\omega), \lambda) \le \hat{\varepsilon}(\omega)/N\}.$$

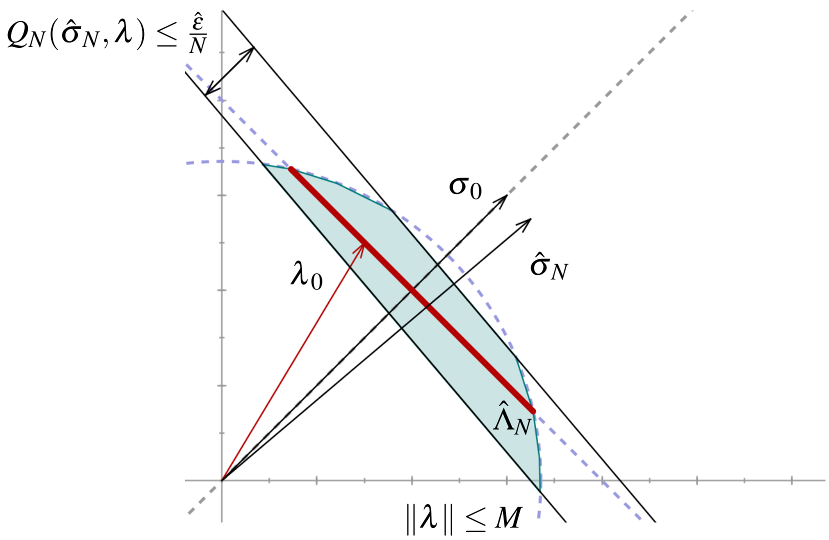
Then $\hat{\Lambda}_N$ *is Effros-measurable with respect to* $\mathscr{F}$, *and* $d_H(\hat{\Lambda}_N, \Lambda_I) = o_p(1)$.

Figure 2 illustrates. As the sample size N increases, the set estimator $\hat{\Lambda}_N$ represented by the shaded region in figure 2 shrinks down to the identified set Λ_I, which is a line segment here.

4.5 Hypothesis testing

Set estimation is useful when interest attaches to the characteristics of the identified set. If instead one wishes to test hypotheses regarding the identified set, it is not necessary to estimate the identified set. Specifically, let R be a closed subset of Θ (or Λ), where R is a set of parameters that

Figure 2: Set Estimator $\hat{\Lambda}_N$ ($d = 1$ and $n = 2$)



satisfy the restrictions of interest. For example, R may represent a set of market prices of risk that are consistent with risk-neutrality or international risk sharing.

As the true coefficient value θ_0 is in the identified set, if θ_0 also satisfies the restrictions, the identified set Θ_I has a non-empty intersection with R. We thus consider hypotheses

$$H_o^{\Theta} : \Theta_I \cap R \neq \emptyset \quad vs. \quad H_A^{\Theta} : \Theta_I \cap R = \emptyset.$$

The null states that there is at least one element in the identified set satisfying the restrictions. Rejection means that none of the parameters in the identified set satisfies the restrictions, implying that θ_0 does not satisfy the restrictions.

In the Black-Scholes example, where interest attaches to λ_0, we consider hypotheses

$$H_o^{\Lambda} : \Lambda_I \cap R \neq \emptyset \quad vs. \quad H_A^{\Lambda} : \Lambda_I \cap R = \emptyset.$$

Because R is a closed subset of the compact parameter space, the hypotheses above are equivalent to

$$H_o^{\Theta} : \inf_{\theta \in \Theta \cap R} \bar{Q}_N(\theta) = 0 \qquad \text{or} \qquad H_o^{\Lambda} : \inf_{\lambda \in R} \bar{Q}_N(\sigma_0, \lambda) = 0.$$

Such hypotheses are considered in the partially identified case by Romano and Shaikh (2008) for parametric inference and by Santos (2007) for nonparametric inference.

To test these hypotheses in our two-stage framework, we replace $\bar{Q}_N$ and Θ with their sample analogs Q_N and $\hat{S}_N \times \Lambda$, which leads to the test statistics

$$\hat{T}_N(\Theta, R) = \inf_{\theta \in (\hat{S}_N \times \Lambda) \cap R} a_N Q_N(\theta) \quad \text{and} \quad \hat{T}_N(\Lambda, R) = \inf_{\lambda \in R} a_N Q_N(\hat{\sigma}_N, \lambda),$$

where a_N is a normalizing constant such that $\sup_{\theta \in \Theta_I} Q_N(\theta) = O_p(1/a_N)$ or $\sup_{\lambda \in \Lambda_I} Q_N(\sigma_0, \lambda) = O_p(1/a_N)$. In our problem, $a_N = N$, so the test statistics can be written

$$\begin{aligned}
\hat{T}_N(\Theta, R) &= \sup_{\theta \in \hat{S}_N \times \Lambda} \sum_{j=1}^{N} \ln f(R_{t_j}; \theta) - \sup_{\theta \in (\hat{S}_N \times \Lambda) \cap R} \sum_{j=1}^{N} \ln f(R_{t_j}; \theta). \\
\hat{T}_N(\Lambda, R) &= \sup_{\lambda \in \Lambda} \sum_{j=1}^{N} \ln f(R_{t_j}; \hat{\sigma}_N, \lambda) - \sup_{\lambda \in R} \sum_{j=1}^{N} \ln f(R_{t_j}; \hat{\sigma}_N, \lambda).
\end{aligned}$$

These can be viewed as log-likelihood ratio statistics for partially identified models.

To maintain a tight focus for the discussion to follow, we now restrict attention to the Black-Scholes case that will be the subject of our empirical examples. This is the case where Assumptions 1 and 2 hold, and the common factor structure of Assumption 3 ensures that σ_0 is point-identified. Thus, we restrict attention to $\hat{T}_N(\Lambda, R)$, where we take $\Lambda = \Lambda^M$. (We leave the notation $\hat{T}_N(\Lambda, R)$ unchanged for simplicity.) To test H_o^Λ, we require asymptotic critical values for $\hat{T}_N(\Lambda, R)$.

Obtaining these critical values presents interesting challenges. Space precludes a rigorous derivation here, as handling all the necessary formalities is fairly involved. Nevertheless, the intuition behind our approach is straightforward, so we offer the following heuristic discussion.

We start by noting that the presence of $\hat{\sigma}_N$ in $\hat{T}_N(\Lambda, R)$ may have an impact on its limiting distribution. To accommodate this, we can proceed in a manner analogous to the fully identified case. There, one can often exploit a two-term mean value or Taylor-like expansion. The following straightforward high-level result applies when θ_0 is interior to Θ and the likelihood function is sufficiently smooth. Analogous but more elaborate results hold even when θ_0 is not interior to Θ.

Proposition 8 *Let $\{a_N\}$ be a sequence of real numbers and for $p \in \mathbb{N}$, suppose that $\theta_0 \in \mathbb{R}^p$ and that $\{\hat{Q}_N : \Omega \to \mathbb{R}\}$, $\{Q_N : \Omega \to \mathbb{R}\}$, $\{\hat{\theta}_N : \Omega \to \mathbb{R}^p\}$, $\{g_N : \Omega \to \mathbb{R}^p\}$, and $\{H_N : \Omega \to \mathbb{R}^{p \times p}\}$ are sequences of measurable functions such that*

$$a_N \hat{Q}_N = a_N Q_N + a_N g_N' \left(\hat{\theta}_N - \theta_0\right) + a_N \left(\hat{\theta}_N - \theta_0\right)' H_N \left(\hat{\theta}_N - \theta_0\right)/2 + o_P(1),$$

where, for random matrices Z_0, Z_1, Z_2, Z_3 of suitable dimension,

$$(a_N Q_N,\ a_N^{1/2}\left(\hat{\theta}_N - \theta_0\right)',\ a_N^{1/2} g_N',\ (vec(H_N))') \xrightarrow{d} (Z_0, Z_1', Z_2', (vec(Z_3))').$$

Then

$$a_N \hat{Q}_N \xrightarrow{d} Z_0 + Z_2' Z_1 + Z_1' Z_3 Z_1/2.$$

In our application, $a_N = N$, $p = dn$, $\theta_0 = vec(\sigma_0)$, $a_N \hat{Q}_N = \hat{T}_N(\Lambda, R)$, $a_N Q_N = T_N(\Lambda, R; \sigma_0)$, where

$$T_N(\Lambda, R; \sigma) := \sup_{\lambda \in \Lambda} \sum_{j=1}^{N} \ln f(R_{t_j}; \sigma, \lambda) - \sup_{\lambda \in R} \sum_{j=1}^{N} \ln f(R_{t_j}; \sigma, \lambda),$$

$\hat{\theta}_N = vec(\hat{\sigma}_N)$, $g_N = N^{-1}(\partial/\partial\sigma)T_N(\Lambda,R;\sigma_0)$, and $H_N = N^{-1}(\partial^2/\partial\sigma\partial\sigma')T_N(\Lambda,R;\sigma_0)$. Under our assumptions, $N^{1/2}(\hat{\sigma}_N - \sigma_0)$ and $N^{-1/2}(\partial/\partial\sigma)T_N(\Lambda,R;\sigma_0)$ will generally jointly obey a central limit theorem, and $N^{-1}(\partial^2/\partial\sigma\partial\sigma')T_N(\Lambda,R;\sigma_0)$ converges in probability to a constant matrix. The desired limiting distribution follows provided $T_N(\Lambda,R;\sigma_0)$ also converges in distribution (jointly with the other random variables).

For this, we can apply results of Liu and Shao (2003), whose theorem 3.1 provides general regularity conditions for the non-identified case ensuring that

$$\lim_{N\to\infty} 2T_N(\Lambda,R_0;\sigma_0) = \sup_{S\in\mathscr{F}_\Lambda} \max(W_S,0)^2,$$

where $R_0 := \{\lambda_0\}$, W_S defines a centered Gaussian process $\{W_S : S \in \mathscr{F}_\Lambda\}$ with uniformly continuous sample paths and covariance kernel

$$E(W_{S_1}W_{S_2}) = E(S_1S_2),$$

and $\mathscr{F}_\Lambda$ is a specific Donsker class of functions, a set of limits of generalized score functions S (see Liu and Shao, 2003, eq.(3.1)).

In our application, interest attaches to

$$T_N(\Lambda,R;\sigma_0) = T_N(\Lambda,R_0;\sigma_0) - T_N(R,R_0;\sigma_0),$$

so Liu and Shao's theorem 3.1 implies that under H_o^Λ,

$$\lim_{N\to\infty} 2T_N(\Lambda,R;\sigma_0) = \sup_{S\in\mathscr{F}_\Lambda} \max(W_S,0)^2 - \sup_{S\in\mathscr{F}_R} \max(W_S,0)^2.$$

Although this gives the asymptotic distribution only for $T_N(\Lambda,R;\sigma_0)$, the extension to the required joint convergence appears straightforward.

Proposition 8 then delivers the asymptotic distribution of $\hat{T}_N(\Lambda,R)$. As this appears to be a complicated distribution, we seek computationally simple methods for obtaining the desired critical values. One particularly appealing approach is to use the method of subsampling, as the existence of the limiting distribution for $\hat{T}_N(\Lambda,R)$ generally suffices to ensure that subsampling can generate valid asymptotic critical values.

