Econometric Evaluation of Asset Pricing Models

Lars Peter Hansen
University of Chicago, NBER, and NORC

John Heaton
MIT and NBER

Erzo G. J. Luttmer
Northwestern University

In this article we provide econometric tools for the evaluation of intertemporal asset pricing models using specification-error and volatility bounds. Weformulate analog estimators of these bounds, give conditions for consistency, and derive the limiting distribution of these estimators. The analysis incorporates market frictions such as short-sale constraints andproportional transactions costs. Among several applications we show how to use the methods to assess specific asset pricing models and to provide nonparametric characterizatons of asset pricing anomalies.

In this article we provide statistical methods for assessing asset pricing models using specification-error and volatility bounds. The statistical procedures can account for market frictions due to transactions costs or short-sale constraints and are easier to interpret

We thank Craig Burnside, Bo Honoré, Andrew Lo, Marc Roston, Whitney Newey, José Scheinkman, Jean-Luc Vila, Jiang Wang, Amir Yaron, and especially Ravi Jagannathan, Richard Green (the editor), and an anonymous referee for helpful comments and discussions. We also received valuable remarks from seminar participants at the 1992 meetings of the Society of Economic Dynamics and Control, at the 1992 NBER Summer Institute, and at Cornell, Duke, LSE, and Waterloo universities. Finally, we gratefully acknowledge the financial assistance of the International Financial Services Research Center at MIT. (Heaton), the National Science Foundation (Hansen and Heaton), and the Sloan Foundation (Heaton and Luttmer). Address correspondence to John Heaton, Sloan School of Management, E52-435, MIT, 50 Memorial Drive, Cambridge, MA 02142.

The Review of Financial Studies Summer 1995 Vol. 8, No. 2, pp. 237-274

than are standard tests of asset pricing models. For the most part these methods are quite easy to implement, even when market frictions are considered. They are designed to provide a better understanding of the statistical failures of some popular asset pricing models and to offer guidance in improving these models.

Models of asset pricing with frictionless markets imply that asset prices can be represented by a stochastic discount factor or pricing kernel. A stochastic discount factor "discounts" payoffs in each state of the world and, as a consequence, adjusts the price according to the riskiness of the payoff. For example, in the capital asset pricing model (CAPM) the discount factor is given by a constant plus a scale multiple of the return on the market portfolio. In the consumption-based CAPM, the discount factor is given by the intertemporal marginal rate of substitution of an investor.

The implications of particular models with observable (up to a finite number of parameters) stochastic discount factors are often tested by looking directly at the average "pricing errors" of the models. Formal statistical tests are performed using a time series of portfolio payoffs and prices by examining whether the sample analogs of the average and predicted prices are significantly different from each other. For examples of this type of procedure, see Hansen and Singleton (1982), Brown and Gibbons (1985), MacKinlay and Richardson (1991), and Epstein and Zin (1991).

While tests such as these can be informative, it is often difficult to interpret the resulting statistical rejections. Further, these tests are not directly applicable when there are market frictions such as transactions costs or short-sale constraints. Extending the models to allow for such frictions entails inequalities instead of the pricing equalities that prevail in frictionless market models (e.g., see Prisman 1986 and Jouini and Kallal 1993). Finally, these tests cannot be used when the candidate discount factor depends on variables unavailable to the econometrician.

As an alternative to considering the average pricing errors of a model, we consider a different set of tests and diagnostics using the specification-error bounds of Hansen and Jagannathan (1994) and the volatility bounds of Hansen and Jagannathan (1991). We also consider extensions of these tests and diagnostics, developed by He and Modest (1993) and Luttmer (1994), that handle transactions costs, short-sale restrictions, and other market frictions. We develop an econometric method to provide consistent estimators of the specification-error and volatility bounds and an asymptotic distribution theory that is easy to implement and that can be used to make statistical inferences.

Among other things, the results in this article allow one to (i) test whether a specific model of the stochastic discount factor satisfies the volatility bounds implicit in asset-market returns, (ii) compare the information about the means and standard deviations of discount factors contained in different sets of asset returns, and (iii) test hypotheses about the size of possible pricing errors of misspecified asset pricing models.

While Burnside (1994) and Cecchetti, Lam, and Mark (1994) have devised tests of models of the discount factor along the lines of application (i), their tests are based on a different parameterization of the volatility bound. Our parameterization yields tests that are simpler to implement and can accommodate market frictions in a straightforward manner. In regard to application (ii), our results permit these comparisons among data sets to be made independent of a specific stochastic discount factor model. Our motivation for application (iii) is to shift the focus of statistical analyses of asset pricing models away from whether the models are correctly specified and toward measurement of the extent to which they are misspecified.

The rest of the article is organized as follows. In Section 1 we review the specification and volatility bounds of Hansen and Jagannathan (1991, 1994), He and Modest (1993), and Luttmer (1994). We show formally that the volatility bound can be viewed as a special case of the specification-error bound. This permits us to develop the underlying econometric tools in a unified way. In Section 2 we provide consistency and asymptotic distribution results for estimators of the bounds. In Section 3 we present two applications of our results of Section 2, each of which can be read independently. Section 3.1 shows how to use the volatility bounds to test models of the discount factor. In Section 3.2 we extend the distribution theory of specification-error bound to the case where there are parameters of the discount factor proxy that are unknown and must be estimated.

Section 4 discusses the limiting distribution of the parameters underlying the bounds both with and without market frictions. Among other things, these results can be used to determine whether the volatility bound is degenerate or more generally whether additional security market data sharpen the bound. Finally, Section 5 describes some extensions and provides some concluding remarks.

1. General Model and Bounds

Our starting point is a model in which asset prices are represented by a stochastic discount factor or pricing kernel. To accommodate security market pricing subject to transactions costs, we permit there to be short-sale constraints for a subset of the securities. Although a short-sale constraint is an extreme version of a transactions. cost, other proportional transactions costs such as bid-ask spreads can also be handled with this formalism. This is done as in Foley (1970), Jouini and Kallal (1993), and Luttmer (1994) by constructing two payoffs according to whether a security is purchased or sold. A short-sale constraint is imposed on both artificial securities to enforce the distinction between a buy and a sell, and a bid-ask spread is modeled by making the purchase price higher than the sale price.

Suppose the vector of security market payoffs used in an econometric analysis is denoted x. The vector x is used to generate a collection of payoffs formed using portfolio weights in a closed convex cone C of $\mathbb{R}^n$:

$$P \equiv \{p : p = \alpha' x \text{ for some } \alpha \in C\}. \tag{1}$$

The cone C is constructed to Incorporate all of the short-sale constraints imposed in the econometric investigation. If there are no market frictions, then C is $\mathbb{R}^n$. More generally, partition x into two components: $x' = [x^{n\prime}, x^{s\prime}]$ where x^n contains the k components not subject to short-sale constraints and x^s contains the ℓ components subject to short-sale constraints. Then the cone C is formed by taking the Cartesian product of $\mathbb{R}^k$ and the nonnegative orthant of $\mathbb{R}^\ell$.

Let q denote the random vector of prices corresponding to the vector x of securities payoffs. These prices are observed by investors at the time that assets are traded and are permitted to be random because the prices may reflect conditioning information available to the investors. In the absence of short-sales constraints, prices can be represented by

$$q = E(mx \mid \mathcal{F}) \tag{2}$$

where m is a stochastic discount factor and $\mathcal{F}$ is the information set available to investors at the time of trade. Since it is difficult to model empirically the conditioning information available to investors, we instead work with the average or expected value of Equation (2):

$$Eq - Emx = 0. \tag{3}$$

Some conditioning information can be incorporated in the usual way by multiplying the original set of payoffs and prices by random variables in the conditioning information of economic agents.

More generally in the case of market frictions, x is partitioned in the manner described previously and the pricing is represented by

$$\begin{aligned} Eq^n - Emx^n &= 0 \\ Eq^s - Emx^s &\geq 0. \end{aligned} \tag{4}$$

For notational simplicity we write Relations (4) as

$$Eq - Emx \in C^* \tag{5}$$

where the elements of C^* are of the form $(0, \beta')'$, β nonnegative.[1] In the absence of frictions, we take C^* to contain only the zero vector so that Relation (5) encompasses Equation (3). The inequality restriction emerges because pricing the vector of payoffs x^s subject to short-sale constraints must allow for the possibility that these constraints bind and hence contribute positively to the market price vector.

1.1 Maintained assumptions

There are three restrictions on the vector of payoffs and prices that are central to our analysis. The first is a moment restriction, the second is equivalent to the absence of arbitrage on the space of portfolio payoffs, and the third eliminates redundancy in the securities.

For pricing relation (5) to have content, we maintain

Assumption 1.1. $E|x|^2 < \infty, E|q| < \infty.$

Assumption 1.2 *For any* $\alpha \in C$, $\alpha' Eq > 0$ *if* $\alpha' x \geq 0$ *and* $Prob\{\alpha' x > 0\} > 0$ *Further for* $\alpha \in C$, $\alpha' Eq \geq 0$ *if* $\alpha' x = 0$.

Recall that C is the Cartesian product of $\mathbb{R}^k$ and the nonnegative orthant of $\mathbb{R}^\ell$, which captures the short-sale restrictions on some of the securities. Assumption 1.2 is a statement of the principle of no-arbitrage applied to expected prices and modified to account for the fact that C need not be linear [e.g., see Kreps (1981), Prisman (1986), Jouini and Kallal (1993), and Luttmer (1994)]. It guarantees that there exists a nonnegative stochastic discount factor m with finite second moment such that Relation (5) holds. Jouini and Kallal (1993) discuss additional assumptions that imply the existence of a positive discount factor satisfying Relation (5); however, in our analysis we consider only nonnegative discount factors.

Next we limit the construction of x by ruling out redundancies in the securities.

Assumption 1.3. *If* $\alpha' x = \alpha^{*\prime} x$ *and* $\alpha' Eq = \alpha^{*\prime} Eq$ *for some* α *and* α^* *in* C, *then* $\alpha = \alpha^*$.

In the absence of transaction costs, Assumption 1.3 precludes the possibility that the second moment matrix of x is singular. Otherwise, there would exist a nontrivial linear combination of the payoff vector x, that is zero with probability one. In light of Relation (5), the (expected)

[1] Formally C^* is the dual cone of C.

price of this nontrivial linear combination would have to be zero, violating Assumption 1.3. To accommodate securities whose purchase price differs from the sale price, we permit the second moment matrix of the composite vector x to be singular. Assumption 1.3 then requires that distinct portfolio weights used to construct the same payoff must have distinct expected prices.[2]

1.2 Minimum-distance problems

Two problems underlie most of our analysis. Let $\mathcal{M}$ denote the set of all random variables with finite second moments that satisfy Equation (5), and let $\mathcal{M}^+$ be the set of all nonnegative random variables in M. Recall that Assumption 1.2 implies that there is a nonnegative discount factor that satisfies Relation (5) so that both sets are nonempty. Let y denote some "proxy" variable for a stochastic discount factor that, strictly speaking, does not satisfy relations (5). Following Hansen and Jagannathan (1994), we consider the following two ad hoc least-squares measures of misspecification:[3]

$$\hat{\delta}^2 = \min_{m \in \mathcal{M}} E[(y - m)^2], \tag{6}$$

and

$$\tilde{\delta}^2 = \min_{m \in \mathcal{M}^+} E[(y - m)^2]. \tag{7}$$

Clearly, the specification-error bound implied by Problem (7) is no smaller than that implied by Problem (6), since it is obtained using a smaller constraint set. The solutions to (6) and (7) are the objects we are interested in estimating and making inferences about. Sections 2 and 3 provide large-sample justifications for the solutions to sample counterparts to these optimization problems.

