Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets

John H. Cochrane, Jesus Saa-Requejo

The Journal of Political Economy, Volume 108, Issue 1 (Feb., 2000), 79-119.

Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28200002%29108%3A1%3C79%3ABAGAPB%3E2.0.CO%3B2-1

Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets

John H. Cochrane

University of Chicago, Federal Reserve Bank of Chicago, and National Bureau of Economic Research

Jesús Saá-Requejo

Vega Asset Management LLC

One often wants to value a risky payoff by reference to prices of other assets rather than by exploiting full-fledged economic models. However, this approach breaks down if one cannot find a perfect replicating portfolio. We impose weak economic restrictions to derive usefully tight bounds on asset prices in this situation. The bounds assume that investors would want to buy assets with high Sharpe ratios—"good deals"—as well as pure arbitrage opportunities. We show how to calculate the price bounds in one-period, multiperiod, and continuous-time contexts. We show that the multiperiod problem can be solved recursively as a sequence of one-period problems. We calculate bounds in option pricing examples including infrequent trading and an option written on a nontraded event, and we use the bounds to explore the economic significance of option pricing predictions. We find that much variation in S&P 500 index option prices over time and across strike prices fits within the bounds.

I. Introduction

The fundamental question of financial economics is how to value uncertain payoffs. For many applications in economics and finance

We thank David Backus, George Constantinides, Kent Daniel, Douglas Diamond, Andrea Eisfeldt, Lars Hansen, Ken Singleton, Steve Ross, many seminar participants, and especially Yacine Aït-Sahalia for useful comments. This research was partially supported by the University of Chicago Graduate School of Business and by a grant from the National Science Foundation administered by the NBER.

[*Journal of Political Economy*, 2000, vol. 108, no. 1]
 0022-3808/2000/10801-0004$02.50

a *relative pricing* approach is appropriate. In these applications we are interested only in the value of a specific payoff, we take as given the prices of other assets without questioning their fundamental economic determinants, and we want to make as few economic assumptions as possible.

Option pricing is the classic case: we want to know the value of an option given the price of the underlying stock. The theory of Black and Scholes (1973) and Merton (1973) is a great success of this approach. Even though an option payoff is quite different from that of a stock or bond, they showed how continuous dynamic trading can complete a market. Therefore, the option payoff can be perfectly replicated by continuous trading in the stock and bond, and the option's value can be determined given that of the stock and bond. Via its role in option pricing theory, the relative method is used in many economic applications of finance theory, such as the real-option value of irreversible investment (e.g., Dixit and Pindyck 1994; Abel and Eberly 1996). Ross's (1976*a*) arbitrage pricing theory tries to determine the expected rate of return of a portfolio given the expected returns of "factor" portfolios, and without preference assumptions. Even apparently economics-based theories such as the capital asset pricing model (CAPM) are typically applied in a relative way since one determines expected returns on assets, portfolios, or investment projects, taking the market expected return as given. A relative approach is, appropriately, used for many applications of asset pricing theory to corporate finance. For example, one examines the returns on "like" securities to determine the cost of capital for a specific project.

Alas, a purely preference-free approach often breaks down. One may not be able to trade continuously; there may be state variables (e.g., stochastic stock volatility and interest rate) that do not correspond to traded assets; and the event on which an option or real payoff depends may not be a traded asset or it may be so thinly traded that it is not useful for hedging. In fact, if options really could be perfectly and costlessly replicated by other liquid assets, it is unlikely that options would be traded in the first place.

We use a *little* economics, a slight strengthening of no-arbitrage and law of one price arguments, to greatly restrict the range of values of a risky payoff in these situations, without having to fall back on completely specified economic models such as the consumption-based asset pricing model.

The basic idea is most simply explained in a one-period environment. We want to learn about the value of a *focus payoff* x^c_{t+1}, taking as given the prices $\mathbf{p}_t$ of a set of *basis payoffs* or hedging assets $\mathbf{x}_{t+1}$.

A discount factor or marginal utility growth rate m_{t+1} generates the value p_t of any payoff x_{t+1} by

$$p = E(mx). \tag{1}$$

Here and below we suppress time subscripts unless they are necessary for clarity. The payoff, value, and discount factor can be real or nominal. The existence of such a discount factor or marginal utility has a portfolio interpretation: it is equivalent to the law of one price that any two ways of constructing the same payoff have the same value. Thus, if the focus payoff x^c can be perfectly replicated from the basis asset payoffs $\mathbf{x}$, we have enough information to determine its value exactly. When replication is less than perfect, however, the existence of a discount factor or law of one price says nothing about the value of the focus asset, and we need more discount factor restrictions.