Specifically, we obtain valid asymptotic critical values for $\hat{T}_N(\Lambda,R)$ by applying a subsampling algorithm proposed by Romano and Shaikh (2008). Let $b := b_N < N$ be a sequence of integers such that $b \to \infty$ and $b/N \to 0$. Let B_N be the number of randomly chosen subsamples of size b from a sample of size N, and let $\hat{T}_{N,b,k}$ be the subsampled test statistic for the kth subsample of size b, specifically

$$\hat{T}_{N,b,k} := \inf_{\lambda\in R} - \sum_{j\in\mathscr{J}_k} \ln f(R_{t_j};\hat{\sigma}_{N,b,k},\lambda) - \inf_{\lambda\in\Lambda^M}\left(- \sum_{j\in\mathscr{J}_k} \ln f(R_{t_j};\hat{\sigma}_{N,b,k},\lambda)\right),$$

where $\mathscr{J}_k$ is the $b-$element set of indexes for the kth subsample. Note that for each k, $\hat{T}_{N,b,k}$ is evaluated using a first stage estimate $\hat{\sigma}_{N,b,k}$, computed for that subsample.

Next, let $\alpha \in (0,1)$ be a prespecified significance level for the test, and define

$$d_{N,1-\alpha} = \inf\left\{ x : B_N^{-1} \sum_{1 \le k \le B_N} 1_{\{\hat{T}_{N,b,k} \le x\}} \ge 1-\alpha \right\}.$$

This is a subsampling estimator for the asymptotic $1-\alpha$ quantile of $\hat{T}_N(\Lambda, R)$. Theorem 3.4(i) of Romano and Shaikh (2008) then ensures that when H_o^{Λ} holds,

$$\liminf_{N\to\infty} P\left(\hat{T}_N(\Lambda,R) \le d_{N,1-\alpha}\right) \ge 1-\alpha,$$

so that $d_{N,1-\alpha}$ provides a valid asymptotic critical value for testing H_o^{Λ}. When the alternative H_A^{Λ} holds, $d_{N,1-\alpha}$ diverges, but at a sufficiently slow rate that the test based on $d_{N,1-\alpha}$ is nevertheless consistent, a consequence of $b/N \to 0$.

As Hall and Jing (1996) and Härdle, Horowitz, and Kreiss (2003) show, subsampling estimators of distributions of statistics converge at a slower rate to the limiting distribution than (block) bootstrap estimators. Despite this drawback, we use subsampling because it provides us with an asymptotically valid procedure and because standard bootstrap procedures are known to fail for some partially identified models, as pointed out by Chernozhukov, Hong, and Tamer (2007) and Bugni (2008)[5].

4.6 Confidence regions

Confidence regions can be constructed using a subsampling procedure proposed by Chernozhukov, Hong, and Tamer (2007) (CHT). For this it suffices that

$$\sup_{\lambda \in \Lambda_{I,0}^M} N\, Q_N(\hat{\sigma}_N, \lambda) \xrightarrow{d} Z,$$

where Z is a random variable.

Care is required in verifying this condition due to the presence of $\hat{\sigma}_N$. One might consider using Proposition 8 to establish this. The natural choices for this are $a_N = N$, $p = dn$, $\theta_0 = vec(\sigma_0)$, $a_N\hat{Q}_N = \sup_{\lambda \in \Lambda_{I,0}^M} N\, Q_N(\hat{\sigma}_N, \lambda)$, $a_N Q_N = \sup_{\lambda \in \Lambda_{I,0}^M} N\, Q_N(\sigma_0, \lambda)$, $\hat{\theta}_N = vec(\hat{\sigma}_N)$, $g_N = (\partial/\partial\sigma)\sup_{\lambda \in \Lambda_{I,0}^M} Q_N(\sigma_0, \lambda)$, and $H_N = (\partial^2/\partial\sigma\partial\sigma')\sup_{\lambda \in \Lambda_{I,0}^M} Q_N(\sigma_0, \lambda)$. It then suffices to verify that $\sup_{\lambda \in \Lambda_{I,0}^M} N\, Q_N(\sigma_0, \lambda)$ converges in distribution jointly with $N^{1/2}(\hat{\sigma}_N - \sigma_0)$ and $N^{1/2}(\partial/\partial\sigma)$

[5]Bugni (2008) studies a modified bootstrap procedure that is valid for a subclass of problems considered in Chernozhukov, Hong, and Tamer (2007).

$\sup_{\lambda \in \Lambda_{I,0}^{M}} Q_N(\sigma_0, \lambda)$, and that $(\partial^2/\partial\sigma\partial\sigma')\sup_{\lambda \in \Lambda_{I,0}^{M}} Q_N(\sigma_0, \lambda)$ converges in probability to a constant matrix. The first condition corresponds to the key primitive condition assumed by CHT (the "CHT condition"), and under mild conditions a central limit theorem holds for $N^{1/2}(\hat{\sigma}_N - \sigma_0)$ and $(\partial^2/\partial\sigma\partial\sigma')\sup_{\lambda \in \Lambda_{I,0}^{M}} Q_N(\sigma_0, \lambda)$ converges as required.

Nevertheless, it is not clear whether $N^{1/2}(\partial/\partial\sigma)\sup_{\lambda \in \Lambda_{I,0}^{M}} Q_N(\sigma_0, \lambda)$ converges in distribution; in particular, nothing appears to ensure that this quantity has (limiting) mean zero, so the central limit theorem need not hold. Accordingly, we seek an alternative approach.

A promising way to proceed is to recast our two-stage estimator as a single-stage estimator; as we show, this yields a straightforward formulation of the CHT condition. If this recasting is indeed possible, one might ask why this approach is not used from the outset. The main reason is that our likelihood-based approach is more robustly applicable to identifying the set of market risk prices of interest than the method of moments-based single-stage approach described next. In general settings, moment-based methods may introduce spurious zeros into the single-stage objective function, thereby possibly altering the apparent identified set in undesired ways. We describe how this can happen below. Using a two-stage approach permits us to ensure that the identified set is that associated with the risk prices of interest. The single-stage recasting can then be used with this identified set to deliver conditions justifying the CHT subsampling procedure.

To recast our two-stage estimator as a single-stage method of moments estimator, let $\beta := (\mu', vech(\Sigma)', \lambda')' \in \mathbb{B}$, say, and define the functions

$$\begin{aligned} m_0(R_{t_j};\beta) &= R_{t_j} - (t_j - t_{j-1})(\mu - (\Sigma_{11}, ..., \Sigma_{dd})'/2) \\ m_1(R_{t_j};\beta) &= vech[m_0(R_{t_j};\beta)m_0(R_{t_j};\beta)' - (t_j - t_{j-1})\Sigma] \\ m_2(R_{t_j};\beta) &= (\partial/\partial\lambda)\ln f(R_{t_j};\varsigma(\Sigma),\lambda), \end{aligned}$$

where Σ is a $d \times d$ symmetric positive semi-definite matrix with diagonal elements Σ_{ii}, $i = 1, ..., d$, and $\varsigma(\Sigma)$ is such that $\varsigma(\Sigma)\varsigma(\Sigma)' = \Sigma$. The functions m_0 and m_1 yield moment equations for estimating μ_0 and Σ_0. The function m_2 is the log-likelihood score with respect to λ.

Let $m := (m_1', m_2', m_3')'$, and define $\hat{m}_N(\beta) := (\hat{m}_{0,N}(\beta)', \hat{m}_{1,N}(\beta)', \hat{m}_{2,N}(\beta)')'$, where

$$\hat{m}_{i,N}(\beta) := N^{-1}\sum_{j=1}^{N} m_i(R_{t_j};\beta), \quad i = 0, 1, 2.$$

Then there generally exists a unique solution $(\hat{\mu}_N, \hat{\Sigma}_N)$ to the first two moment equations satisfying

$$\hat{m}_{0,N}(\hat{\mu}_N, vech(\hat{\Sigma}_N), \lambda) = 0 \quad \text{and} \quad \hat{m}_{1,N}(\hat{\mu}_N, vech(\hat{\Sigma}_N), \lambda) = 0$$

for all λ, as m_0 and m_1 do not depend on λ. This delivers a first-stage estimator $\hat{\Sigma}_N$ that is generally $N^{1/2}-$consistent for Σ_0. From this we construct $\hat{\sigma}_N = \varsigma(\hat{\Sigma}_N)$. Further, for all $\lambda \in \tilde{\Lambda}_N$, say, where

$\tilde{\Lambda}_N$ is the subset of $\hat{\Lambda}_N$ containing the zeros of $Q_N(\hat{\sigma}_N, \lambda)$, we have

$$\hat{m}_{2,N}(\hat{\mu}_N, vech(\hat{\Sigma}_N), \lambda) = 0.$$

That is, the zeros of $\hat{m}_{2,N}$ correspond to (a subset of) our second stage set estimator.

Collecting these facts, we have that for all $\beta \in \{\hat{\mu}_N\} \times \{vech(\hat{\Sigma}_N)\} \times \tilde{\Lambda}_N$,

$$\hat{m}_N(\beta) = 0.$$

Although it is not a typical feature of the likelihood for the Black-Scholes economy, in more general settings, there may be other zeros of $\hat{m}_N(\beta)$, as the likelihood scores may have zeros corresponding to local minima, maxima, or inflection points of the likelihood function. These are the "spurious" zeros referred to above. Nevertheless, because we will not rely on $\hat{m}_N$ to define the identified set of interest, this will not create difficulties.

By making two more identifications, we can state a version of the CHT condition justifying subsampling in the present context. First, we define the single-stage sample objective function

$$\tilde{Q}_N(\beta) := \hat{m}_N(\beta)' \hat{m}_N(\beta).$$

Certain minimizers of this function correspond to our two-stage estimators. Note that this is a standard method of moments objective function; because there are no over-identifying moment conditions, this is also the generalized method of moments objective function (Hansen, 1982). We have $\tilde{Q}_N(\beta) \geq 0$, with the minimum attained at zero because of the lack of over-identification. Finally, define the identified set

$$\mathbb{B}^M_{I,0} := \{\mu_0\} \times \{vech(\Sigma_0)\} \times \Lambda^M_{I,0}.$$

The CHT condition justifying subsampling can now be stated as

$$\sup_{\beta \in \mathbb{B}^M_{I,0}} N\, \tilde{Q}_N(\beta) \xrightarrow{d} Z.$$

To implement the CHT method, we first construct a "preliminary" consistent set estimator, say $\hat{\Lambda}_{N,0}$, and let $l = 1$. Next, we randomly choose B_N subsets of size b, and compute $\hat{\varepsilon}_l$ as the $1-\alpha$ quantile of the statistics

$$\mathscr{Z}_{N,b,k} := \sup_{\lambda \in \hat{\Lambda}_{N,l-1}} bQ_{N,b,k}(\hat{\sigma}_{N,b,k}, \lambda), \quad k = 1, ..., B_N,$$

where $Q_{N,b,k}(\hat{\sigma}_{N,b,k}, \lambda)$ is the criterion function evaluated for the kth $b-$element subset drawn from the full sample of N observations. We then use $\hat{\varepsilon}_l$ to get a new set estimator $\hat{\Lambda}_{N,l} = \{\lambda \in \Lambda : N\ Q_N(\hat{\sigma}_N, \lambda) \leq \hat{\varepsilon}_l\}$.