By setting the proxy y to zero, the specification error problems collapse to finding bounds on the second moment of stochastic dis-

---

[2] A weaker version of this restriction would replace Eq by q. In effect, Assumption 1.3 does more than eliminate redundant securities. It also precludes cases In which distinct portfolio weights give rise to the same payoff, with possibly different prices but the same expected prices.

[3] Hansen and Jagannathan (1994) showed that the least-squares distance between a proxy and the set $\mathcal{M}$ of (possibly negative) stochastic discount factors has an alternative pricing-error interpretation. Formally, the pricing-error interpretation for the least-squares problem (6) Is

$$\hat{\delta} = \inf_{m \in \mathcal{M}} \sup_{\substack{p \in P \\ Ep^2 = 1}} |Emp - Eyp|,$$

and for (7) is

$$\tilde{\delta} = \inf_{m \in \mathcal{M}^+} \sup_{\substack{p \in H \\ Ep^2 = 1}} |Emp - Eyp|$$

where H is the set of payoffs on hypothetical derivative claims.

count factors as constructed by Hansen and Jagannathan (1991), He and Modest (1993), and Luttmer (1994). In particular, the bounds derived in Hansen and Jagannathan (1991) are obtained by setting y to zero and solving Problems (6) and (7) when there are no short-sale constraints imposed (when C is set to $\mathbb{R}^n$); the bound derived in He and Modest (1993) is obtained by solving Problem (6) for y set to zero; and the bound derived by Luttmer (1994) is obtained by solving Problem (7) for y set to zero. These second moment bounds will subsequently be used in deriving feasible regions for means and standard deviations of stochastic discount factors.

In solving the least-squares problems (6) and (7) and in develop-&g econometric methods associated with those problems, it is most convenient to study the conjugate maximization problems. They are given by

$$\hat{\delta}^2 = \max_{\alpha \in C} \{Ey^2 - E[(y - x'\alpha)^2] - 2\alpha' Eq\}, \tag{8}$$

and

$$\tilde{\delta}^2 = \max_{\alpha \in C} \{Ey^2 - E[(y - x'\alpha)^{+2}] - 2\alpha' Eq\} \tag{9}$$

where the notation h^+ denotes $\max\{h, 0\}$. The conjugate problems are obtained by introducing Lagrange multipliers on the pricing constraints in Relation (5) and exploiting the familiar saddle point property of the Lagrangian. The *a's* then have interpretations as the multipliers on the pricing constraints.

The conjugate problems in Problems (8) and (9) are convenient because the choice variables are finitedimensional vectors, whereas the choice variables in the original least-squares problems (6) and (7) are random variables that reside in possibly infinite-dimensional constraint sets. In fact, optimization problem (8) is a standard quadratic programming problem. The specifications of the conjugate problems are justified formally in Hansen and Jagannathan (1991, 1994) and Luttmer (1994). In Section 2 we develop the properties of estimators of $\hat{\delta}^2$ and $\tilde{\delta}^2$ based on time series sample analogs to problems (8) and (9). In so doing we rely on several important proprieties of the solutions to problems (8) and (9) and on an additional identification assumption.

Notice that the criteria for the maximization problems are concave in a and that the first-order conditions for the solutions are given by

$$Eq - E[(y - x'\hat{\alpha})x] \in C^* \tag{10}$$

in the case of problem (8) and

$$Eq - E[(y - x'\tilde{\alpha})^+ x] \in C^* \tag{11}$$

in the case of problem (9), along with the respective complementary slackness conditions. Interpreting the first-order conditions for these problems, observe that associated with a solution to problem (8) is a random variable $\hat{m} = (y - x'\hat{\alpha})$ in $\mathcal{M}$ and associated with a solution to problem (9) is a nonnegative random variable $\tilde{m} = (y - x'\tilde{\alpha})^+$ in $\mathcal{M}^+$. These random variables are the unique (up to the usual equivalence class of random variables that are equal with probability one) solutions to the original least-squares problems (6) and (7).

Consistency of the estimators of $\hat{\delta}^2$ and $\tilde{\delta}^2$ relies upon the fact that the sets of solutions to problems (8) and (9) are compact. Compactness in the case of problem (8) is easily established. Since Assumption 1.3 eliminates redundant securities and the random variable $(y - x'\hat{\alpha})$ is uniquely determined, 'the solution $\hat{\alpha}$ to conjugate problem (8) is also unique. This follows because the value of the criterion must be the same for all solutions, implying that they all must have the same expected price $\hat{\alpha}'Eq$. 'The solution to conjugate problem (9) may not be unique, however. In this case the truncated random variable $(y - x'\tilde{\alpha})^+$ is uniquely determined, as is the expected price $\tilde{\alpha}'Eq$. On the other hand, the random variable $(y - x'\tilde{\alpha})$ is not necessarily unique, so we cannot exploit Assumption 1.3 to verify that the solution $\tilde{\alpha}$ is unique. The set of solutions is convex because of the concavity of the criterion and the convexity of the constraint set. As is shown in Appendix A, it is also compact.

As is typical in asymptotic distribution theory, in Section 2 we will need an identification restriction that there is a unique solution to the conjugate problems, except when all prices are constant. Since the set of solutions is convex, local uniqueness implies global uniqueness. To display a sufficient condition for local uniqueness, let x^*, denote the component of the composite payoff vector x for which the pricing relation is satisfied with equality

$$E\tilde{m}x^* = Eq^* \tag{12}$$

where q^* is the corresponding price vector. Notice that, in addition to x^n, x^* may contain elements of x^s. Let $1_{\{\tilde{m}>0\}}$ be the indicator function for the event $\{\tilde{m} > 0\}$. A sufficient condition for local uniqueness is that

Assumption 1.4. $Ex^*x^{*\prime}1_{\{\tilde{m}>0\}}$ *is nonsingular.*

To see why this is a valid sufficient condition, observe that, from the complementary slackness conditions, we know the multipliers $\tilde{\alpha}$ are zero whenever the pricing constraints are satisfied with a strict inequality. As a result $\tilde{m}$ is given by $(y - x^{*\prime}\tilde{\beta})^+$ for some vector $\tilde{\beta}$, and, consequently,

$$Eq^* = Ey\mathbf{1}_{\{\tilde{m}>0\}} - E(x^* x^{*\prime} \mathbf{1}_{\{\tilde{m}>0\}})\tilde{\beta}. \tag{13}$$

When the matrix $E(x^* x^{*\prime} \mathbf{1}_{\{\tilde{m}>0\}})$ is nonsingular, we can solve (13) for $\tilde{\beta}$.

1.3 Volatility bounds and restrictions on means

The second moment bounds described in the previous subsection can be converted into standard deviation bounds via the formulas:

$$\begin{aligned} \hat{\sigma} &= [\hat{\delta}^2 - (Em)^2]^{1/2} \\ \tilde{\sigma} &= [\tilde{\delta}^2 - (Em)^2]^{1/2} \end{aligned} \tag{14}$$

where $\hat{\delta}^2$ and $\tilde{\delta}^2$ are constructed by setting the proxy to zero. *Em* is equal to the average price of the unit payoff when trade in this payoff is not subject to transaction costs. If there are transactions costs or if no data are available on the price of a conditionally riskless payoff, then *Em* cannot be identified. In these circumstances, volatility bounds can still be obtained for each choice of *Em* by adding a unit payoff to *P* (augmenting *x* with a 1) and assigning a price of *Em* to that payoff (augmenting *Eq* with *Em).* In forming the augmented cone, there should be no short-sale constraints imposed on the additional security. Mean-specific volatility bounds can then be obtained using (8), (9), and (14).

Although *Em* may not be identified, the principle of no-arbitrage does put bounds on the admissible values of Em: $Em \in [\lambda_0, v_0]$ where λ_0 is the lower arbitrage bound and v_0 the upper arbitrage bound. These bounds are computed using formulas familiar from derivative claims pricing:

$$\lambda_0 \equiv -\inf\{\alpha' Eq : \alpha \in C \text{ and } \alpha' x \geq -1\} \tag{15}$$

and

$$v_0 \equiv \inf\{\alpha' Eq : \alpha \in C \text{ and } \alpha' x \geq 1\}. \tag{16}$$

While λ_0 is always well defined via Formula (15), v_0 may not be because there may not exist a payoff in *P* that dominates a unit payoff. In such circumstances, we define v_0 to be $+\infty$. In Section 2 we show how to consistently estimate λ_0 and v_0. Consistent estimation of these bounds is important since the standard deviation bound $\tilde{\sigma}$ is infinite for choices of *Em* outside these bounds.

2. Estimation of the Bounds

In this section we develop consistency and asymptotic distribution results for the specification-error bounds presented in Section 1. A key presumption underlying our analysis is that the data on asset payoffs and prices are replicated over time in some stationary fashion. That is, associated with the composite vector $(x', q', y)'$ is a stochastic process $\{(x_t', q_t', y_t)'\}$ whose sequence of empirical distributions approximate the joint distribution of $(x', q', y)'$. We denote integration with respect to the empirical distribution for sample size T as Σ_T More precisely, for any z that is a (Borel measurable) function of (x', q', y) with a finite first moment, we will approximate Ez by $\Sigma_T z$ where

$$\Sigma_T z \equiv (1/T)\Sigma_{t=1}^{T} z_t. \tag{17}$$

Among other things, we require that this approximation becomes arbitrarily good as the sample size T gets large. That is, we presume that $\{z_t\}$ obeys a Law of Large Numbers. A sufficient condition for this is

Assumption 2.1. *The composite process $\{(x_t', q_t', y_t)\}$ is stationary and ergodic.*

Under this assumption, we can think of (x', q', y) as (x_0', q_0', y_0).[4]

To estimate the specification-error bounds, we suppose that a sample of size T is available and that the empirical distribution implied by these data is used in place of the population distribution. [Thus, we are applying the Analogy Principle of Goldberger (1968) and Manski (1988)]. We introduce two random functions $\hat{\phi}$ and $\tilde{\phi}$:

$$\hat{\phi}(\alpha) \equiv y^2 - (y - \alpha' x)^2 - 2\alpha' q, \tag{18}$$

and

$$\tilde{\phi}(\alpha) \equiv y^2 - (y - \alpha' x)^{+2} - 2\alpha' q. \tag{19}$$

The sample analog estimators for $\hat{\delta}^2$ and $\tilde{\delta}^2$ are given by

$$(\hat{d}_T)^2 = \max_{\alpha \in C} \Sigma_T[\hat{\phi}(\alpha)] \tag{20}$$

and

$$(\tilde{d}_T)^2 = \max_{\alpha \in C} \Sigma_T[\tilde{\phi}(\alpha)]. \tag{21}$$

---

[4] Assumption 2.1 could be weakened in a variety of ways, but it is maintained for pedagogical simplicity. More generally, we might imagine that the process $\{(x_t', q_t', y_t)\}$ is asymptotically stationary, where the convergence to the stationary distribution is sufficiently fast to ensure that the Law of Large Numbers applies to averages of the form (17). In this case, the joint distribution of (x', q', y) is given by the stationary limit point of the process $\{(x_t', q_t', y_t)\}$.