The more we can restrict the discount factor, the more we can learn about asset values. We require that the discount factor price a set of basis assets, require that it be nonnegative, and impose an upper bound on its volatility. Thus the lower good-deal bound solves

$$\begin{aligned} \underline{C} &= \min_{\{m\}} E(mx^c) \\ \text{subject to } \mathbf{p} &= E(m\mathbf{x});\ m \geq 0;\ \sigma(m) \leq \frac{h}{R^f}, \end{aligned} \tag{2}$$

where $\underline{C}$ is the lower good-deal bound; m is the discount factor; x^c is the focus payoff to be valued; for example, $x^c = \max(S_T - K, 0)$ for a call option; $\mathbf{p}$ and $\mathbf{x}$ are the price and payoffs of basis assets (vectors), for example stock and bond; h is the prespecified volatility bound; R^f is the risk-free interest rate; E and σ are the conditional mean and variance; and the upper good-deal bound $\overline{C}$ solves the corresponding maximum.

"Value" means how much a particular investor with marginal utility growth m would be willing to pay for a marginal quantity of the payoff. This sense of value makes no assumptions about whether the focus payoff is traded or not, or whether the investor already holds some of it or not. However, this sense of value is not the equilibrium price of a heretofore untraded security after agents buy all they want of it.

The first constraint, $\mathbf{p} = E(m\mathbf{x})$, enforces the relative pricing idea that we take as given the prices of a set of basis assets. We use the prices of these assets to learn about the discount factor, not vice versa.

The second constraint, $m \geq 0$, is a classic and weak characteriza-

tion of marginal utility. (The distinction between $m > 0$ and $m \geq 0$ is unimportant for our results.) This assumption also has a portfolio interpretation (Ross 1976*b*; Hansen and Richard 1987): it is equivalent to the absence of arbitrage opportunities, which means that if a payoff is nonnegative in every state of nature, its value must also be nonnegative. If we know, perhaps by a risk attitude survey, that an investor will take part in any arbitrage opportunity, then we know that his marginal utility is nonnegative. The problem above with the first two constraints leads to well-known *arbitrage bounds* on the value.

The volatility constraint $\sigma(m) \leq h/R^f$ is our innovation. We intend it as a similar weak restriction on marginal utility, a natural next step when absence of arbitrage alone does not give precise enough answers. It also has a portfolio interpretation. Hansen and Jagannathan (1991) show that a discount factor volatility restriction is equivalent to an upper limit on the *Sharpe ratio* of mean excess return to standard deviation. Precisely, they show that

$$E(mR^e) = 0 \quad \text{if and only if} \quad \frac{|E(R^e)|}{\sigma(R^e)} \leq \frac{\sigma(m)}{E(m)},$$

and we have $E(m) = 1/R^f$ if there is a risk-free rate. Therefore, if we know, perhaps by a risk attitude survey, that an investor will take part at the margin in any portfolio that delivers a Sharpe ratio greater than h, then we know that his marginal utility satisfies $\sigma(m) \leq h/R^f$. We call the bounds on value calculated with this additional constraint *good-deal* bounds because they assume that investors would want to trade any good deals—large Sharpe ratios—as well as pure arbitrage opportunities.

There is a long tradition in finance that regards high Sharpe ratios as "good deals" that are unlikely to survive, since investors would quickly grab them up. Ross (1976*a*) bounds asset pricing theory residuals by assuming that no portfolio can have more than twice the market Sharpe ratio, which Shanken (1992) calls "approximate arbitrage." MacKinlay (1995) criticizes Fama and French (1993) by noting what seem like excessively high Sharpe ratios. Ledoit (1995) calls a high Sharpe ratio a "δ arbitrage" and rules it out.

The discount factor volatility constraint is also a way of imposing weak or robust predictions of economic models. One may not wish to impose the full structure of an economic asset pricing model, for example a utility function and a specification of the joint distribution of consumption and asset payoffs. Still, a wide range of such models imply that marginal rates of substitution are not outrageously volatile, as well as positive. Furthermore, the standard deviation of the discount factor, while a weak prediction of such a

model, may be more robust to model and data specification errors than are the covariances of payoffs with the discount factor by which such models generate values.