We may repeat this process for $l = 2, ..., L$. The final set estimator

$$\hat{\Lambda}_N = \{\lambda \in \Lambda : NQ_N(\hat{\sigma}_N, \lambda) \leq \hat{\varepsilon}\}$$

is a $1-\alpha$ confidence set for $\Lambda_{I,0}^M$, taking $\hat{\varepsilon} = \hat{\varepsilon}_L$. That is, $\liminf_{N\to\infty} P(\Lambda_{I,0}^M \subseteq \hat{\Lambda}_N) \geq 1-\alpha$.

Moreover, $\hat{\Lambda}_N$ is a consistent set estimator when we choose $\hat{\varepsilon} = \min(\hat{\varepsilon}_L, q_N + \kappa_N)$ for any $\kappa_N \propto \ln N$, where $q_N := \inf_{\lambda\in\Lambda} Q_N(\hat{\sigma}_N, \lambda)$.

5 Applications

In this section, we illustrate set estimation and hypothesis testing with two examples. The first studies international risk sharing. The second studies risk premia for market capitalization range index returns.

5.1 International risk sharing

5.1.1 A three-country asset price process

Consider three portfolios with prices S_t^1, S_t^2, and S_t^3, each of which is traded in the domestic market of each country $i = 1, 2, 3$. We assume that investors can potentially participate in all three markets. In addition, we assume there is an international risk free asset with a known rate of return r. Let $S_t = (S_t^1, S_t^2, S_t^3)'$. Suppose $\{S_t\}$ is generated by a multivariate Black-Scholes process with $d = 3$ and $n = 4$. Suppose further that Assumptions 2 and 3 hold. The identifying restriction on the diffusion coefficient, therefore, is $\rho(\sigma_0) = \rho_0 vec(\sigma_0) = 0$, as described in Section 3.2. We thus interpret the first three elements of dW_t as country-specific risks and the fourth element as international risk.

The true market price of risk λ_0 is a 4×1 vector that satisfies $\sigma_0\lambda_0 = \mu_0 - r\iota$ and the bound $\|\lambda_0\| \leq M$. The first three elements of λ_0 represent risk premia on the country specific risks, and the fourth element represents a risk premium on the international risk. Because $d = 3$ and $n = 4$, λ_0 is not point identified. The identified set for the market price of risk, therefore, is $\Lambda_{I,0}^M = \{\lambda : \sigma_0\lambda - \mu_0 - r\iota, \|\lambda\| \leq M\}$. Using set estimation, we can estimate the set of market prices of risk (and therefore risk neutral measures) that are compatible with the behavior of portfolio returns.

In this example, Assumption 3 fully identifies σ_0, so the identified set for the diffusion coefficients is a singleton, $S = \{\sigma_0\}$. Let $\hat{\Sigma}_N$ be the standard sample covariance estimator. This is a $\sqrt{N}-$consistent estimator of Σ_0 under Assumption 1. Using the relationship $\sigma_0\sigma_0' = \Sigma_0$, we define a first stage estimator $\hat{\sigma}_N$ to be the (unique) estimator such that $\hat{\sigma}_N\hat{\sigma}_N' = \hat{\Sigma}_N$. This estimator $\hat{\sigma}_N$ is then a $\sqrt{N}-$consistent estimator of σ_0. Given the first-stage estimator $\hat{\sigma}_N$, we estimate the identified set $\Lambda^M_{I,0}$ using eq. (8). This gives a set of market prices of risk compatible with the observed domestic portfolio returns across the three countries. Assumptions 1-3 ensure that the regularity conditions of Theorem 7 hold, so this is a consistent set estimator of $\Lambda^M_{I,0}$.

We turn now to hypothesis testing. If there is an integrated international financial market, country specific risks should be diversified away. This implies a simple hypothesis that risk premia for the country-specific risks are zero, whereas those who accept the international aggregate risk receive a nonzero risk premium as a reward. According to Lewis (1995), complete markets and optimal risk-sharing imply that the stochastic discount factor varies only with the common international component and is independent of any country specific disturbances. She tests this hypothesis by regressing consumption growth on a constant (the common international component) and domestic output growth (a proxy for country-specific risk), using cross-country data.

In our framework, the international risk sharing hypothesis can be tested using panel data on portfolio returns rather than consumption data[6]. If asset prices are determined in general equilibrium for the integrated world market as described by Lewis (1995), the stochastic discount factor $m(W_t)$ depends only on the international risk. This implies that the elasticities of the pricing kernel with respect to the country-specific risks are zero: i.e., $-\partial \ln m(W_t)/\partial W^i_t = 0$ for $i = 1,2,3$. Recall that $m(W_t) = e^{-rt}\xi_t$ and that ξ_t is given by (3) in the multivariate Black-Scholes economy. Thus, the vector of elasticities equals λ. If the country-specific risks are fully diversified away, the first three elements of λ should be 0. We thus let R be the subset of Λ such that

$$R := \{\lambda \in \Lambda : \lambda_1 = \lambda_2 = \lambda_3 = 0\}.$$

The null hypothesis is that there is at least one element λ in the identified set that is consistent with full international risk sharing. This is equivalent to $\Lambda^M_{I,0} \cap R \neq \emptyset$. We can test this hypothesis using the statistic $\hat{T}_N(\Lambda,R) = \inf_{\lambda\in R} NQ_N(\hat{\sigma}_N,\lambda)$ and the subsampling procedure described in Section 4.5.

5.1.2 Empirical results

For our empirical study, we consider the financial markets of the U.S., Japan, and Europe. For these three regions, we use standard publicly available data obtained from the Global Financial Database. For the U.S., we use the S&P 500 composite price index, a value weighted index that represents about 75% of the market capitalization of the New York Stock Exchange. To ensure

[6] An approach similar to ours was taken by Campbell and Hamao (1992). They used U.S. and Japanese stock returns to investigate capital market integration based on a single factor model.

comparability across countries, we use the Tokyo Stock Exchange Price Index (TOPIX) for Japan and the Morgan Stanley Capital International (MSCI) Europe Price Index for Europe. The TOPIX is a value weighted index of all securities traded on the first section of the Tokyo Stock Exchange. The MSCI Europe Price Index is a free-float-adjusted market capitalization-weighted index constructed from indices in 16 developed markets: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.

We use the monthly 1-month T-Bill yield taken from the Center for Research on Security Prices (CRSP) to construct the short term risk-free rate. Specifically, we take the average yield over the whole sample period as our constant risk-free rate r, which is 5.6232% per annum.

The first and last months for which we are able to obtain complete data for all three portfolios and the T-Bill yield are January 1970 and December 2007, for a total of 456 observations. We remove the top and bottom 2.5% of returns from our sample to ensure that the results are not influenced by large outliers. This reduces the sample size to 405. Panel A in Table 1 shows summary statistics for the four variables over the full sample period. Panel B reports their variance and correlation coefficients.

Table 2 reports the first stage estimate $\hat{\sigma}_N$ for σ_0 with standard errors in parentheses, computed by the delta method. The estimates for the full sample (70:1-07:12) used in the second stage estimation appear in the last column. To assess the stability of the sample, we also report estimates for four sub-periods (70:1-79:12, 80:1-89:12, 90:1-99:12, and 00:1-07:12) in columns 2-5. The estimated coefficients are stable across sub-periods in most cases, although the diffusion coefficient of the MSCI index on its idiosyncratic risk is poorly estimated, especially during the 1980's and 1990's.

To consistently estimate the identified set $\Lambda^M_{I,0}$, we must choose $\hat{\varepsilon}$ satisfying the conditions of Theorem 6. Any $\hat{\varepsilon}$ that grows slower than N ensures the consistency of our estimator. We thus choose $\hat{\varepsilon}_0 = q_N + \kappa_N$, where $q_N = \inf_{\lambda\in\Lambda} Q_N(\hat{\sigma}_N, \lambda)$ and $\kappa_N \propto \ln N$, and we form the consistent set estimator $\hat{\Lambda}_{N,0} = \{\lambda : NQ_N(\hat{\sigma}_N, \lambda) \leq \hat{\varepsilon}_0\}$. In this example, $\Lambda^M_{I,0}$ is a line segment in four-dimensional Euclidean space. The set estimator, therefore, is a four-dimensional cylinder that shrinks down to this line with probability approaching 1.

We project this cylinder to lower dimensional Euclidean spaces to understand its shape. Figure 3 shows the convex hulls of boundary points of the four-dimensional cylinder projected onto three-dimensional spaces. Note that the surface of the original set is smooth, but the set is approximated by a polygon because of the discretization of the grid. Figure 4 shows our second stage set estimator projected onto two-dimensional subspaces.

We can construct a confidence region or another consistent set-estimator using the CHT subsampling procedure described above, using $\hat{\Lambda}_{N,0}$ as our preliminary estimator. The 95% confidence region is smaller than the preliminary set estimator and contains the origin, as depicted in Figures 5 and 6.

Finally, we formally test the international risk sharing hypothesis. The statistic $\hat{T}_N(\Lambda, R)$ is 9.12 in our sample. We estimate the critical value for $\hat{T}_N(\Lambda, R)$ by subsampling. Table 3 provides critical values for different choices of b and B_N. For all of these critical values, we reject the null hypothesis of international risk sharing. Figure 7 shows the corresponding subsampling distribution with $b = 40$ and $B_N = 5,000$.

As an experiment, we also computed subsampled critical values (not reported here) always using the full sample first-stage estimator in the subsampling exercise. The critical values for the different choices of b and B_N are largely similar to those in Table 3, suggesting that the first-stage estimation is not having much impact on the asymptotic distribution of our test statistic.

The fact that this test rejects, whereas the 95% confidence interval contains the origin provides mixed evidence for the risk-sharing hypothesis. Possibly, the direct hypothesis testing approach is more powerful; but without further investigation, we cannot rule out the possibility that this mixed result is due to variations associated with the method of subsampling that would be mitigated in larger samples.

5.2 Risk Premia on Cap Range Index Returns

5.2.1 An asset price process for three cap range indexes

Our second illustration concerns risk premia for market capitalization range (”cap range”) index returns. Since the seminal work of Fama and French (1993, 1996), many empirical studies have shown that there are sources of priced risk beyond just that associated with movements in the market portfolio. One of these risk factors is known to be related to firm size. Using cap range index returns, we study risk premia on both size-specific risk factors and the market factor without assuming the uniqueness of the risk price (or, equivalently, the risk neutral measure).

For this, suppose Assumption 1 holds and that there are three portfolios (”large cap,” ”mid cap,” and ”small cap”) whose returns are generated by a multivariate Black-Scholes economy with $d = 3$ and $n = 4$. As above, we impose Assumptions 2 and 3, so that the index return for each cap range is driven only by its idiosyncratic factor and the market factor. Because this structure is exactly the same as in the previous example, we can use the same set-estimation methods.

Previous studies have found that the small cap ($j = 3$) risk and the market risk are priced in the market (see, e.g., Fama and French, 1993, 1996; and Liew and Vassalou, 2000). If the other risks are diversified away, we expect $\lambda_1 = \lambda_2 = 0$. We thus let R be the subset of Λ^M such that

$$R := \{\lambda \in \Lambda^M : \lambda_1 = \lambda_2 = 0\}.$$

Therefore, we consider the null hypothesis that there is at least one parameter value in the identified set that is compatible with the irrelevance of the large cap and mid cap risks: $H_o : \Lambda^M_{I,0} \cap R \neq \emptyset$. Again, we can test this hypothesis using the framework in section 4.5.