2.1 Consistent estimation of the specification-error bounds

We first establish the statistical consistency of the estimator sequences $\{\hat{d}_T\}$ and $\{\tilde{d}_T\}$:

Proposition 2.1. *Under Assumptions 1.1-1.3 and 2.1, $\{\hat{d}_T\}$ and $\{\tilde{d}_T\}$ converge almost surely to $\hat{\delta}$ and $\tilde{\delta}$, respectively.*

The proof of this proposition is given in Appendix A and is not complicated by the presence of short-sale constraints. The basic idea is that the population and sample criterion functions for the conjugate problems are concave and the sets of maximizers are convex. By Assumptions 1.1 and 2.1, the criterion functions converge pointwise (in a and β) almost surely to the population criterion functions introduced in Section 1.2. In light of the concavity of the criterion functions, this convergence is uniform on compact sets almost surely [e.g., see Rockafellar (1970)]. Finally, since the sets of maximizers of the limiting criterion functions are compact, for sufficiently large T one can find a compact set such that the maximizers of the sample and population criteria are contained in that compact set [e.g., see Hildenbrand (1974) and Haberman (1989)]. Hence, the conclusion follows from the uniform convergence of the criteria on a compact set.

2.2 Asymptotic distribution of the estimators of the bounds

We consider next the limiting distribution of the analog estimator sequences of the specification-error bounds. Our ability to express the objects of interest as solutions to the conjugate problems permits us to obtain results very similar to those in the literature on using likelihood ratios as devices for model selection in environments when models are possibly misspecified [e.g., see Vuong (1989)].. We show that, when the specification error bounds are positive, we obtain a limiting distribution that is equivalent to the one obtained by ignoring parameter estimation, and, when the specification error bound is zero; the limiting distribution is degenerate. [See Theorem 3.3 of Vuong (1989), p. 307, for the corresponding result for likelihood ratios.]

Let $\hat{\alpha}_T$ be a maximizer of $\Sigma_T\hat{\phi}$, $\hat{\alpha}$ a maximizer of $E\hat{\phi}$, $\tilde{\alpha}_T$ a maximizer of $\Sigma_T\tilde{\phi}$, and $\tilde{\alpha}$ be a maximizer of $E\tilde{\phi}$. To study the limiting behavior of the estimators, we use the decompositions

$$\sqrt{T}[(\hat{d}_T)^2-\hat{\delta}^2]=\sqrt{T}\Sigma_T[\hat{\phi}(\hat{\alpha}_T)-\hat{\phi}(\hat{\alpha})]+\sqrt{T}\Sigma_T[\hat{\phi}(\hat{\alpha})-E\hat{\phi}(\hat{\alpha})], \tag{22}$$

and

$$\sqrt{T}[(\tilde{d}_T)^2-\tilde{\delta}^2]=\sqrt{T}\Sigma_T[\tilde{\phi}(\tilde{\alpha}_T)-\tilde{\phi}(\tilde{\alpha})]+\sqrt{T}\Sigma_T[\tilde{\phi}(\tilde{\alpha})-E\tilde{\phi}(\tilde{\alpha})]. \tag{23}$$

We make the following assumptions:

Assumption 2.2.

$$\left\{\sqrt{T}\Sigma_T\left[\begin{array}{c}\hat{\phi}(\hat{\alpha}) - E\hat{\phi}(\hat{\alpha}) \\ [(\hat{m}x - q) - E(\hat{m}x - q)]\end{array}\right]\right\}$$

converges in distribution to a normally distributed random vector with mean zero and covariance matrix $\hat{V}$.

Assumption 2.3.

$$\left\{\sqrt{T}\Sigma_T\left[\begin{array}{c}\tilde{\phi}(\tilde{\alpha}) - E\tilde{\phi}(\tilde{\alpha}) \\ [(\tilde{m}x - q) - E(\tilde{m}x - q)]\end{array}\right]\right\}$$

converges in distribution to a normally distributed random vector with mean zero and covariance matrix $\tilde{V}$.

More primitive assumptions that imply the central limit approximations underlying Assumptions 2.2 and 2.3 are given by Gordin (1969) and Hall and Heyde (1980).

Let u denote a *selection vector* with a one in its first position followed by $k+\ell$ zeros: The limiting distributions for the specification-error bound estimators are given by Proposition 2.2.

Proposition 2.2. *Suppose that* $\hat{\delta} \neq 0$ *and* $\tilde{\delta} \neq 0$. *Under Assumptiom 1.1-1.3, 2.1, and 2.2,* $\sqrt{T}[\hat{d}_T - \hat{\delta}]\}$ *converges to a normally distributed random vector with mean zero and variance* $u'\hat{V}u/(4\hat{\delta}^2)$.

Under Assumptions 1.1-1.4, 2.1, and 2.3, $\{\sqrt{T}[\tilde{d}_T - \tilde{\delta}]\}$ *converges in distribution to a normally distributed random vector with mean zero and variance* $u'\tilde{V}u/(4\tilde{\delta}^2)$.

As Proposition 2.2 indicates, the limiting distributions for the maximized values depend only on the second terms of the decompositions in Equations (22) and (23). In other words, the effect of replacing the unknown population maximizers by the sample maximizers in the sample criterion functions is negligible. As a consequence, the presence of short-sale constraints does not complicate the limiting distribution.

To see why the asymptotic distribution in Proposition 2.2 is not affected by sampling error in the estimation of the multipliers, consider the case of the sequence $\{(\tilde{d}_T)^2\}$. By the concavity of $\tilde{\phi}$, we have the following gradient inequalities:

$$\begin{aligned}\tilde{\phi}(\tilde{a}_T) - \tilde{\phi}(\tilde{\alpha}) &\leq (\tilde{m}x - q)\cdot(\tilde{a}_T - \tilde{\alpha}) \\ &= [(\tilde{m}x - q) - E(\tilde{m}x - q)]\cdot(\tilde{a}_T - \tilde{\alpha}) \\ &\quad + E(\tilde{m}x - q)\cdot(\tilde{a}_T - \tilde{\alpha}).\end{aligned} \tag{24}$$

However, it follows from the first-order conditions (including the complementary slackness conditions) for the population conjugate problem that

$$E(\tilde{m}x - q) \cdot (\tilde{a}_T - \tilde{\alpha}) = E(\tilde{m}x - q) \cdot \tilde{a}_T \leq 0. \tag{25}$$

The inequality in (25) is obtained because $E(q - \tilde{m}x)$ is in C^* while $\tilde{a}_T$ is constrained to be in C. Combining Relations (24) and (25) we have that

$$\begin{aligned} 0 &\leq \sqrt{}\Sigma_T[\tilde{\phi}(\tilde{a}_T) - \tilde{\phi}(\tilde{\alpha})] \\ &\leq \sqrt{T}\Sigma_T[(\tilde{m}x - q) - E(\tilde{m}x - q)] \cdot (\tilde{a}_T - \tilde{\alpha}). \end{aligned} \tag{26}$$

Since, by Assumption 2.3, the sample counterparts of the pricing errors obey a central limit theorem, $\sqrt{T}\Sigma_T[\tilde{\phi}(\tilde{a}_T) - \tilde{\phi}(\tilde{\alpha})]\}$ converges in probability to zero if the maximizers can be chosen so that $\{(\tilde{a}_T - \tilde{\alpha})\}$ converges almost surely to zero. This latter convergence can be demonstrated by exploiting the concavity of the population criterion function and the convexity of the constraint set [e.g., see the discussion on page 1635 of Haberman (1989) and in Appendix A].

To use Proposition 2.2 in practice requires consistent estimation of $u'\hat{V}u$ or $u'\tilde{V}u$. Consider the case of $u'\tilde{V}u$. For each T form the scalar sequence $\{\tilde{\phi}_t(a_T) : t = 1, 2, 3, \ldots T\}$ and use one of the frequency zero spectral density estimators described by Newey and West (1987) or Andrews (1991), for example.

As is shown in Appendix A, when the price vector q is a vector of real numbers (degenerate random variables), the asymptotic distribution for $T[\tilde{d}_T - \tilde{\delta}]\}$ remains valid even when the population version of the conjugate maximum problem fails to have a unique solution (Assumption 1.4 is violated). In this case, the lack of identification of the parameter vector $\tilde{\alpha}$ does not alter the distribution theory for the specification-error bound. While this special case is of considerable interest, it rules out the possibility of using conditioning information to form synthetic payoffs with random prices as described in Section 1.

Notice that, if $\hat{\delta} = 0$ or $\tilde{\delta} = 0$, Proposition 2.2 breaks down. This occurs if y is a valid stochastic discount factor, in which case the solutions to the population conjugate problems are $\hat{\alpha} = \tilde{\alpha} = 0$. As a consequence, $\hat{\phi}_t(\hat{\alpha})$ and $\tilde{\phi}_t(\tilde{\alpha})$ are both identically zero, giving rise to a degenerate limiting distribution for $\{\sqrt{T}(\hat{d}_T)^2\}$ and $\sqrt{T}(\tilde{d}_T)^2\}$ Our results in 'Section 4 on the convergence of the parameter estimators can be used to establish that the rate of convergence of $\{(\hat{d}_T)^2\}$ and $\{(\tilde{d}_T)^2\}$ is T, and the limiting distribution is given by a weighted sum of chi-squared distributions (see Vuong 1989). As a result $\{\hat{d}_T\}$ and

$\{\bar{d}_T\}$ converge at the rate $\sqrt{T}$, although the limiting distribution is not normal.

2.3 Consistent estimation of the arbitrage bounds

As we discussed in Section 1, the second moments bounds can be converted into standard deviation bounds if the mean of m is known or if it can be estimated using the price of a risk-free asset. When Em is not known, it must be prespecified. Let v be the hypothesized mean of m when a risk-free asset is not available. Proposition 2.1 can be applied to establish the consistency of the second moment bound estimators for each admissible price assignment v. In the case of $\tilde{\delta}^2$, for the price assignment to be admissible, it must not induce arbitrage opportunities onto the augmented collection of asset payoffs and prices. Any price (mean) assignment in the open interval (λ_0, v_0) is admissible in this sense.

The final question we explore in this section is whether the arbitrage bounds, λ_0 and v_0 given in Formulas (15) and (16), can be consistently estimated using the sample analogs:

$$\begin{aligned} \ell_T \equiv & -\inf\{\alpha'\Sigma_T q : \alpha \in C \text{ and} \\ & \alpha' x_t \geq -1 \text{ for all } t = 1, 2, \ldots, T\} \end{aligned} \tag{27}$$

and

$$\begin{aligned} u_T \equiv & \inf\{\alpha'\Sigma_T q : \alpha \in C \text{ and} \\ & \alpha' x_t \geq 1 \text{ for all } t = 1, 2, \ldots, T\}. \end{aligned} \tag{28}$$

The estimated upper arbitrage bound u_T is always finite when there is a payoff on a limited liability security that is never observed to be zero in the sample. Our estimated range of the admissible values for the (average) price of a unit payoff and hence mean of m is $[\ell_T, u_T]$. Notice that these bounds can be computed by solving simple linear programming problems. In Appendix A we prove

Proposition 2.3 *Under Assumptions 1.1-1.3 and 2.1, $\{\ell_T\}$ converges to λ_0 almost surely. If v_0 is finite, then $\{u_T\}$ converges to v_0 almost surely; and, if $v_0 = +\infty$, then (u_T) diverges to $+\infty$ almost surely.*

2.4 Consistent estimation of the feasible region of means and standard deviations

We now consider consistent estimation of the set of feasible means and standard deviations of stochastic discount factors. Previously, we showed that, for a given mean of the stochastic discount factor, the standard deviation bound can be consistently estimated. However, the mean of the stochastic discount factor typically is not known. As a result it is important to understand the sense in which the entire feasible region can be approximated. Such a region can be computed with or without imposing the no-arbitrage restriction that the stochastic discount factors be positive. Let S_0 denote the feasible region without positivity and S_0^+ the (closure) of the feasible region with positivity. Similarly, let S_T and S_T^+ denote the sample counterparts. The question we now turn to is in what sense are S_T and S_T^+ good approximations to S_0 and S_0^+?