Similarly, the CAPM specifies that the market portfolio is mean-variance efficient. If we view the CAPM theory and market proxy for the wealth portfolio as approximations, then we believe that the market portfolio should not be too inefficient: Sharpe ratios of other portfolios should not be dramatically higher than that of the market portfolio. On the other hand, if the market return is even slightly inefficient, covariances with the market return can generate arbitrarily large pricing errors (Kandel and Stambaugh 1995).

Finally, the volatility constraint is an easy way to prune unreasonable discount factors within the arbitrage bounds. For example, the lower arbitrage bound $C = 0$ for a call option requires marginal utility that is zero for all states of nature in which the option finishes in the money. The upper arbitrage bound $C = S$ requires a discount factor that is nonzero only in the two states of nature with the most extreme stock prices. Though nonnegative, these discount factors are unlikely characterizations of anyone's marginal utility. These discount factors vary a great deal across states of nature and have a high variance. The volatility constraint weeds out some of these arbitrage-free but still "unreasonable" discount factors and their corresponding option prices.

As a simple example, consider a call option on the Standard & Poor's 500 index with strike price $K = \$100$, three months to expiration, and no intermediate trading. Figure 1 presents the upper and lower good-deal bounds for this case.

We use parameter values $E(R) = 13$ percent, $\sigma(R) = 16$ percent for the stock index return, and a risk-free rate $R^f = 5$ percent. (We report all parameters at an annual frequency and adjust to the appropriate time horizon in the calculations.) To calibrate the discount factor volatility constraint, we assume that the investor would take any opportunity with a Sharpe ratio twice that of the S&P 500, $h = 2 \times [E(R - R^f)/\sigma(R)] = 1.0$. Since most fund managers seem desperate for average returns a few percent above the S&P 500 index, this value seems conservative. This value doubles the already troubling equity premium puzzle; it implies that the standard deviation of marginal utility growth $\sigma(m) = h/R^f$ is equal to its mean, $E(m) = 1/R^f$, and nearly 100 percent per year. However, this is the central parameter that a user must input to the calculation; it is easy to change it, and our contribution is to show how to calculate good-deal bounds for whatever value of this limit that the user thinks is appropriate, not to advocate a specific value.

The figure includes the lower arbitrage bounds $C \geq 0$, $C \geq K/R^f$.

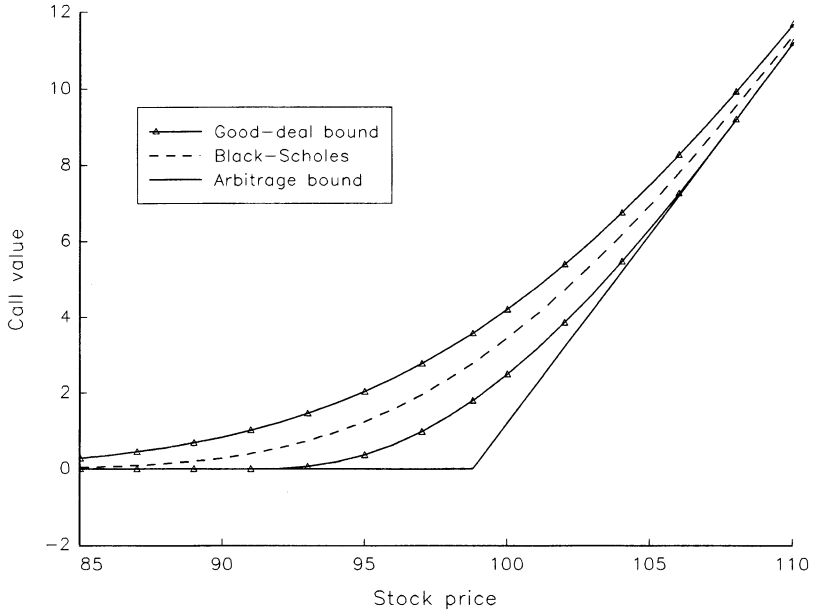

FIG. 1.—Option price bounds as a function of stock price. Options have three months to expiration and strike price $K = \$100$. The bounds assume no trading until expiration and a discount factor volatility bound $h = 1.0$ corresponding to twice the market Sharpe ratio. The stock is lognormally distributed with parameters calibrated to an index option.

The upper arbitrage bound states that $C \leq S$, but this 45-degree line is too far up to fit on the vertical scale and still see anything else. As in many practical situations, the arbitrage bounds are so wide that they are of little use.