5.2.2 Empirical results

For our empirical study, we consider three subclasses of firm sizes using the S&P/Citigroup Global Cap Range Index Returns. There, stocks are classified on the basis of their float-adjusted market capitalization. We examine daily returns for the following three indexes: *large cap* ($>$ \$5 billion), *mid cap* (\$1-\$5 billion), and *small cap* ($<$ \$1 billion). The first and last days for which we are able to obtain complete data for all three index returns are August 1, 1989 and December 31, 2007, for a total of 4,805 observations. After removing the top and bottom 2.5% of returns, we obtain 4,420 observations. Once again, we use the monthly 1-month T-Bill yield from CRSP to construct the short term risk-free rate for the same sample period. Our constant risk-free rate r is the average of these rates, 4.0633% per annum.

Table 4 reports summary statistics. Table 5 reports the first stage estimate $\hat{\sigma}_N$, with standard errors computed using the delta method. The estimated coefficients are stable across sub-periods.

The second-stage set estimator with $M = 20$ is depicted in Figures 8 and 9. The 95% confidence region is depicted in Figures 10 and 11 We observe that a non-zero premium on the market risk and a zero premium on the small cap risk is plausibly compatible with the returns distribution. For the premia on the large cap and mid cap indexes, the upper left panels of Figures 9 and 11 show that the origin is in the set estimator and the confidence region, implying that the irrelevance of the large cap and the mid cap risks is compatible with the returns distribution.

Next, we formally test the irrelevance of large cap and mid cap risks. The test statistic $\hat{T}_N(\Lambda, R)$ is 0.12 in our sample. As before, we estimate the critical value for $\hat{T}_N(\Lambda, R)$ by subsampling. Table 6 provides critical values for different choices of b and B_N. For example, with $b = 80$ and $B_N = 1,000$ the critical value is 1.70. Figure 12 shows this subsampling distribution. For none of the tabulated critical values do we reject the null hypothesis. Therefore, the irrelevance of the large cap and mid cap risks is statistically compatible with the observed returns distribution.

6 Modeling More General Asset Price Processes

Our discussion so far has focused mainly on the Black-Scholes case for clarity and conciseness. To the extent that this case is overly simplistic, our empirical results constitute only an illustrative first step in the study of risk pricing in incomplete markets. Nevertheless, much of our analysis and discussion extends to more general processes, providing the foundation for more sophisticated empirical studies. In this section we discuss some aspects of this extension.

A more general data generating process whose special cases are often used in applications is the following geometric process:

Assumption 4 (Multivariate Geometric Process) *Let $\{W_t\}$ be a vector of $n \in \mathbb{N}$ independent standard Brownian motions under P adapted to the filtration $\{\mathscr{F}_t\}$. Let $\{S_t\}$ be a vector of $d \in \mathbb{N}$ assets such that $S_0^i = 1$ and solving the stochastic differential equations*

$$dS_t^i = \mu_{0t}^i S_t^i dt + \sigma_{0t}^i S_t^i dW_t, \quad t \in [0,T], \quad i = 1,...,d,$$

where μ_{0t} has elements $\mu_{0t}^i : \Omega \to \mathbb{R}$ and σ_{0t} has $1 \times n$ rows $\sigma_{0t}^i : \Omega \to \mathbb{R}^n$, adapted to $\mathscr{F}_t$, $i = 1,...,d$. Further, $\{S_t\}$ does not admit arbitrage.

Under this assumption, $\{\mu_{0t}\}$ and $\{\sigma_{0t}\}$ are general adapted processes. For example, one may posit that a version of Assumption 3 holds, such that

$$\sigma_{0t} = \begin{bmatrix} \theta_{0\sigma}^1 (S_t^1)^{v-1} & 0 & 0 & \cdots & 0 & \theta_{0\sigma}^{d+1} (S_t^1)^{w-1} \\ 0 & \theta_{0\sigma}^2 (S_t^2)^{v-1} & 0 & \cdots & 0 & \theta_{0\sigma}^{d+2} (S_t^2)^{w-1} \\ \vdots & & \ddots & \cdots & & \vdots \\ \vdots & & & \ddots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 & \theta_{0\sigma}^d (S_t^d)^{v-1} & \theta_{0\sigma}^{2d} (S_t^d)^{w-1} \end{bmatrix},$$

where $v, w \in [0,1]$, and $(S_t^i)^{v-1}$ denotes the price of the ith security at t raised to the power $v-1$. Letting $\{\lambda_{0t}\}$ denote the true risk price process, suppose λ_{0t} has elements $\lambda_{0t}^i = \theta_{0\lambda}^i (S_t^i)^{1-v}$ for $i = 1,...,d$ and $\lambda_{0t}^n = \theta_{0\lambda}^n$ for $n = d+1$. Then, by the no arbitrage condition, the drift is determined by $\mu_{0t}^i = \theta_{0\sigma}^{d+i} \theta_{0\lambda}^n (S_t^i)^{w-1} + (\theta_{0\sigma}^i \theta_{0\lambda}^i + r)$ for $i = 1,...,d$. The process is indexed by the coefficient vector $\theta_0 = (\theta_{0\sigma}^1, \cdots, \theta_{0\sigma}^{2d}, \theta_{0\lambda}^1, \cdots, \theta_{0\lambda}^n) \in \mathbb{R}^{2d+n}$.

For the Black-Scholes economy, $v = w = 1$. The general case in which $v = w$ corresponds to a multivariate version of the constant elasticity of variance (CEV) process often used to model stock prices, short rates, forward rates, and stochastic volatilities. Various other cases of interest arise by varying v and w. For instance, choosing $v = 1/2$ and $w = 0$ gives a process whose idiosyncratic component follows a square root process, as in Cox, Ingersoll, and Ross (1985), and whose aggregate component follows a Brownian motion.

Further, because the $\sigma-$field $\mathscr{G}_t := \sigma(W_\tau, \tau \in [0,t])$ generated by the $t-$history of the multivariate Brownian motion $\{W_t\}$ may be a proper subset of $\mathscr{F}_t$, this assumption also covers certain more general stochastic volatility processes.

The analog of Assumption 2 becomes

Assumption 5 (Envelope Process) *There exists an adapted process* $\{M_t\}$ *such that* $0 < M_t < \infty$ *for* $t \in [0,T]$, $||\lambda_{0t}|| \leq M_t$, *and* $E^P\left[\exp\left(\int_0^T M_t^2 dt\right)\right] < \infty$.

This condition ensures that $var(\xi_T) < \infty$, where

$$\xi_T = \exp\left(-\int_0^T \lambda_{0s} \cdot dW_s - \frac{1}{2}\int_0^T \|\lambda_{0s}\|^2 ds\right).$$

Under these assumptions, the market prices of risk at time t belong to the random set

$$\Lambda_{I,t}^M := \Lambda_{I,t} \cap \Lambda^{M_t},$$

where

$$\Lambda_{I,t} := \{\lambda : \sigma_{0t}\lambda = \mu_{0t} - r\iota\} \quad \text{and} \quad \Lambda^{M_t} := \{\lambda : ||\lambda|| \leq M_t\}.$$

To apply maximum likelihood methods, we parameterize λ_{0t} and σ_{0t} as follows:

Assumption 6 (Parametric Specification) *Let* Θ *be a compact subset of* $\mathbb{R}^p$, $p \in \mathbb{N}$. *(i) For* $t \in [0,T]$, *the functions* $\ell_t : \Omega \times \Theta \to \mathbb{R}^n$ *and* $s_t : \Omega \times \Theta \to \mathbb{R}^{d \times n}$ *are such that for each* $\theta \in \Theta$, $\ell_t(\cdot,\theta)$ *and* $s_t(\cdot,\theta)$ *are measurable*$-\mathscr{F}_t$, *and for each* $\omega \in \Omega$, $\ell_t(\omega,\cdot)$ *and* $s_t(\omega,\cdot)$ *are continuous on* Θ*; (ii) for each* $\theta \in \Theta$ *and* $t \in [0,T]$, $||\ell_t(\cdot,\theta)|| \leq M_t$; *(iii) there exists* $\theta_0 \in \Theta$ *such that for* $t \in [0,T]$, $\sigma_{0t} = s_t(\cdot,\theta_0)$ *and* $\lambda_{0t} = \ell_t(\cdot,\theta_0)$.

Similar to our discussion above, Θ is the parameter space; for convenience, we assume that it implicitly embodies any prior restrictions known to hold for θ_0, such as $\rho(\theta_0) = 0$. The first part of this assumption specifies the parametric functions ℓ_t and s_t. In the second part, we require that the bound of Assumption 5 holds for all θ in Θ. The third part ensures that this specification is correct, in that there is a parameter value θ_0 in Θ corresponding to the true arbitrage-free process generating asset returns.

This assumption implies a parameterization for μ_{0t} of the form $m_t(\cdot,\theta) = s_t(\cdot,\theta)\ell_t(\cdot,\theta) + r\iota$. Alternatively, one may directly parameterize μ_{0t} instead of λ_{0t}; the no arbitrage condition then implies a parameterization for λ_{0t}. For brevity, we leave aside this possibility here.

Successive conditioning yields a likelihood function for returns $R_{t_j} := \ln S_{t_j} - \ln S_{t_{j-1}}$ defined by

$$f_N(R_{t_1}, \cdots, R_{t_N}; \theta) = \prod_{j=1}^{N} f_{t_j}(R_{t_j}; \theta \mid \mathscr{H}_{t_{j-1}}),$$

where $f_{t_j}(R_{t_j}; \theta \mid \mathscr{H}_{t_{j-1}})$ defines the likelihood for returns in t_j given the information $\mathscr{H}_{t_{j-1}}$, where $\sigma(R_{t_1}, \cdots, R_{t_{j-1}}) \subseteq \mathscr{H}_{t_{j-1}}$. This likelihood function does not necessarily have a closed form expression. In such cases, we may rely on an approximation of the likelihood function. See Aït-Sahalia (2002, 2008) and Kristensen (2008), for example.

Analogous to the Black-Scholes case, the criterion function $\bar{Q}_N : \Theta \to \bar{\mathbb{R}}_+$ is the shifted expected negative average log-likelihood defined by

$$\bar{Q}_N(\theta) := E^P\left[-N^{-1}\sum_{j=1}^{N} \ln f_{t_j}(R_{t_j}; \theta \mid \mathscr{H}_{t_{j-1}})\right] - q_{0,N},$$

where

$$q_{0,N} := E^P\left[-N^{-1}\sum_{j=1}^{N} \ln f_{t_j}(R_{t_j}; \theta_0 \mid \mathscr{H}_{t_{j-1}})\right].$$

The identified set is again the set of zeros of $\bar{Q}_N$,

$$\Theta_I := \{\theta \in \Theta : \bar{Q}_N(\theta) = 0\}.$$

The identified market prices of risk at time t are then given by the Effros-measurable set

$$\lambda_t(\Theta_I) := \{\lambda : \lambda = \ell_t(\theta), \ \ \theta \in \Theta_I\} \subset \Lambda^M_{I,t}.$$

It is not immediately obvious that $\lambda_t(\Theta_I) = \Lambda^M_{I,t}$. Ensuring that this holds may require further conditions. Nevertheless, the correct specification assumption ensures that $\lambda_{0t} \in \lambda_t(\Theta_I)$.