When there is a unit payoff, all four feasible regions are vertical rays in mean and standard deviation space because the (average) price of this payoff is the mean discount factor. In this case the points of origin of the rays S_0 and S_0^+ can be estimated consistently by the points of origin of the corresponding rays S_T and S_T^+.

In the more usual case, when data on the price of a unit payoff are not available, matters are a little more complicated. The feasible regions are no longer vertical rays but instead are unions of such rays resulting in convex sets with nonempty interiors. The boundaries of these sets can be represented as (possibly extended) real-valued functions of the ordinate (hypothetical mean), and our previous analysis implies pointwise (in the mean) convergence of the sample analog functions to their population counterparts. This result implies uniform convergence of the sample analog functions in following sense.

Since the lower and upper-arbitrage bounds can be consistently estimated, for large enough T, the sample analog functions under positivity are finite on any compact subset of (λ_0, υ_0). When positivity is ignored the functions are finite on any compact subset of R. Further, these functions are convex functions of the hypothetical mean of the discount factor. As a result [see Theorem 10.8 of Rockafellar (1970)]. the sample analog functions converge uniformly, almost surely, on any compact subset of (λ_0, υ_0) in the case of positivity and on any compact set when positivity is ignored. One difficulty is that the approximations deteriorate as the mean assignment, v, approaches the arbitrage bounds in the case of positivity, or when v gets large when positivity is ignored.

The deterioration of the sample analog to S_0^+ in the vicinity of a finite arbitrage bound turns out not to be problematic. To see this, instead of viewing the boundaries of the feasible regions 'as functions' of the ordinate, we explore the approximation error from a set-theoretic vantage point in $\mathbb{R}^2$. Consider first the case in which $\upsilon_0 < +\infty$.

Associated with a sample of size T is an approximation error as measured by the Hausdorff metric:

$$\eta_T \equiv \max\{\pi(S_T^+, S_0^+), \pi(S_0^+, S_T^+)\} \tag{29}$$

where

$$\pi(K_1, K_2) \equiv \sup_{(v_1, w_1)\in K_1} \inf_{(v_2, w_2)\in K_2} |(v_1, w_1) - (v_2, w_2)|. \tag{30}$$

Measuring the approximation error via the Hausdorff metric allows ordered pairs to get close without restricting them to having the same ordinate.[5] In other words, we no longer confine our attention to "vertical" measures of distance, as is the case when we view the boundaries of the feasible regions as functions of the hypothetical (expected) prices of a unit payoff. The added flexibility in the Hausdorff metric permits us to exploit better the consistent estimation of the upper and lower arbitrage bounds (Proposition 2.3).

When v_0 is infinite, the approximation error η_T defined by Equation (29) will be infinite. As a remedy, we replace π by

$$\pi_\rho(C_1, C_2) \equiv \sup_{\substack{(v_1, w_1)\in K_1 \\ 0\le v_1 \le \rho}} \inf_{\substack{(v_2, w_2)\in K_2 \\ 0\le v_2 \le \rho}} |(v_1, w_1) - (v_2, w_2)| \tag{31}$$

where p is any arbitrary positive number greater than the lower arbitrage bound λ_0.

Proposition 2.4. *Under Assumption 1.1-1.3, and 2.1,* $\{\eta_T\}$ *converges to zero almost surely.*

This proposition follows since the arbitrage bounds can be consistently estimated (Proposition 2.3) and the lower boundaries of $\{S_T^+\}$ approach the lower boundary of S_0^+ uniformly on any compact interval within the arbitrage bound. This latter convergence follows from Proposition 2.1 and the convexity of the boundary.

3. Applications

In this section we discuss two applications of the analysis of Section 2. First we show how the feasible regions for the means and standard deviations can be used to test a specific model of the discount factor. Burnside (1994) and Cecchetti, Lam, and Mark (1994) have developed a version of this test when there are no assets subject to short-sale constraints or transactions costs. We demonstrate how this test can be implemented in a relatively simple manner by exploiting the results of Section 2. Further, we formulate the test so that it is also applica-

[5] The Hausdorff metric is usually employed for compact sets to ensure that the resulting distance is finite. Because of the vertical character of the regions and the existence of finite arbitrage bounds, the Hausdorff distance will be finite even though the sets are not bounded. The Euclidean distance in (30) could be replaced by the square root of a quadratic form in the differences between two points as long as a positive weight is given to both dimensions.

ble when there are assets subject to short-sale constraints. As a result this provides (large-sample) statistical foundation to the tests of asset pricing models suggested by He and Modest (1993) and Luttmer (1994).

Second, we outline an extension of the specification-error bound analysis that is useful when the discount factor proxy under consideration depends upon a vector of unknown parameters. We consider how these results can be used to select between two nonnested models by comparing the minimized values of the specification-errors.

3.1 Testing a specific model of the discount factor using volatility bounds

Suppose that in addition to asset-market data, a model of the discount factor is posited and a time series of observations of the discount factor is available: $\{m_t: t = 1, \dots, T\}$. One way to test the model is to examine whether it satisfies the volatility bounds discussed in Sections 1 and 2. Since observations of the discount factor are available, the average price of a unit payoff can be estimated by the mean of *m*. Specifically, form *x* by augmenting the original vector of payoffs with a unit payoff; form *q* by augmenting the original vector of prices with the random variable *m;* and form *C* by constructing the Cartesian product of the original cone with **R.** In effect, we have added a unit payoff with an average price *m* that *is* not subject to a short-sale constraint. In forming a test, we can apply the results of Sections 1 and 2.2 with one minor modification. The random functions $\hat{\phi}$ and $\tilde{\phi}$ are now constructed by setting the proxy *y* to zero and subtracting m^2:

$$\hat{\phi}(\alpha) \equiv -(-\alpha' x)^2 - 2\alpha' q - m^2, \tag{32}$$

and

$$\tilde{\phi}(\alpha) \equiv -(-\alpha' x)^{+2} - 2\alpha' q - m^2. \tag{33}$$

Subtracting m^2 does not alter the solutions to either the sample or population maximization problems. It does, however, change the maximized values of the criteria functions. The volatility bounds for *Em* will be satisfied if, and only if,

$$\hat{\xi} \equiv \max_{\alpha \in C} E\hat{\phi}(\alpha) \leq 0, \tag{34}$$

when positivity is ignored, or

$$\tilde{\xi} \equiv \max_{\alpha \in C} E\tilde{\phi}(\alpha) \leq 0, \tag{35}$$

when positivity is imposed. The limiting distribution reported in Proposition 2.2 (appropriately modified) can be applied to construct a test of these hypotheses using sample analog estimators of $\hat{\xi}$ and $\tilde{\xi}$.[6] Again, we have formulated the problem so that approximation error due to parameter estimation plays no role in the limiting distributions for these sample analogs.

In practice we find the solutions for the sample maximization problems, estimate the asymptotic standard errors, and form one-sided tests. In particular, let $\tilde{c}_T$ be the maximized value of $\Sigma_T(\tilde{\phi})$ over the constraint set C. Then $\{\sqrt{T}[\tilde{c}_T - \tilde{\xi}]$ converges in distribution to a normal random variable with mean zero and variance $u'\tilde{V}u$. This variance can be estimated in the manner described in Section 2.2. Since $\tilde{\xi}$ is not specified under the null hypothesis (35), the "conservative" choice of $\tilde{\xi} = 0$ is used in constructing the test statistic.[7]

Similar approaches to testing a model of the discount factor can be applied when a time series for m can be constructed from simulated data instead of actual data. In this case the randomness of $\tilde{\phi}(\tilde{\alpha})$ can be decomposed additively into two components, one due to the randomness of the security market payoffs and prices and the other due to the simulation of m. As in the work of McFadden (1989), Pakes and Pollard (1989), Lee and Ingram (1991), and Duffie and Singleton (1993), the asymptotic variance in the limiting distribution will now have an extra component because of the sampling error induced by simulation. For an example of this approach and a more extensive discussion; see Heaton (1994).

When the first two moments of m can be computed numerically with an arbitrarily high degree of accuracy, we can proceed as follows. Augment the price vector with Em instead of m and subtract Em^2 from the criteria instead of m^2 as in Equations (32) and (33). This same strategy can be used to assess the accuracy of the estimated feasible region for means and standard deviations of stochastic discount factors. For any hypothetical mean-standard deviation pair for m, one can compute the corresponding test statistic and probability value.

3.2 Minimizing the specification-error bound for parameterized families of models

Recall that the specification-error bounds provide a way to assess the usefulness of an asset pricing model even when it is technically

[6] Even if hypothesis (35) is satisfied, the sample analog may be infinite, making, implementation problematic. This happens when the sample mean is outside the estimated arbitrage bounds. This phenomenon does not arise for hypothesis (34).

[7] Burnside (1994) and Cecchetti, Lam, and Mark (1994) developed and stud&d alternative versions of the volatility bounds tests when no transactions costs are introduced. The test used by Cochrane and Hansen (1992) abstracted from positivity and can be formulated equivalently using $\hat{\phi}$ in (34). See Burnside (1994) for a Monte Carlo comparison of various volatility tests.

misspecified. In many situations the discount factor proxy depends on unknown parameters. For example, in a representative consumer model with constant relative risk aversion preferences, the pure rate of time preference and the coefficient of relative risk aversion are typically unknown. In this case, one way to estimate the parameters of the model is to minimize the specification error. Alternatively, in an observable factor model, the discount factor proxy depends on a linear combination of the factors with unknown coefficients. As in the work of Shanken (1987), one could imagine selecting factor coefficients to minimize the specification error. We now sketch how the results of Section 2.2 extend in a straightforward manner to obtain a distribution theory for the minimized value of the specification-error bound. In Section 4 we discuss distribution theory for the resulting estimators of the parameters of the discount factor.

Suppose that the discount factor proxy y depends on the parameter vector $\beta \in B$ where B is a compact set. The population optimization problems of interest are now

$$\hat{\delta}^2 \equiv \min_{\beta \in B} \max_{\alpha \in C} \left(Ey(\beta)^2 - E\{(y(\beta) - x'\alpha)^2\} - 2\alpha' Eq \right), \tag{36}$$

and

$$\tilde{\delta}^2 \equiv \min_{\beta \in B} \max_{\alpha \in C} \left(Ey(\beta)^2 - E\{[(y(\beta) - x'\alpha)^+]^2\} - 2\alpha' Eq \right). \tag{37}$$

When $\hat{\delta}$ and $\tilde{\delta}$ are strictly positive and the parameterized family of stochastic discount factors satisfies the appropriate smoothness and moment restrictions, an extended version of Proposition 2.2 can be obtained for the sample analog estimators of $\hat{\delta}$ and $\tilde{\delta}$. Again, the limiting distribution will be the same as if the solutions to the population optimization problems were known a priori.

The approach can be extended to compare the smallest specification-errors for two nonnested families of models. Such a comparison potentially can be used as a device for selecting between the two families of models. Vuong (1989) examined a very similar problem by using the large-sample behavior of likelihood ratios for two nonnested families of misspecified models [in particular, see the discussion in Section 5 of Vuong (198911, and we can imitate and adapt his analysis to our problem. More precisely, let $\tilde{\delta}_1$ and $\tilde{\delta}_2$ denote the specification-error bounds associated with two such families. Take the null hypothesis to be

$$\tilde{\delta}_1 = \tilde{\delta}_2. \tag{38}$$

Under the null hypothesis, the smallest specification-error associated with each parameterized family is the same. As a consequence, the performance of the two parameterized families cannot be ranked once sampling error is accounted for. This hypothesis can be tested by using the corresponding distribution theory for the difference between the analog estimators of $\tilde{\delta}_1$ and $\tilde{\delta}_2$ scaled by the square root of the sample size.

4. Asymptotic Distribution of the Multipliers

Recall that the asymptotic distribution for the estimators of $\tilde{\delta}^2$ and $\hat{\delta}^2$ discussed in Sections 2 and 3 depends only on the population solutions to the conjugate problems (8) and (9). In other words, sampling error in the estimated multipliers does not contribute to the limiting distribution of the specification error bounds. However, as elsewhere in econometric practice, the magnitudes of the multipliers remain interesting in their own right as a measure of the importance of components of the pricing relations to the bounds. Consequently, it is advantageous to be able to make statistical Inferences about their magnitude.

To amplify this point, one use of the asymptotic distribution for the multiplier estimators is to test whether the specificationerror or volatility bounds remain the same when a subset of assets is omitted from the analysis. A special case of such a test is a *region subset test* where the question of interest is whether, given an initial set of asset returns, additional asset returns result in an increase in the volatility bounds. It is problematic to construct a test directly in terms of the difference between a bound computed with the full set of securities and the corresponding bound calculated with the more limited army of securities because the resulting statistic has a degenerate distribution under the null hypothesis. This degeneracy follows from the fact that sampling error in the multipliers does not contribute to the limiting distribution for the bounds and is circumvented by instead basing the statistical test directly on the estimated multipliers. That is, it is advantageous to reformulate the null hypothesis to be

$$R\tilde{\alpha} = 0, \text{ or } R\hat{\alpha} = 0 \tag{39}$$

where R is an appropriately constructed selection matrix (see below), and to use the limiting distribution for the estimated multipliers in constructing an asymptotic chi-square test.

Several versions of region subset tests have been used in the literature.[8] For example, Snow (1991) considered the small firm effect

[8] The impetus for this work was the econometric discussion in an unpublished precursor to this article [Hansen and Jagannathan (1988)].

by examining whether the returns on small capitalization stocks have incremental importance in determining the volatility bounds over and above the returns of large capitalization stocks. Other examples can be found in Braun (1992), Cochrane and Hansen (1992), De Santis (1993), and Knez (1993).

The limiting distribution for the multipliers shows up in other applications as well. For instance, testing Relations (39) in conjunction with specification-error analysis is helpful in ascertaining which assets are important contributors to model misspecification. Also, when a researcher uses the specification-error bounds to select among a parameter&d family of discount factor proxies, it is desirable to make inferences about the parameter vector chosen. As we will see in this section, the asymptotic distribution for the parameter estimator selected in this fashion interacts with the limiting distribution for the estimated multipliers.- As a consequence, our characterization of the multiplier limit distribution is an important component in the derivation of the limiting distribution for estimators of parametric stochastic discount factor proxies.

The remainder of this section is organized as follows. In Section 4.1 the distribution theory for the estimated multipliers is developed assuming that there are no assets subject to short-sale constraints. In Section 4.2 we comment briefly on how the theory can be extended to the case where short-sale constraints are imposed on some of the assets.

4.1 Distribution without market frictions

In the absence of short-sale constraints, the cone C is $\mathbb{R}^n$. As a consequence the estimation problem for the multipliers is posed as an unconstrained maximization problem, and the limiting covariance matrices for the asymptotic distribution of the coefficient estimators have a form that is familiar both from M estimation [e.g., see Huber (1981)] and from *GMM* estimation [e.g., see Hansen (1982)]. The only complication occurs when considering the bounds in which the no-arbitrage restriction is fully exploited because of the kink induced by the nonnegativity restriction.

The moment conditions of interest are given by the first-order conditions (10) and (11) for the conjugate maximum problems:

$$E[x(y - x'\hat{\alpha}) - q] = 0, \tag{40}$$

and

$$E[x(y - x'\tilde{\alpha})^+ - q] = 0. \tag{41}$$

The analog estimators of $\hat{\alpha}$ and $\tilde{\alpha}$ ($\hat{\alpha}_T$ and $\tilde{\alpha}_T$) satisfy

$$\Sigma_T[x(y - x'\hat{a}_T) - q] = 0 \tag{42}$$

and

$$\Sigma_T[x(y - x'\tilde{a}_T)^+ - q] = 0. \tag{43}$$

While the equations for $\hat{\alpha}$ are linear, those for $\tilde{\alpha}$ are nonlinear. In the latter case, we use a linear approximation to the moment conditions in deriving the central limit approximation for the parameters:

$$x(y - x'\alpha)^+ - q \tag{44}$$

Notice that the function of a on the left side of Relation (44) is differentiable except at values of $\tilde{\alpha}$ such that $y - x'\tilde{\alpha} = 0$. We assume that such sample points are "unusual":

Assumption 4.1. $Pr\{y - x'\tilde{\alpha} = 0\} = 0$.

As is shown in Appendix C, Assumptions 1.1 and 4.1 are sufficient for us to study the asymptotic behavior of the estimator $\{a_T\}$ using the linearization on the right side of Relation (44):[9]

$$E[x(y - x'\alpha)\mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}} - q] = 0. \tag{45}$$

To use linear equation system (45) to identify $\tilde{\alpha}$, the matrix $E(xx'\mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})$ must be nonsingular. Given Assumption 4.1, this rank condition is equivalent to Assumption 1.4 because x and x^* must coincide when no short-sale constraints are imposed. The counterpart to this rank condition for $\hat{\alpha}$ is that the second moment matrix (xx') be nonsingular as required by Assumption 1.3.

Working with the two linear moment conditions, we obtain the approximations:

$$\sqrt{T}(\tilde{a}_T - \tilde{\alpha}) \approx -[E(xx'\mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})]^{-1}\sqrt{T}\Sigma_T[x(y - x'\tilde{\alpha})^+ - q]$$
$$\sqrt{T}(\hat{a}_T - \hat{\alpha}) \approx -[E(xx')]^{-1}\sqrt{T}\Sigma_T[x(y - x'\hat{\alpha}) - q] \tag{46}$$

where the notation $\approx$ is used to denote the fact that the differences between the left and right sides of Relations (46) converge in probability to zero. Let $w \equiv [0, I_n]$. Combining approximations (46) with Assumptions 2.2 and 2.3 gives us the asymptotic distribution of the analog estimators.

[9] Our formal derivation of the distribution theory uses a result from Pakes and Pollard (1989). A by product from our analysis in the Appendix is a (modest) weakening of the assumptions imposed in Hansen (1982) to accomodate kinks in the moment conditions used in estimation.

Proposition 4.1. *Suppose Assumptions 1.1-1.3, 2.1, and 2.2 are satisfied. Then* $\{\sqrt{T}(\hat{a}_T - \hat{\alpha})\}$ *converges in distribution to a normally distributed random vector with mean zero and covariance matrix:* $[E(xx')]^{-1} w\hat{V}w' [E(xx')]^{-1}$. *Suppose Assumptions 1.1-1.4, 2.1 ,2.3, and 4.1 are satisfied. Then* $\sqrt{T}(\tilde{a}_T - \tilde{\alpha})\}$ *converges in distribution to a normally distributed random vector with mean zero and covariance matrix:*

$$[E(xx'\mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})]^{-1} w\tilde{V}w' [E(xx'\mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})]^{-1}.$$

To apply these limiting distributions in practice requires consistent estimators of the asymptotic covariance matrices. The terms $w\hat{V}w'$ and $w\tilde{V}w'$can be estimated using one of the spectral methods referenced previously. Under assumptions maintained in Proposition 4.1, *the matrices* $E(xx')$ and $E(xx'\mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})$ can be estimated consistently by their sample analogs, where the estimator $\tilde{a}_T$ is used in place of $\tilde{\alpha}$ in estimating the second of these matrices.

We now briefly consider two extensions of Proposition 4.1.

(i) Estimation of the parameters of a discount factor proxy. Suppose that the discount factor proxy depends on the parameter vector β, and let $\tilde{\beta}$ be the parameter vector that minimizes the specification error when positivity is imposed. Assume that the parameterized family satisfies the appropriate smoothness and moment restrictions and that $\tilde{\beta}$ is in the interior of the parameter space. The population moment conditions are given by

$$E\{x[y(\tilde{\beta}) - x'\tilde{\alpha}]^+ - q\} = 0; \tag{47}$$

and

$$E\left(\frac{\partial y}{\partial \beta}(\tilde{\beta})\{y(\tilde{\beta}) - [y(\tilde{\beta}) - \tilde{\alpha}'x]^+\}\right) = 0. \tag{48}$$

The distribution theory for the analog estimators $\{\tilde{b}_T\}$ and $\{\tilde{a}_T\}$ of $\tilde{\beta}$ and $\tilde{\alpha}$ respectively, can be deduced by taking linear approximations to the sample moment conditions (47) and (48) and appealing to the results of Appendix C.

(ii) Region subset tests. Let z denote an $(n-1)$-dimensional vector of assets under consideration with price vector s, and let f be the k-dimensional subvector of z, including the $k - 1$ asset payoffs that are to be used to construct the bound augmented by a unit payoff.

Formally, the region subset test can be represented as the hypothesis:

$$E[z(f'\tilde{\theta})^{+} - s] = 0,$$
$$E[(f'\tilde{\theta})^{+} - v_0] = 0. \quad (49)$$

Hence, the hypothesis of interest is whether $(f'\tilde{\theta})$ has mean v_0 and prices the vector of securities z correctly. One possibility is to test this hypothesis for a prespecified v_0, and the other is to test whether it is satisfied for some $v_0 > 0$.

Consider first the case in which v_0 is prespecified. To map this into our previous setup, form the n-dimensional vector x by augmenting z with a unit payoff, and form q by augmenting s with the 'price" v_0. Then hypothesis (49) can be interpreted as a zero restriction on the coefficient vector $\tilde{\alpha}$ employed in Sections 1, 2, and 4.1. Thecomponents of coefficient vector $\tilde{\alpha}$ set to zero correspond to the entries of z that are omitted from f. A large sample Wald test of this zero restriction can be formed by applying the limiting distribution in Proposition 4.1.

Alternatively, we could estimate the parameter vector θ based on the overidentified system (49) using *GMM*. The analysis leading up to Proposition 4.1 can be easily modified to show that the minimized value of the criterion function is distributed as a chi-square random variable with $n - k$ degrees of freedom [see Hansen (198211. When the hypothesis of interest is altered to be for some $v_0 > 0$, this *GMM* approach is modified by minimizing the criterion function by choice of θ and v with a corresponding loss in the degrees of freedom.

4.2 Distribution with market frictions

We now briefly describe how the distribution theory is modified when some short-sale constraints are imposed (C is a proper subset of $\mathbf{R}^n$). We will focus on the limiting behavior of $\sqrt{T}(\tilde{a}_T - \tilde{\alpha})$, but the results for $\sqrt{T}(\hat{a}_T - \hat{\alpha})$ are very similar. As in Section 1, we partition x by whether or not $\tilde{m}$ prices the payoffs with equality or not; that is, by whether

$$E\tilde{m}x^i = Eq^i, \text{ or } E\tilde{m}x^i < Eq^i. \quad (50)$$

The component coefficient estimators that multiply x^i's for which there is strict inequality will equal zero with arbitrarily high probability as the sample size gets large. Hence, the limiting distribution is degenerate for these component estimators.