The sample criterion function is given by $Q_N : \Omega \times \Theta \to \bar{\mathbb{R}}_+$, defined by

$$Q_N(\theta) = -N^{-1}\sum_{j=1}^{N} \ln f_{t_j}(R_{t_j}; \theta \mid \mathscr{H}_{t_{j-1}}) - q_N,$$

where $q_N = \inf_\Theta -N^{-1}\sum_{j=1}^{N} \ln f_{t_j}(R_{t_j}; \theta \mid \mathscr{H}_{t_{j-1}})$.

As when we apply the RS-CHT approach above, we define an ε-level set for the sample criterion function by

$$\hat{\Theta}_N(\varepsilon) := \{\theta \in \Theta : N \cdot Q_N(\theta) \leq \varepsilon\}.$$

When we choose ε properly, the random set $\hat{\Theta}_N(\varepsilon)$ is Effros measurable and is a consistent set estimator or confidence region for Θ_I. The estimated prices of risk at time t are given by $\lambda_t(\hat{\Theta}_N(\varepsilon))$.

Hypothesis tests for the risk price process $\{\lambda_t(\Theta_I)\}$ can be conducted by inverting the confidence interval process $\{\lambda_t(\hat{\Theta}_N(\varepsilon))\}$ or using a likelihood ratio test for $\theta \in \Theta_R$, where $\Theta_R \subset \Theta$ expresses the restrictions specified by a null hypothesis of interest, e.g.,

$$H_o^{\Lambda} : \lambda_{0|\mathscr{T}} \in \Lambda_R, \;\; t \in \mathscr{T},$$

where $\mathscr{T}$ is a given subset of $[0,T]$, $\lambda_{0|\mathscr{T}}$ denotes the process $\{\lambda_{0t}\}$ restricted to $\mathscr{T}$, and Λ_R is a given subset of $\Lambda^{M_{\mathscr{T}}} := \{\{\Lambda^{M_t}\}, t \in \mathscr{T}\}$.

As before, two-stage estimation can help in mitigating the challenges arising in estimating θ_0. Specifically, let $\theta_0 := (\theta_{01}, \theta_{02}) \in \Theta_1 \times \Theta_2 =: \Theta$, and let $\hat{\Theta}_{1N}$ be a first-stage set estimator for θ_{01}.

Using this, let the second-stage set estimator for the identified set be defined by

$$\hat{\Theta}_N(\omega) := \{(\theta_1, \theta_2) \in \Theta : NQ_N(\omega, \theta_1, \theta_2) \leq \hat{\varepsilon}(\omega), \;\; \theta_1 \in \hat{\Theta}_{1N}(\omega)\},$$

where $\hat{\varepsilon}$ may be random, as before.

When the available restrictions suffice to fully identify θ_{01}, we have $\hat{\Theta}_{1N} = \{\hat{\theta}_{1N}\}$, say. The second-stage set estimator is then $\hat{\Theta}_N = \{\hat{\theta}_{1N}\} \times \hat{\Theta}_{2N}$, where

$$\hat{\Theta}_{2N}(\omega) := \{\theta_2 \in \Theta_2 : NQ_N(\omega, \hat{\theta}_{1N}(\omega), \theta_2) \leq \hat{\varepsilon}(\omega)\}.$$

Subsampling remains an appealing method for constructing confidence regions here. To provide conditions ensuring its validity, in particular the CHT condition, recasting a two-stage procedure as a single-stage procedure may again prove convenient. Depending on the particular circumstances, it may be possible to use a method of moments approach analogous to that discussed in Section 4.6. In other cases, it may be helpful to exploit an exponentially tilted likelihood, along the lines proposed by Kitamura and Stutzer (1997).

Formally ensuring consistency and convergence in distribution of these estimators will require careful specification of further regularity conditions appropriate to the specific context of interest. Nevertheless, the framework sketched here should prove helpful in pursuing these results.

7 Concluding remarks

In this paper, we study an econometric framework useful for estimating and testing hypotheses about the price of risk in the absence of complete markets. We state results ensuring the Effros-measurability and consistency of set estimators for the vector of market risk prices, and we discuss the construction of hypothesis tests and confidence sets using subsampling.

Our results build on the seminal work of Romano and Shaikh (2006, 2008) and Chernozhukhov, Hong, and Tamer (2007) for estimation and testing in partially identified models. To handle the challenges associated with jointly estimating all parameters of the model, we apply a two-stage method introduced by Kaido and White (2008). For the present application, we estimate covariance parameters in the first stage and risk prices in the second stage. To illustrate, we apply our methods to estimate market risk prices and test hypotheses concerning international risk sharing and market capitalization range indexes.

By providing new methods for inference on risk neutral measures in incomplete markets, our work thus complements that of Aït-Sahalia and Lo (1998), Chernov and Ghysels (2000), Clement, Gourieroux, and Monfort (2000), and Abadir and Rockinger (2003), among others.

An interesting direction for further research is to study investor risk preferences in the absence of the identification of market risk prices. This may create an opportunity to extend the work of Aït-Sahalia and Lo (2000), Jackwerth (2000), and Rosenberg and Engle (2002).

One of the key assumptions in our framework is a bound, M_t, on the market price of risk. Not only does this bound sharpen our set estimators, but it also plays a key role when using the estimated risk neutral measure to price non-redundant securities. Cochrane and Saá-Requejo (2000) show that this type of L^2 bound on the Radon-Nikodym derivative (or SDF) delivers sharper upper and lower bounds on the price of the non-redundant security. In related work, Bernardo and Ledoit (2000) consider a L^∞ bound on the Radon-Nikodym derivative. Further investigation of the choice of M_t, particularly the use of empirical evidence to choose M_t, is an interesting topic for further research.

Yet another interesting topic is the development of tests for market completeness *per se*. Such tests will require careful specification of the nature of the alternative complete and incomplete market structures, together with a theory of estimation and inference for parameters partially identified only under the alternative, possibly on the boundary of the parameter space. This will require extension of work of Davies (1977, 1987) and Andrews (1999, 2001) to the context of partial identification.

To maintain a sharp focus for our results, we have considered in detail the multivariate Black-Scholes economy. Nevertheless, our framework applies more broadly, and we sketch some features of its application to more general geometric processes. Extension to asset prices generated by Levy processes or subordinated processes are other interesting possibilities deserving attention in future work. Methods of estimation and inference for such potentially more realistic asset-price generating processes will then make possible increasingly refined empirical studies of risk pricing in incomplete markets.

A Tables

A.1 Tables for International Risk Sharing

Table 1: Summary Statistics for Stock Index Returns and the T-Bill Yield

A:	S&P500	TOPIX	MSCI	T-Bill
Mean	0.0061	0.0067	0.0069	0.0047
Std. Dev.	0.0343	0.0504	0.0378	0.0023
Min	-0.0893	-0.1141	-0.1027	0.0007
Max	0.0829	0.1382	0.0926	0.0135
Skewness	-0.0425	-0.0986	-0.0167	0.2800
Kurtosis	-0.1081	-0.0484	0.1247	0.3700
Obs	405	405	405	405

B:	S&P500	TOPIX	MSCI
S&P500	0.00118		
TOPIX	0.23588	0.00254	
MSCI	0.50724	0.38051	0.00143

The sample period is 1970:1-2007:12 with 456 monthly observations. We remove observations corresponding to the top and bottom 2.5% of returns for each series. This reduces the sample size to 405. We compute returns from the S&P500, TOPIX, and MSCI Europe Price Indexes obtained from the Global Financial Database for stock index returns. For the risk-free rate, we average monthly 1-month T-Bill yields taken from the CRSP database. We report robust measures of skewness and kurtosis. Skewness is computed as $SK = (Q_3 + Q_1 - 2Q_2)/(Q_3 - Q_1)$, where Q_i is the ith quartile of the return. Kurtosis is computed as $KR = (E_7 - E_5 + E_3 - E_1)/(E_6 - E_2) - 1.23$, where E_i is the ith octile. See Kim and White (2004) for details. Panel B reports variance (diagonal) and correlation (off-diagonal) coefficients for the returns.

Table 2: First-Stage Diffusion Coefficient Estimates

	2	3	4	5	6
	70:1-79:12	80:1-89:12	90:1-99:12	00:1-07:12	70:1-07:12
σ_{11}	0.1106	0.1101	0.0925	0.0520	0.0982
	(0.0292)	(0.0324)	(0.0336)	(0.0616)	(0.0337)
σ_{22}	0.1327	0.1546	0.1833	0.1518	0.1580
	(0.0531)	(0.0410)	(0.0428)	(0.0450)	(0.0441)
σ_{33}	0.0915	0.0000	0.0460	0.0676	0.0572
	(0.0514)	(0.2288)	(0.1348)	(0.0618)	(0.1188)
σ_{14}	0.0554	0.0534	0.0748	0.1038	0.0698
	(0.0472)	(0.0433)	(0.0458)	(0.0401)	(0.0447)
σ_{24}	0.1029	0.0690	0.0731	0.0648	0.0769
	(0.0733)	(0.0602)	(0.0595)	(0.0494)	(0.0607)
σ_{34}	0.0945	0.1606	0.1176	0.1112	0.1198
	(0.0575)	(0.1057)	(0.0593)	(0.0472)	(0.0632)
Number of Obs.	111	100	105	89	405

Columns 2-5 report the estimated diffusion coefficients (with standard errors in parentheses) for the following sample periods: 1970:1-1979:12, 1980:1-1989:12, 1990:1-1999:12, and 2000:1-2007:12. The last column reports the estimation results for the full sample.

Table 3: Critical Values for Various Choices of b and B_N

b	B_N	$d^{\Lambda}_{405,0.95}$
80	5000	6.36
80	1000	6.03
80	500	6.36
40	5000	4.63
40	1000	4.70
40	500	5.13
20	5000	3.81
20	1000	4.06
20	500	4.12

The third column reports 95% critical values of the test statistic. For each subsample, we estimate the diffusion coefficient $\hat{\sigma}_{N,b,k}$ and compute the statistic $\hat{T}_{N,b,k}$.

A.2 Tables for Market Cap Index Returns

Table 4: Summary Statistics for Cap Range Index Returns and the T-Bill Yield

A:	Large Cap	Mid Cap	Small Cap	T-Bill
Mean	0.00042	0.00051	0.00052	0.00016
Std. Dev.	0.00732	0.00706	0.00761	0.00007
Min	-0.02043	-0.02019	-0.02214	0.00003
Max	0.02021	0.01942	0.02117	0.00031
Skewness	0.03916	0.03459	-0.02127	-0.41912
Kurtosis	0.25411	0.22744	0.26137	0.04876
Obs	4,420	4,420	4,420	221

B:			
Large Cap	0.000054		
Mid Cap	0.860242	0.000050	
Small Cap	0.731907	0.904943	0.000058

The sample period is 08/01/1989-12/31/2007 with 4,805 daily observations for the cap range index returns. We remove observations corresponding to the top and bottom 2.5% of returns for each series. This reduces the sample size to 4,420. We compute returns from the S&P/Citigroup Global Cap Range Index data for: *large cap* (> \$5 billion), *mid cap* (\$1-\$5 billion), and *small cap* (< \$1 billion). For the risk-free rate, we average monthly 1-month T-Bill yields taken from the CRSP database for the same sample period (221 monthly observations). We report robust measures of skewness and kurtosis. Skewness is computed as $SK = (Q_3 + Q_1 - 2Q_2)/(Q_3 - Q_1)$, where Q_i is the ith quartile of the return. Kurtosis is computed as $KR = (E_7 - E_5 + E_3 - E_1)/(E_6 - E_2) - 1.23$, where E_i is the ith octile. See Kim and White (2004) for details. Panel B reports variance (diagonal) and correlation (off-diagonal) coefficients for the returns.