Consider next the estimator of the remaining subvector of $\tilde{\alpha}$, which we denote $\tilde{\gamma}$. Because of the degeneracy just described, we can, in effect, treat the limiting distribution of the estimator of $\tilde{\gamma}$ separately. Let C be the lower-dimerional cone associated with estimating $\tilde{\gamma}$. If $\tilde{\gamma}$ is an interior point of $\tilde{C}$, then the argument leading up to Proposition 4.1 can be imitated to deduce a limiting normal distribution for the parameter estimator, However, if $\tilde{\gamma}$ is at the boundary of the cone C, the limiting distribution may be a nonlinear function of a normally distributed random vector [see Haberman (1989), p. 1645].[10]

5. Conclusions and Extensions

In this article we provided statistical methods for assessing asset pricing models based on specification-error and volatility bounds. Two significant advantages of the statistical methods are that they are easy to interpret and that they are simple to implement even in the presence of transactions costs and short-sales constraints. The results can be used in a variety of ways. For example, they can be used to test specific models of the discount factor, to examine the information contained in different sets of asset-market data, and to assess misspecified asset pricing models.

There are several interesting extensions of the econometric methods in this article, including the following.

(i) The short-sale constraint formulation could be generalized to include "solvency constraints" whereby portfolios are restricted so that portfolio payoffs are nonnegative. This amounts to imposing a form of borrowing constraint on consumers [see, e.g., Hindy (1993) and Luttmer (1994)]. As in Section 1, the constraints on portfolio weights in the presence of solvency constraints can be formulated as a closed convex cone. However, without knowledge of the distribution of the payoffs on primitive securities, the cone would be subject to estimation error and hence not directly covered by the results in this article. The consistency proof in Section 2.3 for the arbitrage bounds might well be adaptable to approximating constraint sets more generally. It is then of interest to understand how approximation errors for the constraint sets affect the distribution of the specification and/or volatility bounds.

(ii) A difficult feature of the limiting distribution theory for the Lagrange multipliers in the presence of short-sales constraints (Section 4.3) is the manner in which it depends on the true parameter vector and the associated discontinuities. This feature makes the distribution theory harder to use in practice and, in other settings, has led researchers to compute approximate bounds on proba-

[10] Haberman characterized this nonlinear function as a particular projection onto a dosed convex set formed by translating $\tilde{C}$ by $-\tilde{\gamma}$. Although Haberman (1989) considers only the case in which the data are *iid,* his characterization of the limiting distribution applies mote generally with a covariance matrix replaced by a spectral density matrix at frequency zero.

bilities of test statistics [see, e.g., Wolak (1991) and Boudoukh, Richardson, and Smith (1992)]. Perhaps similar probability bounds could be derived for the region subset tests with frictions.

(iii) Using tools similar to those described in Section 1, Chen and Knez (1995) have developed nonparametric measures of market integration. It should be possible to extend the econometric methods in this article to incorporate transactions costs in the market integration measures they propose.

Appendix A: Consistency

In this Appendix we demonstrate formally the results of Sections 2.1 and 2.4. We maintain Assumptions 1.1-1.3 and 2.1 throughout.

Let U denote a compact set in $\mathbb{R}^n$. For any subset h of U, we let $cl(h)$ denote the closure of h. Let $\mathcal{K}$ denote the collection of all nonempty closed subsets of U. We use the Hausdorff metric η on $\mathcal{K}$ given by

$$\eta(b_1, b_2) \equiv \max\{\sup_{\alpha_1 \in b_1} \inf_{\alpha_2 \in b_2} |\alpha_1 - \alpha_2|, \sup_{\alpha_2 \in b_2} \inf_{\alpha_1 \in b_1} |\alpha_1 - \alpha_2|\}. \tag{A1}$$

to define notions of convergence of compact sets. For some of our results, we will use the construct of a lim sup of a sequence in $\mathcal{K}$. We follow Hildenbrand (1974) and define:

Definition A.1. For a sequence $\{b_j\}$ in U,

$$\limsup b_j \equiv \bigcap_m cl\left(\bigcup_{j \geq m} b_j\right).$$

Since the lim sup is the intersection of a decreasing sequence of closed sets, it is closed and not empty. An alternative way to characterize the lim sup is to imagine forming sequences of points by selecting a point from each h_j. All of the limit points of convergent subsequences are in the lim sup, and, in fact, ail of elements of the lim sup can be represented in this manner.

We shall make reference to an implication of a corollary on page 30 of Hildenbrand (1974) that characterizes the set of minimizers of an "approximating" function over an "approximating set."

Lemma A1. *Suppose*

(i) $\{\psi_j\}$ *is a sequence of continuous functions mapping U into* $\mathbb{R}$ *that converges uniformly to* ψ_∞; *and*

(ii) $\{b_j\}$ *converges to* b_∞.

Then lim sup $g_j \subset g_\infty$ *where*

$$g_j \equiv \{u \in b_j : \psi_j(u) \leq \psi_j(u') \text{ for all } u' \in b_j\},$$

and

$$\lim \min_{b_j} \psi_j = \min_{b_\infty} \psi_\infty.$$

Proof. To verify that this follows from the corollary in Hildenbrand, let $\mathcal{J}$ denote the set of positive integers augmented by $+\infty$ and endow $\mathcal{J}$ with the usual metric for a one-point compactification. Then, in light of (i), the sequence $\{\psi_j\}$ in conjunction with ψ_∞ defines a continuous function on $\mathcal{J} \times \mathcal{U}$; and in light of (ii), the sequence $\{b_j\}$ in conjunction with b_∞ defines a continuous compact correspondence mapping $\mathcal{J}$ into $\mathcal{U}$. The conclusion of the Lemma A1 then follows from the corollary together with part (ii) of Proposition 1 of page 22 in Hildenbrand. ∎

Turning to the result in Sections 1 and 2.1, we first establish the compactness of the set of solutions to the conjugate problems.

Lemma A2. *The set of solutions to conjugate maximization problems (8) and (9) are compact.*

Proof. We consider the case of problem (9). The proof for the case of problem (8) is similar. The set of solutions is closed because the constraint set is closed and the criterion function is continuous. Boundedness of the set of solutions can be demonstrated by investigating the tail properties of the criterion function. We consider two cases: directions ζ for which $\zeta' x$ is negative with positive probability and directions ζ for which $\zeta' x$ is nonnegative. To study the former case, we take the criterion in problem (9) and divide it by $1+|\alpha|^2$. For large values of $|\alpha|$, the scaled criterion is approximately

$$-E[(-x'\zeta)^{+2}] \text{ where } \zeta \equiv \alpha/[1+|\alpha|^2]^{1/2}. \tag{A2}$$

Hence $|\zeta|$ is approximately one for large values of $|\alpha|$. Moreover, $\zeta' x$ is a payoff in P. Consequently, the unscaled criterion will decrease (to $-\infty$) quadratically for large values $|\alpha|$.

Consider next directions ζ for which $\zeta' x$ is nonnegative. From Assumption 1.2 we have that $\zeta' E\iota$ must be strictly positive unless $\zeta' x$ is identically zero. When $\zeta' x = 0$, it follows from Assumptions 1.2 and 1.3 that $\zeta' Eq$ is again strictly positive.

For directions ζ for which the payoff $\zeta' x$ is nonnegative, we study the tail behavior of the criterion after dividing by $(1+|\alpha|^2)^{1/2}$, which yields approximately $-\zeta' Eq$ for large values of $|\alpha|$. Hence, in these directions the unscaled criterion must diminish (to $-\infty$), at least linearly in $|\alpha|$. Thus, in either case we find that the set of solutions to conjugate problem (9) is bounded. ∎

We now formally establish Proposition 2.1.

Proof of Proposition 2.1. We treat only the consistency of $\{\tilde{d}_T\}$ because the corresponding argument for $\{\hat{d}_T\}$ is very similar. Assumptions 1.1 and 2.1 imply that $\{\Sigma_T\tilde{\phi}(\alpha)\}$ converges almost surely to $E\tilde{\phi}(\alpha)$ for each $\alpha \in C$. Since, for each T, $\Sigma_T\tilde{\phi}$ is concave, as is $E\tilde{\phi}$, Theorem 10.8 of Rockafellar implies that $\Sigma_T\tilde{\phi}\}$ converges uniformly on any compact set in $\mathbb{R}^n$. Further, from Lemma A.2, the set of maximizers of $E\tilde{\phi}$ is bounded. For a positive number N, define $C_N \equiv \{\alpha \in C : |\alpha| \leq N\}$ and $D_N \equiv \{\alpha \in C : |\alpha| = N\}$. Then C_N and D_N are compact. By choosing N to be sufficiently large, we can ensure that C_N contains all of the maximizers of $E\tilde{\phi}$ over the constraint set C and that none of the maximizers are in D_N. Let δ_N be the maximized value of $E\tilde{\phi}$ over D_N. Then by choice of N we have that $\delta_N < \bar{\delta}$. Since $\{\Sigma_T\tilde{\phi}\}$ converges uniformly to $E\tilde{\phi}$ on C_N almost surely, for sufficiently large T, the maximizers of $\Sigma_T\tilde{\phi}$ over C_N are also not in D_N. By the convexity of C and concavity of $\Sigma_T\tilde{\phi}$, it follows that for sufficiently large T, the maximizers of $\Sigma_T\tilde{\phi}$ over C_N coincide with those over C. Consequently, the almost sure convergence of $\{\tilde{d}_T\}$ to $\bar{\delta}$ follows from the almost sure uniform convergence of $\{\Sigma_T\tilde{\phi}\}$ on C_N. ∎

We now turn to the results in Section 2.3 and investigate the statistical consistency of sample analog estimators $\{\ell_T\}$ and $\{u_T\}$ for the arbitrage bounds λ_0 and v_0. Recall that the arbitrage bounds are representable as solutions to linear programming problems. Since there is no natural compact set for the choice variables in these problems, we must explore "directions to infinity." We study these "directions" using a compactification of the parameter space.

First consider any $\alpha \in C$ such that $\alpha'x \geq -1$ with probability one. Then with probability one $\alpha'x_t \geq -1$ for all t with probability one and $\{\alpha'\Sigma_T q\}$ converges almost surely to $\alpha' Eq$. Define $\ell_T^- \equiv -\ell_T$ and $\lambda_0^- \equiv -\lambda_0$. Since $\ell_T^- \leq \alpha'\Sigma_T q$, it follows that lim sup $\ell_T^- \leq \lambda_0^-$ with probability one. Similarly, if $v_0 < \infty$, lim sup $u_T \leq v_0$. Hence our interest is in the lim inf ℓ_T^- and lim inf u_T.

To construct a compact parameter space, we map the original parameter space for each problem into the closed unit ball in $\mathbb{R}^n$, which we denote as U. We consider explicitly the case of u_T. The proofs for the case of ℓ_T are completely analogous to the case for u_T and are omitted.