Table 5: First-Stage Diffusion Coefficient Estimates

	2	3	4
	08/01/1989-12/31/1999	01/03/2000-12/31/2007	08/01/1989-12/31/2007
σ_{11}	0.0662	0.0622	0.0654
	(0.0044)	(0.0043)	(0.0044)
σ_{22}	0.0000	0.0000	0.0000
	(0.0055)	(0.0078)	(0.0059)
σ_{33}	0.0515	0.0618	0.0591
	(0.0038)	(0.0049)	(0.0043)
σ_{14}	0.0961	0.1044	0.0991
	(0.0069)	(0.0067)	(0.0068)
σ_{24}	0.1079	0.1302	0.1181
	(0.0058)	(0.0065)	(0.0062)
σ_{34}	0.0848	0.1383	0.1084
	(0.0061)	(0.0076)	(0.0071)
Number of Obs.	2,603	1,817	4,420

Standard errors in parentheses.

Table 6: Critical Values for Various Choices of b and B_N

b	B_N	$d_{4,420,0.95}$
320	1000	1.68
320	500	1.55
160	1000	1.54
160	500	1.74
80	1000	1.70
80	500	1.59
40	1000	1.66
40	500	1.62

The third column reports 95% critical values of the test statistic. For each subsample, we estimate the diffusion coefficient $\hat{\sigma}_{N,b,k}$ and compute the statistic $\hat{T}_{N,b,k}$.

B Figures

B.1 Figures for International Risk Sharing

Figure 3: Second-Stage Set Estimator (Convex Hulls of Projections to Three-Dimensional Subspaces)

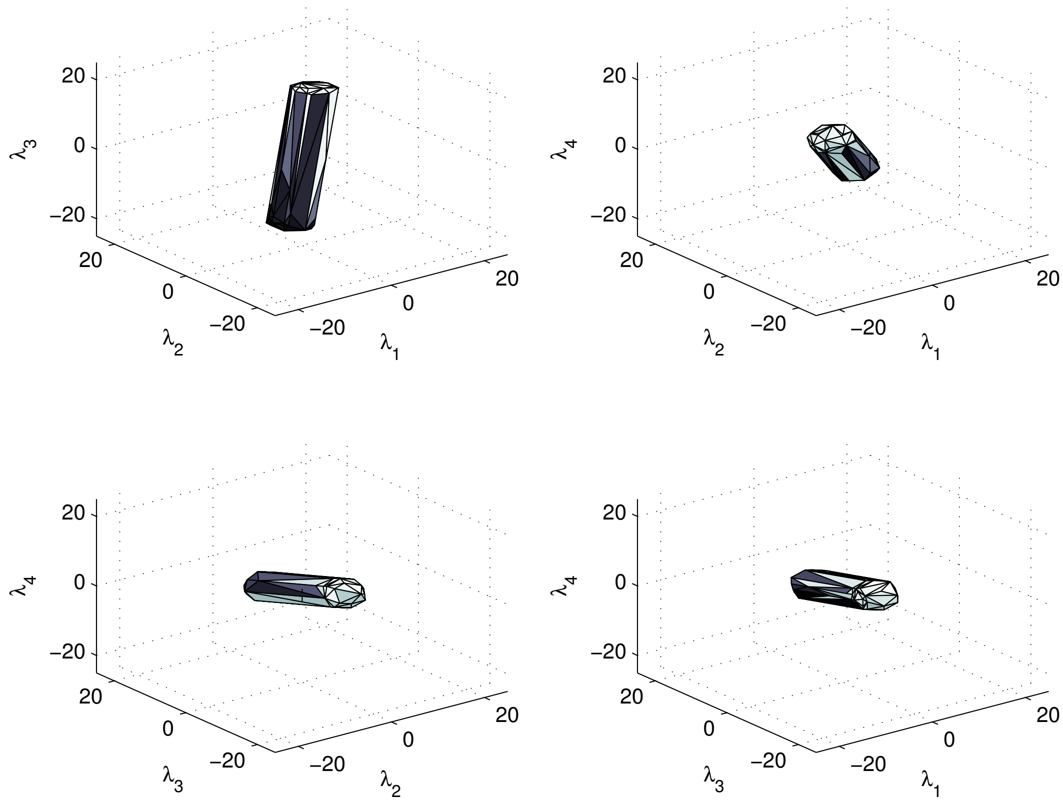


Figure 4: Second-Stage Set Estimator (Projections of Boundary Points to Two-Dimensional Sub-spaces)

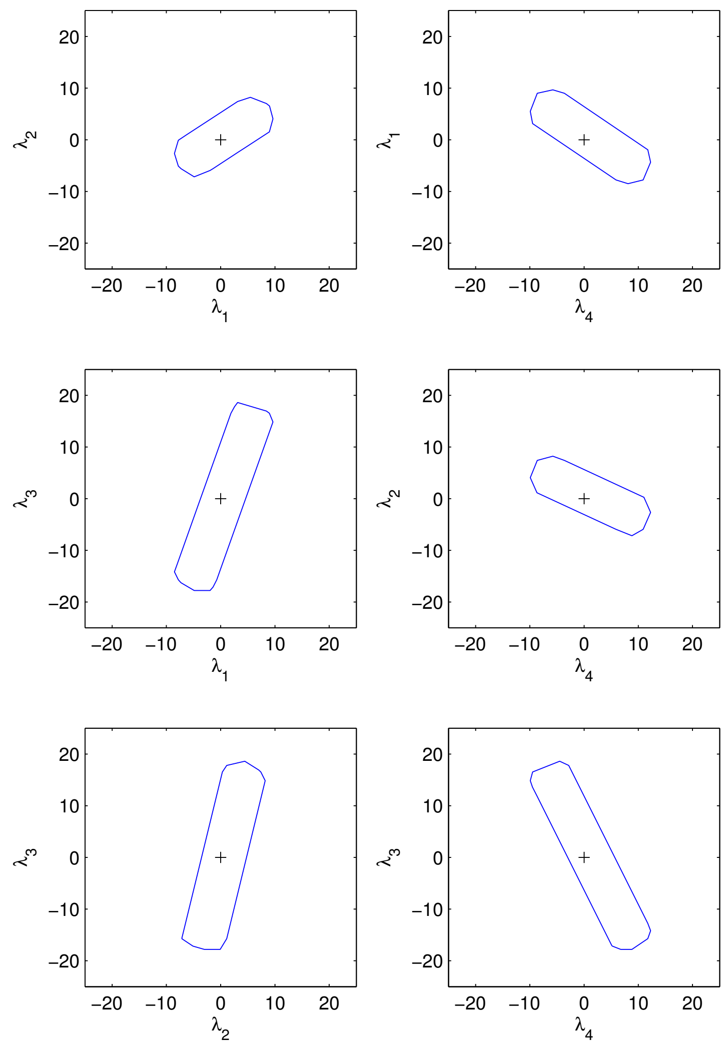


Figure 5: 95% Confidence Region (Convex Hulls of Projections to Three-Dimensional Subspaces)

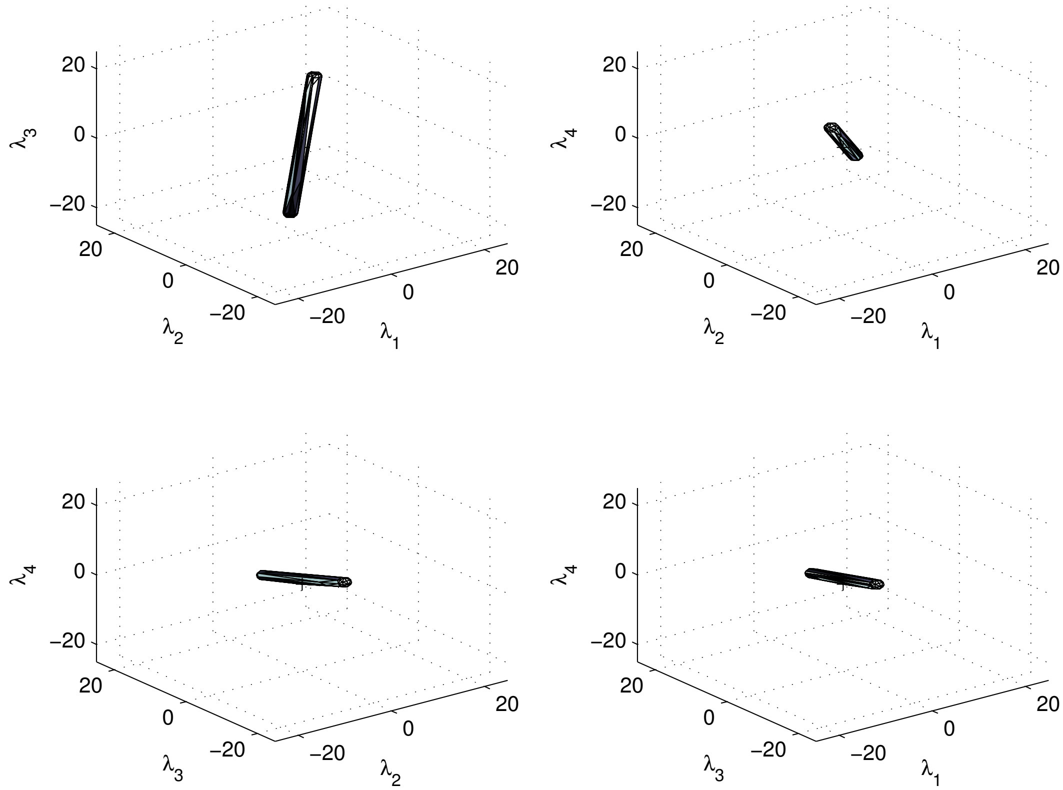


Figure 6: 95% Confidence Region (Projections of Boundary Points to Two-Dimensional Subspaces)

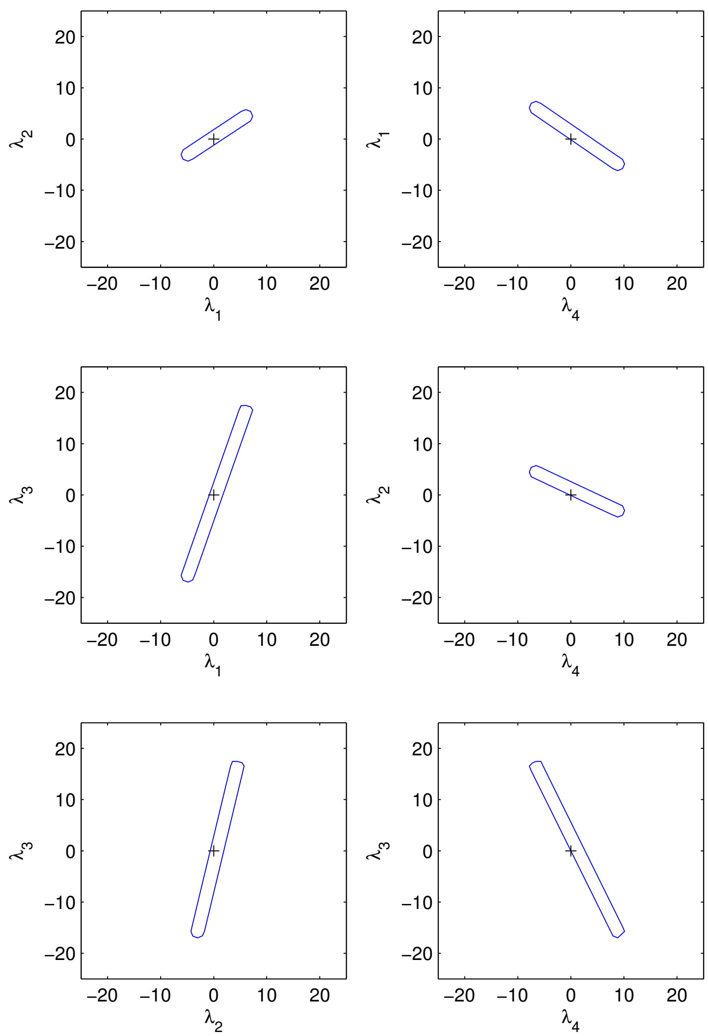
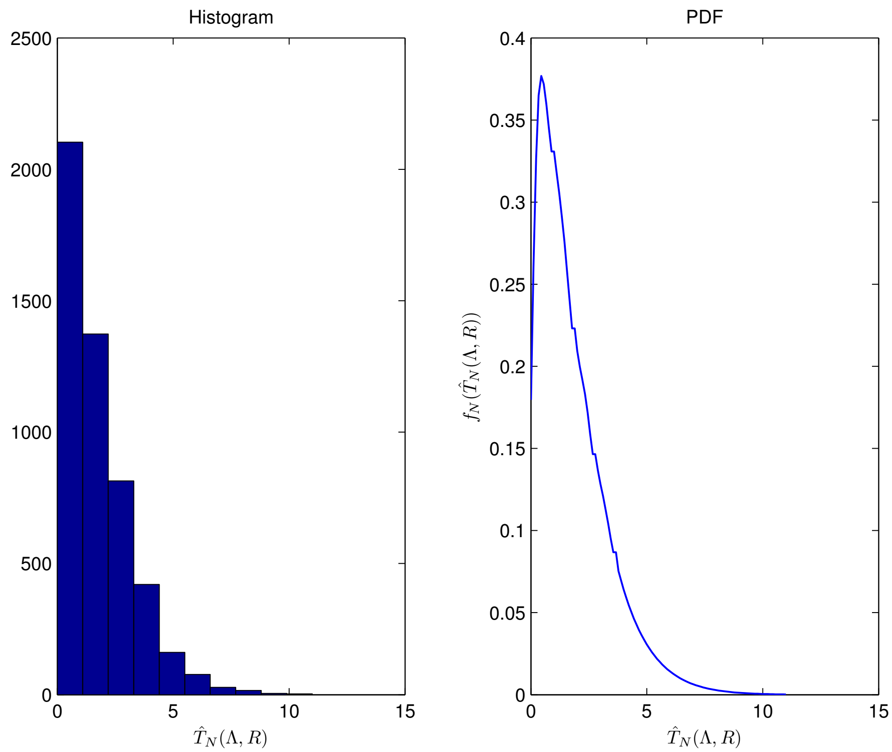


Figure 7: Subsampling Distribution of the LR Test Statistic $(N = 405, b = 40, B_N = 5,000)$



B.2 Figures for Returns on Market Cap Index Returns

Figure 8: Second-Stage Set Estimator (Convex Hulls of Projections to Three-Dimensional Subspaces)

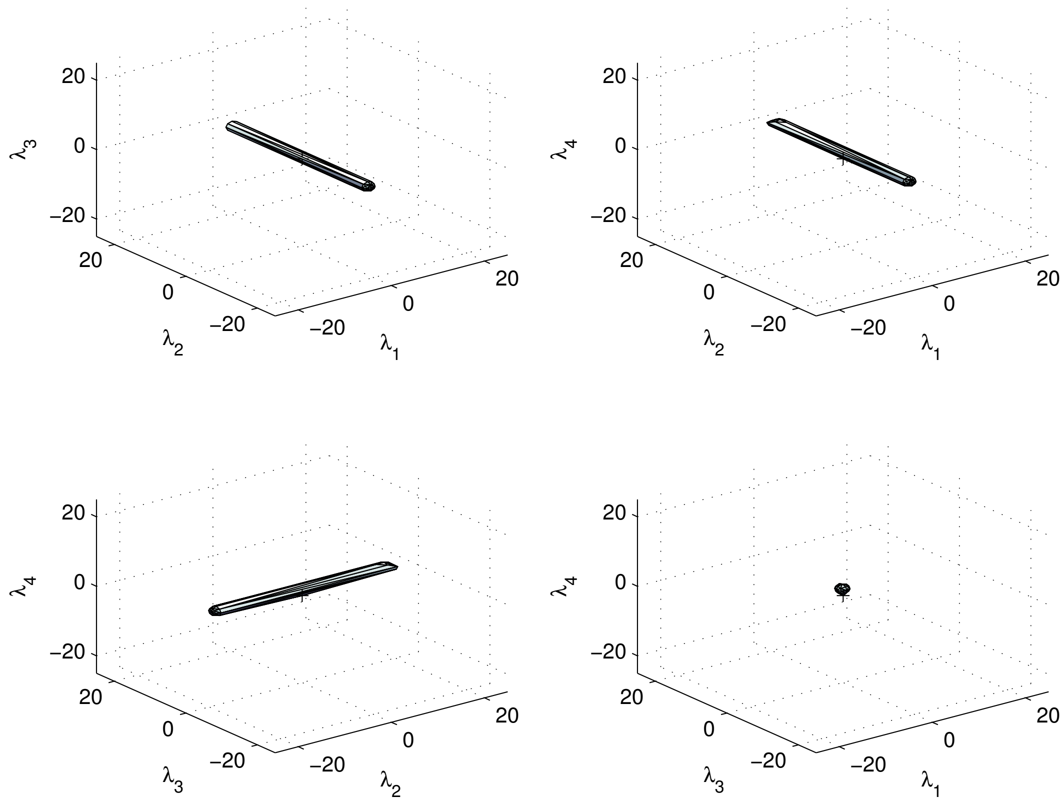


Figure 9: Second-Stage Set Estimator (Projections of Boundary Points to Two-Dimensional Subspaces)

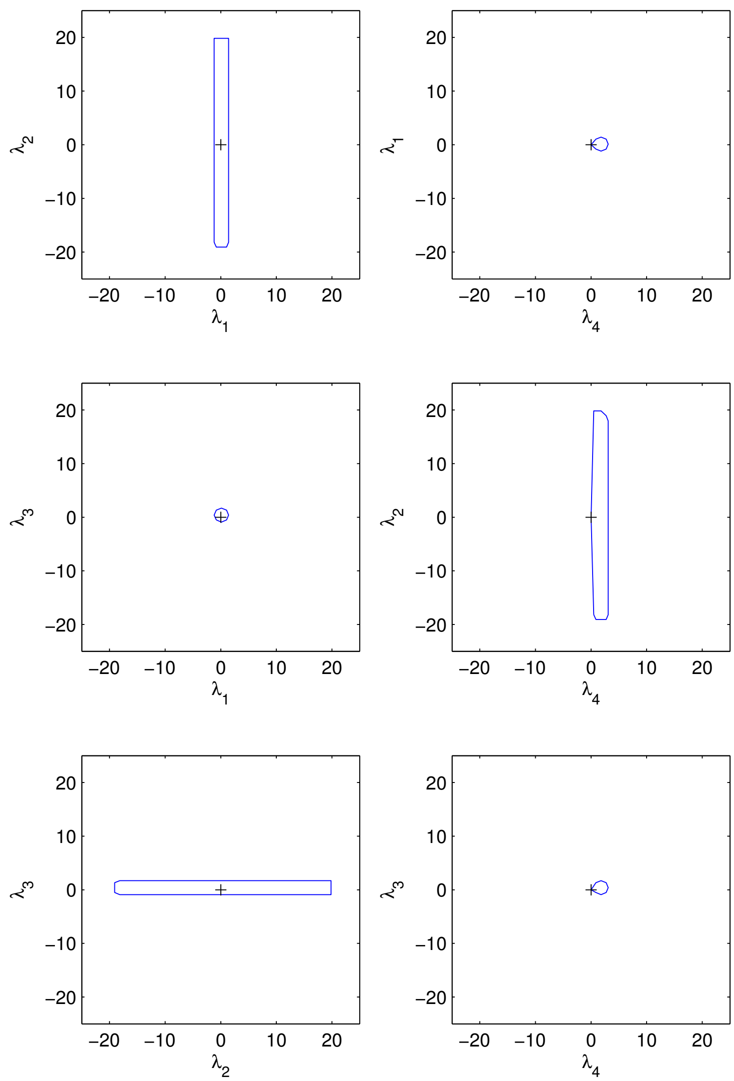


Figure 10: Confidence Region (Convex Hulls of Projections to Three-Dimensional Subspaces)

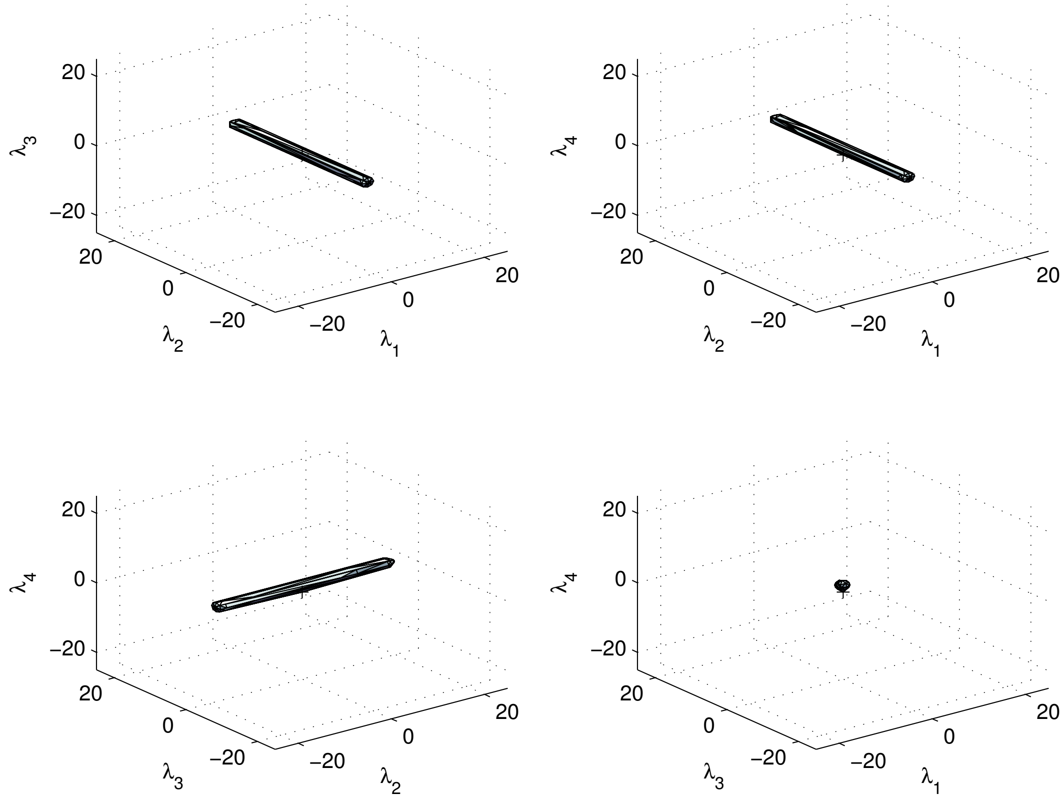


Figure 11: Confidence Region (Projections of Boundary Points to Two-Dimensional Subspaces)