Notice that the constraint set used in defining v_0 can be represented as the set of all $\alpha \in C$ satisfying the equation

$$E[(1 - \alpha' x)^+] = 0. \tag{A3}$$

As in the proof of Lemma A.2, we map the parameter space into the unit ball (with a slightly different transformation). The transformation $\zeta \equiv \alpha/(1 + |\alpha|)$ maps $\mathbf{R}^n$ into the open unit ball. To compactify the transformed parameter space, we consider adding the boundary points of the unit ball. Notice that we can recover the original parameterization by considering the inverse mapping:

$$\alpha \equiv \zeta/(1 - |\zeta|) \tag{A4}$$

for $|\zeta| < 1$. Using the transformation in Formula (A4), instead of considering those *a's* that solve Equation (A3) we consider

$$D^* \equiv \{\zeta \in \mathcal{U} \cap C \mid E\{[(1 - |\zeta|) - x'\zeta]^+\} = 0\}. \tag{A5}$$

This transformation potentially adds solutions to, Equation (A3) by including the boundary of the unit ball. The potentially problematic values of ζ are those for which $x'\zeta \geq 0$, $\zeta \in C$ and $|\zeta| = 1$. We rule this out by limiting attention to values of ζ in

$$\hat{D} \equiv \{\zeta \in \mathcal{U} \cap C \mid \zeta' Eq \leq (1 - |\zeta|)(v_0 + 1)\}. \tag{A6}$$

Notice that any ζ in $\hat{D}$ for which $|\zeta| \neq 1$ satisfies

$$\zeta' Eq/(1 - |\zeta|) \leq (v_0 + 1).$$

in effect, by focusing on ζ's in $\hat{D}$ we are eliminating ζ's corresponding to payoffs with "high" prices. This does not cause us problems because we are concerned with estimated upper arbitrage bounds that are too low, not too high. Also, any ζ in $\hat{D}$ for which $|\zeta| = 1$ must have an (average) price that is nonpositive. This elites the troublesome points (directions) from D^*. Let D_T^* be the sample analog of D^* and $\hat{D}_T$ be the sample analog of $\hat{D}$. We first consider the limiting behavior of $D_T^* \cap \hat{D}_T$:

Lemma A.3. *Suppose that* $v_0 < \infty$. *Then*

$$\limsup D_T^* \cap \hat{D}_T \subset D^* \cap \hat{D}.$$

Proof. First notice that, since $\Sigma_T q$ converges to Eq almost surely, then $\eta(\hat{D}_T, \hat{D})$ converges almost surely to 0. We next establish that $\lim D_T^* = D^*$. To do this we first show that $\Sigma_T[(1 - |\zeta|) - x'\zeta]^+$ converges uniformly to $E[(1-|\zeta|)-x'\zeta]^+$ on $\mathcal{U}$. Note that

$$\mathcal{U} \cap C$$

is compact and that

$$E\left\{|[(1-|\zeta_1|)-x'\zeta_1]^+-[(1-|\zeta_2|)-x'\zeta_2]^+|\right\} \le (1+(E|x|^2)^{1/2})|\zeta_1-\zeta_2|. \tag{A7}$$

This is sufficient for the Uniform Law of Large Numbers of Hansen (1982) to apply. Hence, from Lemma A.1, the lim sup of the sequence of minimizers of $\Sigma_T[(1-|\zeta|)-x'\zeta]^+$ over $\mathcal{U} \cap C$ is contained in the set of minimizers of $E[(1-|\zeta|)-x'\zeta]^+$. Since $v_0 < \infty$ the set D^* is not empty and D^* is the set of minimizers of $E[(1-|\zeta|)-x'\zeta]^+$. With probability one, any point in D^* must also be in D_T^* for all $T \ge 1$. Since D^* is separable, a common probability measure one set of sample points can be selected so that $D^* \subseteq D_T^*$ for all $T \ge 1$. As a result $\lim D_T^* = D^*$. The conclusion follows. ∎

Lemma A.4. *Suppose that* $v_0 < \infty$. *Then*

$$\liminf u_T \ge v_0.$$

Proof: First note that

$$u_T = \min\{\zeta'\Sigma_T q/(1-|\zeta|) \mid \zeta \in D_T^* \cap \hat{D}_T\} \quad \text{for sufficiently large } T, \tag{A8}$$

and

$$v_0 = \min\{\zeta' Eq/(1-|\zeta|) \mid \zeta \in D^* \cap \hat{D}\}.$$

Hypothetical expansions of the constraint set $D_T^* \cap \hat{D}_T$ for u_T can only result in smaller values of the maximized criterion. For instance, suppose the constraint set is augmented to include all of the points in $D^* \cap \hat{D}$. Then Lemma A.2 implies that this sequence of augmented constraint sets converges to $D^* \cap \hat{D}$. The conclusion then follows from Lemma A.1. ∎

Finally, we consider the case in $v_0 = \infty$.

Lemma A.5. *Suppose that* $v_0 = \infty$ *Then* $\{u_T\}$ *diverges with probability one.*

Proof. Since $v_0 = \infty$, there are no values of $\alpha \in C$ such that $\alpha' x \ge 1$ with probability one. Consequently, the only values of ζ in D^* are ones for which $|\zeta| = 1$. We consider two cases. First suppose that $D^* = \emptyset$. The uniform convergence of $\Sigma_T[(1-|\zeta|) - x'\zeta]^+$ to $E[(1-|\zeta|) - x'\zeta]^+$ implies that for sufficiently large T, $D_T^* = \emptyset$ and $u_T = \infty$. Next suppose that $D^* \neq \emptyset$. Since there are no arbitrage opportunities (Assumption 1.2), $\zeta' Eq > 0$ for any ζ in D^* such that $\|\zeta' x\| > 0$. Also, Assumption 1.2 together with the no-redundancy Assumption 1.3 imply that $\zeta' Eq > 0$ for any ζ in D^* such that $\|\zeta' x\| = 0$. Furthermore, D^* is closed, implying that

$$\varepsilon \equiv \inf\{\zeta' Eq : \zeta \in D^*\} > 0. \tag{A9}$$

Since $\{\Sigma_T q\}$ converges to Eq almost surely and D_T^* converges almost surely to D^*, it follows from Lemma A.1 that, with probability one for sufficiently large T, $\zeta'\Sigma_T q > \varepsilon/2$ for all $\zeta \in D_T^*$. he convergence of $\{D_T^*\}$ to D^* coupled with the fact that all elements of D^* have norm one then implies that $\{u_T\}$ diverges almost surely. ■

Taken together, Lemmas A.3, A.4, and A.5 imply proposition 2.3.

Appendix B: Asymptotic Distribution of the Bounds Estimators

In this Appendix we show that, in the case in which the prices of the payoffs are constant, the asymptotic distribution of the estimated bounds can be demonstrated even when the parameter vector is not uniquely identified (even when Assumption 1.4 is not satisfied).

Proof of Proposition 2.2.

We consider the case of $\tilde{d}_T$. The case of $\hat{d}_T$ is similar. Let h_T be the set of maximizers of $\Sigma_T\tilde{\phi}$ and let b_∞ be the set of maximizers of $E\tilde{\phi}$. For each T, let $\tilde{a}_T$ be a measurable selection from h_T [see Theorem 1 of Hildenbrand (1974), p. 54). Since lim sup $b_T = b_\infty$ almost surely and b_∞ is compact, there is a sequence $\{\alpha_T\}$ in b_∞ such that

$$\lim |\tilde{\alpha}_T - \alpha_T| = 0$$

almost surely (see Appendix A). Further, an implication of Lemma A.1 of Hansen and Jagannathan (1991) is that all $\alpha \in b_\infty$ result in the same random variable

$$\tilde{m} = (y - \alpha' x)^+.$$

Also, the complementary slackness condition for problem (9) implies that, for a $\in b_\infty$, $\alpha' q = E\{y(y-\alpha' x)^+ - (y - \alpha' x)^{+2}\}$, so that $\alpha' q$ the same for all $\alpha \in b_\infty$. As a result, the random variable $\tilde{\phi}(\alpha)$ is the same for all $\alpha \in b_\infty$. Now consider the decomposition of

$$\sqrt{T}\Sigma_T[(\tilde{d}_T)^2 - \tilde{\delta}^2]$$

as in Equation (23):

$$\sqrt{T}[(\tilde{d}_T)^2 - \tilde{\delta}^2] = \sqrt{T}\Sigma_T[\tilde{\phi}(\tilde{a}_T) - \tilde{\phi}(\tilde{\alpha}_T)] + \sqrt{T}\Sigma_T[\tilde{\phi}(\tilde{\alpha}_T) - E\tilde{\phi}(\tilde{\alpha}_T)]. \quad \text{(B1)}$$

As in relation (26), we have

$$0 \leq \sqrt{T}\Sigma_T[\tilde{\phi}(\tilde{a}_T) - \tilde{\phi}(\tilde{\alpha}_T)] \leq \sqrt{T}\Sigma_T[(\tilde{m}x - q) - E(\tilde{m}x - q)] \cdot (\tilde{a}_T - \tilde{a}_T).$$

Since $|\tilde{a}_T - \alpha_T|$ converges almost surely to 0, the result follows. ■

Appendix C: Asymptotic Distribution of the Multipliers

In this Appendix we consider the asymptotic distribution of our estimator of the Lagrange multipliers when there are no transaction costs. We begin by demonstrating that restrictions used in Hansen (1982) can be extended along the lines of Pollard (1985) and Pakes and Pollard (1989) to accommodate "kink" in the functions used to represent the moment conditions. We then show how to use this result to prove Proposition 4.1.

The notation used in our initial proposition for *GMM* estimators conflicts with some of the notation used elsewhere in the article. We let β_0 denote the parameter vector of interest and β any hypothetical point in the underlying parameter space $\mathcal{P}$. The parameter space is restricted to satisfy:

Assumption C.1 *$\mathcal{P}$ contains an open ball in $\mathbf{R}^k$ about β_0.*

We will use the construct of a random function. A random function ψ maps the set of sample points into the space of vector-valued continuous functions on $\mathcal{P}$. We require that $\psi(\beta)$ be an *n*- dimensional random vector for each β in $\mathcal{P}$.

We also consider an approximating function

$$\psi_t^a(\beta) = \psi_t(\beta_0) + \Delta_t(\beta - \beta_0)$$

that is linear β. The composite random function satisfies:

Assumption C.2. $\{(\psi_t', \psi_t^{a\prime})'\}$ *is stationary and ergodic and has finite first moments.*

We now specify the sense in which ψ_t^a is required to approximate ψ_t. The approximation error induced by using ψ_t^a in place of ψ_t is

$$r_t(\beta) \equiv |\psi_t(\beta) - \psi_t^a(\beta)|.$$

Define

$$dmod_t(\delta) \equiv \sup\{r_t(\beta)/|\beta - \beta_0| : |\beta - \beta_0| < \delta, \beta \neq \beta_0\}.$$

Note that $dmod_t(\cdot)$ is monotone in δ. Therefore, we can take almost sure limits as δ declines to zero. We impose the following restrictions on mod,.

Assumption C3. $\lim_{\delta \searrow 0} dmod_t(\delta) = 0$ *almost surely.*

Assumption C4. $E[dmod_t(\delta)] < \infty$ for some $\delta > 0$.

To satisfy Assumptions C.3 and C.4, Δ_t is typically taken to be the matrix of partial derivatives of ψ_t at β_0 when ψ_t is differentiable at β_0 and be *well behaved* for the other sample points. The random variable $dmod_t(\delta)$ is interpreted as the *modulus of differentiability* for ψ_t at β_0.

The approach adopted in Hansen (1982) is to restrict the modulus of continuity of the derivative of ψ_t to converge almost surely to zero and to have a finite expectation for some neighborhood of the parameter. It follows from the mean-value theorem that restrictions imposed in Hansen (1982) on the local behavior of ψ_t imply Assumptions C.3 and C.4.

We use Assumptions C.3 and C.4 to study the sense in which $\Sigma_T\psi$ is stochastically differentiable. Hence, look at the approximation error

$$\varepsilon_T(\delta) \equiv \sup\{|\Sigma_T\psi(\beta) - \Sigma_T\psi^a(\beta)|/|\beta - \beta_0| : |\beta - \beta_0| < \delta, \beta \neq \beta_0\}.$$

By the triangle inequality we have that

$$\varepsilon_T(\delta) \leq \Sigma_T\, dmod(\delta).$$

Thus, by Assumptions C.1-C.2, we have that

$$\lim_{\delta \searrow 0} \limsup_{T \to \infty} \varepsilon_T(\delta) \leq \lim_{\delta \searrow 0} E\, dmod(\delta) \qquad \text{(Cl)}$$
$$= 0.$$

This in turn implies the stochastic differentiability condition in Pollard (1985) because the counterpart to $\varepsilon_T(\delta)$ in Pollard's condition is scaled by

$$\sqrt{T}|\beta - \beta_0|/(1 + \sqrt{T}|\beta - \beta_0|),$$

which is less than one. Also, the iterated limit in Equation (Cl) implies the limit taken in Pollard's condition because ε_T is monotone in δ. 'The differentiability of limiting moment function $E\psi$ follows directly from Assumption C.4. Therefore, $\Sigma_T\psi - E\psi$ satisfies the stochastic differentiability condition with derivative at β_0 given by $\Sigma_T\Delta - E\Delta$. Since $\{\psi_t^a\}$ is stationary and ergodic, $\{\Sigma_T\Delta - E\Delta\}$ converges almost surely to zero; hence, the derivative is asymptotically negligible.

Next we impose a *global identification* condition on the approximating function ψ_t^a. Since the approximation of ψ_t by ψ_t^a is local, this condition can also be viewed as a *local identification* condition on the original function ψ_t.

Assumption C.5. $E|\Delta_t| < \infty$ and $E\Delta_t$ *has full rank k.*

This rank condition on the derivative together with the stochastic differentiability conditions already established imply the equicontinuity condition (iii) in Theorem 3.3 of Pakes and Pollard (1989) (see the discussion on p. 1043 of Pakes and Pollard).

We study the behavior of an estimator h_T that solves the equations

$$a_T\Sigma_T\psi(b_T) = 0$$

for sufficiently large *T*. The $(k \times n)$ random matrix a_T selects the linear combination of moment conditions to be used in estimation.

Assumption C.6. $\{b_T\}$ *converges in probability to* β_0.

Assumption C.7. $\{a_T\}$ *converges in probability to a nonrandom matrix* a_0 *where* $a_0E\Delta$ *is nonsingular.*

Finally, to obtain a limiting distribution for $\{b_T\}$ we assume

Assumption C.8.

$$\{\sqrt{T}\Sigma_T\psi(\beta_0)\}$$

converges in distribution to a normally distributed random vector with mean zero and nonsingular covariance matrix V_0.

Sufficient conditions for Assumption C.8 can be obtained using martingale approximations as described by Gordin (1969), Hall and Heyde (1980), and Hansen (1985). This condition implies that $E\psi(\beta_0)$ is equal to zero.

The following extension of Theorem 3.1 in Hansen (1982) is now a direct consequence of Theorem 3.3 and Lemma 3.5 in Pakes and Pollard (1989).

Proposition C.1. *Suppose that Assumptions C.1-C.8 are satified. Then* $\{\sqrt{T}(b_T - \beta_0)\}$ *converges in distribution to a normally distributed random vector with mean zero and covariance matrix*

$$[a_0 E(\Delta)]^{-1} a_0 V_0 a_0' [E(\Delta') a_0]^{-1}.$$

Estimation of $E\Delta$ follows as in Hansen (1982) as long as A can be expressed in terms of a random matrix function D that satisfies $\Delta = D(\beta_0)$ where D is continuous at β_0 with probability one and has a modulus of continuity with a finite first moment for some $\delta > 0$. In this case, $\{\Sigma_T D(b_T)\}$ converges in probability to $E\Delta$.

Proof of Proposition 4.1. In light of Proposition C.1, we now verify that our approximation in Equation (44) satisfies Assumption C.3. Let t(a) denote the random approximation error:

$$r(\alpha) \equiv |x(y - x'\alpha)(\mathbf{1}_{\{y-x'\alpha\geq 0\}} - \mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})|. \tag{C2}$$

It follows from the Cauchy-Schwarz inequality that

$$\begin{aligned} r(\alpha) &\leq |x(y - x'\alpha)||(\mathbf{1}_{\{y-x'\alpha\geq 0\}} - \mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})| \\ &\leq |xx'\alpha - xx'\tilde{\alpha}||(\mathbf{1}_{\{y-x'\alpha\geq 0\}} - \mathbf{1}_{\{y-x'\tilde{\alpha}\geq 0\}})| \\ &\leq |x|^2 |\alpha - \tilde{\alpha}| \end{aligned} \tag{C3}$$

where the second inequality follows because $|x'\alpha - x'\tilde{\alpha}|$ dominates, $|x(y - x'\alpha)|$ whenever $y - x'\alpha$ and $y - x'\tilde{\alpha}$ have opposite signs. Therefore, the random approximation error satisfies

$$r(\alpha)/|\alpha - \tilde{\alpha}| \leq |x|^2 \tag{C4}$$

for $\alpha \neq \tilde{\alpha}$ implying that the *modulus of differentiability*

$$dmod(\delta) = \sup\{r(\alpha)/|\alpha - \tilde{\alpha}| : |\alpha - \tilde{\alpha}| < \delta, \text{ for } \alpha \neq \tilde{\alpha}\} \tag{C5}$$

is dominated by $|x|^2$. Combined with Assumption 1.1 this implies that, for any positive value of δ, $E[dmod(\delta)]$ is finite. As $\delta \searrow 0$, $dmod(\delta)$ goes to zero except when

$$\mathbf{1}_{\{y-x'\tilde{\alpha}=0\}} = 1.$$

In this case it is possible to choose a such that

$$|\alpha - \tilde{\alpha}| < \delta$$

and

$$\mathbf{1}_{\{y-x'\alpha<0\}} = 1$$

so that $r(\alpha) = |xx'|$. However, Assumption 4.1 implies that this occurs with probability zero so that, as $\delta \searrow 0$, $dmod(\delta)$ converges almost surely to zero. ∎

References

Andrews, D. W. K., 1991, "Heteroskedasticity and Autocorrelation Consistent Consistent Covariance Matrix Estimation," *Econometrica,* 59, 817-858.

Boudoukh, J., M. Richardson, and T. Smith, 1992, "Testing Inequality Restrictions Implied from Conditional Asset Pricing Models," working paper, New York University.

Braun, P. A.. 1992, "Tests of Consistent Pricing and No Arbitrage," working paper, Kellogg Graduate School of Management, Northwestern University.

Brown, D. P., and M. R. Gibbons, 1985, "A Simple Econometric Approach for Utility-based Asset Pricing Models," *Journal of Finance,* 40, 359-381.

Burnside, C., 1994, "Hansen-Jagannathan Bounds as Classical Tests of Asset Pricing Models, *Journal of Business and Economic Statistics,* 12, 57-79.

Cecchetti, S. G., P. Lam, and N. C. Mark, 1994, "Testing Volatility Restrictions on Intertemporal Marginal Rates of Substitution Implied by Euler Equations and Asset Returns," *Journal of Finance,* 49, 123152.

Chen, Z., and P. J. Knez, 1995, 'Measurement of Market Integration and Arbitrage," forthcoming in *Review of Financial Studies.*

Cochrane, J. H., and L. P. Hansen, 1992, "Asset Pricing Explorations for Macroeconomics," *NBER Macroeconomics Annual 1992,* MIT Press, Cambridge, 115-165.

De Santis, G., 1993, "Volatility Bounds for Stochastic Discount Factors: Tests and Implications from International Stock Returns; Ph.D. Dissertation, University of Chicago.

Duffie, D., and K. J. Singleton, 1993, Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica,* 61, 929-952.

Epstein, L G., and S. E. Zin, 1991, "Substitution, Risk Aversion, and Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy,* 99,263 286.

Foley, D. K., 1970, "Economic Equilibrium with Costly Marketing," *Journal of Economic Theory,* 2, 276-291.

Goldberger, A., 1968, *Topics In Regression Analysis,* Macmillan, New York.

Gordin, M. I., 1969, "The Central Limit Theorem for Stationary Processes," *Soviet Mathematics* Doklady, 10, 1174-1176.

Haberman, S. J., 1989. "Concavity and Estimation," *Annals of Statistics,* 17, 1631-1661.

Hall, P., and C. C. Heyde, 1980, Martingale Limit Theory and Its Application, New York, Academic Press.

Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* 50, 1029-1054.

Hansen, L P., 1985, "A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators," *Journal of Econometrics,* 30, 203-238.

Hansen, L. P., and R. Jagannathan, 1988, "Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Prices,' working paper, University of Chicago.

Hansen, L. P., and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy.* 99, 225-262.

Hansen. L P., and R. Jagannathan, 1994, "Assessing Specification Errors in Stochastic Discount Factor Models," NBER Technical Working Paper no. 153.

Hansen, L. P., and K. J. Singleton, 1982, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica,* 50, 1269-1286.

He, H., and D. M. Modest, 1993, "Market Frictions and Consumption-Based Asset Pricing," working paper, University of California at Berkeley.

Heaton, J., 1994, "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications," forthcoming in *Econometrica.*

Hildenbrand, W., 1974, *Core and Equilibria of a Large Economy,* Princeton University Press, Princeton, N.J.

Hindy, A.. 1993, "Viable Prices in Markets with Solvency Constraints." forthcoming in *Journal of Mathematical Economics.*

Huber, P., 1981, *Robust Statistics,* Wiley, New York.

Jouini, E., and H. Kallal, 1993, "Martingales, Arbitrage and Equilibrium in Securities Markets with Transaction Costs,' forthcoming, *Journal of Economic Theory.*

Knez, P. J., 1993, "Pricing Money Market Securities with Stochastic Discount Factors," working paper, University of Wisconsin.

Kreps, D. M., 1981, "Arbitrage and Equilibrium in Economies with Infinitely Many Commodities," *Journal of Mathematical Economics,* 15-35.

Lee, B., and B. Ingram, 1991, "Simulation Estimation of Time Series Models," *Journal of Econometrics,* 47,197-205.

Luttmer, E. G. J., 1994, "Asset Pricing in Economies with Frictions," working paper, Kellogg Graduate School of Management, Northwestern University.

MacKinlay, A. C, and M. P. Richardson, 1991. "Using Generalized Method of Moments To Test Mean-Variance Efficiency," *Journal of Finance,* 46, 511-527.

Manski, C, 1988, *Analog Estimation Methods in Econometrics,* Chapman and Hall, New York.

McFadden, D., 1989, "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica,* 57, 995-1026.

Newey, W. K., and K. West, 1987. "A Simple Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* 55,703-708.

Pakes, A., and D. Pollard, 1989, "Simulation and the Asymptotics of Optimization Estimators," *Econometrica,* 57, 1027-1057.

Pollard, D., 1985, "New Ways to Prove Central Limit Theorems," *Econometric Theory,* 1, 295-314.

Prisman, E. Z., 1986, "Valuation of Risky Assets in Arbitrage Free Economies with Frictions," *Journal of Finance,* 41, 545-559.

Rockafellar, R. T., 1970, *Convex Analysis,* Princeton University Press, Princeton, NJ.

Shanken, J., 1987, "Multivariate Proxies and Asset Pricing Relations: Living with the Roll Critique, *Journal of Financial Economics,* 18, 91-110.

Snow, K. N., 1991, "Diagnosing Asset Pricing Models Using the Distribution of Asset Returns," *Journal of Finance,* 46, 955-984.

Vuong, Q. H., 1989, "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica,* 57, 307-333.

Wolak, F. A., 1991, "The Local Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models," *Econometrica,* 59, 981-995.