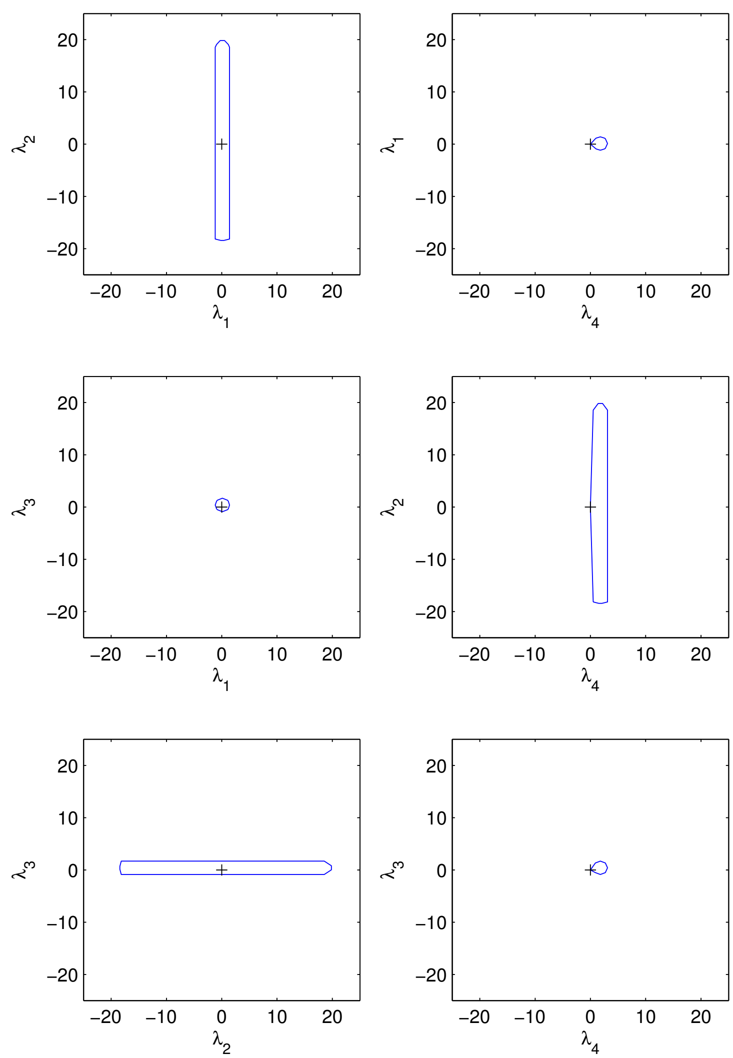
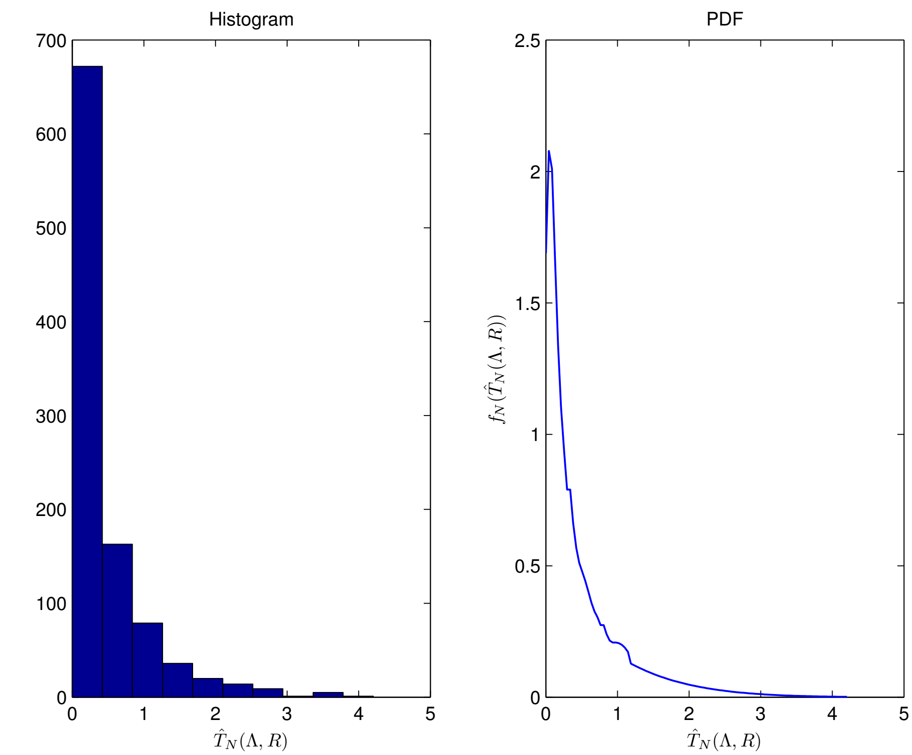


Figure 12: Subsampling Distribution of the LR Test Statistic ($N = 4,420, b = 320, B_N = 1,000$)



References

ABADIR, K., AND M. ROCKINGER (2003): "Density Functionals, with an Option-Pricing Application," *Econometric Theory*, 19(05), 778–811.

AÏT-SAHALIA, Y. (2002): "Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach," *Econometrica*, 70(1), 223–262.

——— (2008): "Closed-form Likelihood Expansions for Multivariate Diffusions," *Annals of Statistics*, 36(2), 906–937.

AÏT-SAHALIA, Y., AND A. W. LO (1998): "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," *The Journal of Finance*, 53(2), 499–547.

——— (2000): "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics*, 94(1-2), 9–51.

ANDREWS, D. W. K. (1999): "Estimation When a Parameter is on a Boundary," *Econometrica*, 67(6), 1341–1383.

——— (2001): "Testing When a Parameter is on the Boundary of the Maintained Hypothesis," *Econometrica*, 69(3), 683–734.

BERNARDO, A., AND O. LEDOIT (2000): "Gain, Loss and Asset Pricing," *Journal of Political Economy*, 108(1), 144.

BREEDEN, D. T. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7(3), 265–296.

BUGNI, F. A. (2008): "Bootstrap Inference in Partially Identified Models," Working Paper, Northwestern University.

CAMPBELL, J. Y., AND Y. HAMAO (1992): "Predictable Stock Returns in the United States and Japan: A Study of Long-Term Capital Market Integration," *The Journal of Finance*, 47(1), 43–69.

CHERNOV, M., AND E. GHYSELS (2000): "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 56(3), 407–458.

CHERNOZHUKOV, V., H. HONG, AND E. TAMER (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75(5), 1243–1284.

CLEMENT, E., C. GOURIEROUX, AND A. MONFORT (2000): "Econometric Specification of the Risk Neutral Valuation Model," *Journal of Econometrics*, 94(1-2), 117–143.

COCHRANE, J., AND J. SAÁ-REQUEJO (2000): "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets," *Journal of Political Economy*, 108(1), 79.

COX, J. C., J. INGERSOLL, JONATHAN E., AND S. A. ROSS (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53(2), 385–408.

DAVIES, R. B. (1977): "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64(2), 247–254.

DAVIES, R. B. (1987): "Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74(1), 33–43.

DELBAEN, F., AND W. SCHACHERMAYER (1994): "A General Version of the Fundamental Theorem of Asset Pricing," *Mathematische Annalen*, V300(1), 463–520.

DUFFIE, D. (2001): *Dynamic Asset Pricing Theory.* 3rd ed. Princeton: Princeton University Press.

FAMA, E. F., AND K. R. FRENCH (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33(1), 3–56.

——— (1996): "Multifactor Explanations of Asset Pricing Anomalies," *The Journal of Finance*, 51(1), 55–84.

HALL, P., AND B. JING (1996): "On Sample Reuse Methods for Dependent Data," *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(4), 727–737.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–1054.

HÄRDLE, W., J. HOROWITZ, AND J.-P. KREISS (2003): "Bootstrap Methods for Time Series.," *International Statistical Review*, 71(2), 435–459.

HARRISON, J. M., AND D. M. KREPS (1979): "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20(3), 381–408.

HARRISON, J. M., AND S. R. PLISKA (1981): “Martingales and Stochastic Integrals in the Theory of Continuous Trading,” *Stochastic Processes and their Applications*, 11(3), 215–260.

JACKWERTH, J. (2000): “Recovering Risk Aversion from Option Prices And Realized Returns,” *Review of Financial Studies*, 13(2), 433–451.

KAIDO, H., AND H. WHITE (2008): “A Two-Stage Procedure for Partially Identified Models,” Discussion Paper, University of California, San Diego.

KIM, T.-H., AND H. WHITE (2004): “On More Robust Estimation of Skewness and Kurtosis,” *Finance Research Letters*, 1(1), 56–73.

KITAMURA, Y., AND M. STUTZER (1997): “An Information-Theoretic Alternative to Generalized Method of Moments Estimation,” *Econometrica*, 65(4), 861–874.

KRISTENSEN, D. (2008): “Pseudo-Maximum Likelihood Estimation in Two Classes of Semiparametric Diffusion Models,” Working Paper, Columbia University.

LEWIS, K. K. (1995): “Puzzles in International Financial Markets,” in G. M. Grossman, and K. Rogoff (eds.) *Handbook of International Economics*, vol. 3, New York: Elsevier, pp. 1913–1971.

LIEW, J., AND M. VASSALOU (2000): “Can Book-to-Market, Size, and Momentum be Risk Factors that Predict Economic Growth?” *Journal of Financial Economics*, 57(2), 221–245.

LINTNER, J. (1965): “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *The Review of Economics and Statistics*, 47(1), 13–37.

LIU, X., AND Y. SHAO (2003): “Asymptotics for Likelihood Ratio Tests under Loss of Identifiability,” *The Annals of Statistics*, 31(3), 807–832.

LUCAS, R. E. (1978): “Asset Prices in an Exchange Economy,” *Econometrica*, 46(6), 1429–1445.

MANSKI, C. F. (2003): *Partial Identification of Probability Distributions*. Berlin: Springer.

MOLCHANOV, I. (2005): *Theory of Random Sets*. Berlin: Springer.

PROTTER, P. E. (2005): *Stochastic Integration and Differential Equations*, 3rd ed. Berlin: Springer.

ROMANO, J. P., AND A. SHAIKH (2006): "Inference for the Identified Set in Partially Identified Econometric Models," Working Paper, University of Chicago.

——— (2008): "Inference for Identifiable Parameters in Partially Identified Econometric Models," *Journal of Statistical Planning and Inference*, Special Issue in Honor of T. W. Anderson, forthcoming.

ROSENBERG, J. V., AND R. F. ENGLE (2002): "Empirical Pricing Kernels," *Journal of Financial Economics*, 64(3), 341–372.

SANTOS, A. (2007): "Inference in Nonparametric Instrumental Variables with Partial Identification," Working Paper, University of California, San Diego.

SHARPE, W. F. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, 19(3), 425–442.

WILLIAMS, R. J. (2006): *Introduction to the Mathematics of Finance*. New York: American Mathematical Society.