The upper good-deal bound is much tighter than the upper arbitrage bound. For example, if the stock price is \$95, the entire range of option prices between the upper bound of \$2 and the upper arbitrage bound of \$95 is ruled out. The lower good-deal bound is the same as the lower arbitrage bound for stock prices less than about \$90 and greater than about \$110. In between \$90 and \$110, the good-deal bound improves on the lower arbitrage bound.

The width of the bounds is larger, about \$1, at the money than it is far in the money or out of the money. Options are hardest to hedge at the money because the nonlinearity of the option payoff as a function of stock price is greatest here. Therefore, the residual—option payoff less best approximate hedge—is largest in this region. However, the width of the bounds is a much larger *fraction* of the call option value for out-of-the-money options on the left-

hand side of the graph. In this sense, as well as when translated to implied volatilities, the bounds are wider for out-of-the-money options.

Though one is naturally inclined to look for small bounds, large bounds can be as interesting as small bounds, and maybe more so. Small bounds confirm that replication arguments are good approximations. Large bounds warn us that replication arguments are poor approximations, that assumptions about unmeasured market prices of risk are important to the answer, and that further discount factor restrictions would be useful to restrict the range of values.

Figure 1 includes the Black-Scholes option value for reference, although it does not apply to this example since the investor cannot trade continuously. The good-deal bounds converge to Black-Scholes as the rebalancing interval is made more frequent, but they converge to a line that differs from the Black-Scholes value as the volatility bound is lowered while maintaining a fixed trading interval. Still, it is nice to see that the good-deal bounds include the Black-Scholes value and have the same general shape, since the Black-Scholes value is often an excellent approximation to observed option prices.

Not all values outside the good-deal bounds imply high Sharpe ratios or arbitrage opportunities. Such values might be generated by a positive but highly volatile discount factor, and generated by *another* less volatile but sometimes negative discount factor, but no discount factor generates these values that is *simultaneously* nonnegative and respects the volatility constraint.

It makes sense to define bounds as we do—and to rule out these values—to intersect discount factor restrictions ($m > 0$, $\sigma(m) \leq h/R^f$) rather than to intersect the value regions (no-arbitrage, limited Sharpe ratio) formed from each discount factor restriction simultaneously. If we know that an investor will invest in any arbitrage opportunity or take any Sharpe ratio greater than h, then we know that his marginal utility satisfies both restrictions. He would find a utility-improving trade for values outside the good-deal bounds, even though those values may not imply a high Sharpe ratio, an arbitrage opportunity, or any other simple portfolio interpretation. Simple portfolio interpretations, while historically important, are likely to fall by the wayside as we add more discount factor restrictions or intersect simple ones. Furthermore, our method for extending the problem to handle multiple periods requires a nonnegative discount factor, so again we must impose both constraints.

On the other hand, in many applications the good-deal region is only very slightly smaller than the intersection of a limited Sharpe ratio region and the arbitrage-free region, and we have analytic for-

mulas for these two regions. The difference gets smaller as the time interval gets shorter and disappears entirely in the continuous-time limit. Therefore, intersecting the limited Sharpe ratio and arbitrage-free regions may be a convenient approximation for many applications.

Good-deal bounds should be useful in many situations in which a relative pricing approach is appropriate but perfect replication is not possible. A few examples follow: (1) A trader can use the bounds as buy and sell points in the search for "good deals" in asset markets (with the usual warning question of why the market leaves good deals undiscovered). (2) A bank or other institution that markets or synthesizes nontraded securities can use good-deal bounds as bid and ask prices. (3) Good-deal bounds can be used as economic measures of the accuracy of option pricing formulas. Arbitrage-based formulas predict no error, so that "measurement errors" in prices must be tacked on to the models. The bounds can tell us which option prices should lie close to arbitrage-based formulas and which can lie far from the predictions of those formulas, using an economic measure of distance. (4) Option pricing techniques are increasingly applied to "real options" in capital budgeting, investment with irreversibilities, and policy questions. A relative pricing approach is appropriate, but the focus payoffs typically cannot be perfectly replicated. (5) Option pricing formulas are often used in risk assessment to quantify the exposure of a position or institution to various risk factors. It is useful to assess such risks when perfect replication is impossible and to quantify the importance of the market price of risk assumptions.

Section II shows how to calculate good-deal bounds in single-period, multiperiod, and continuous-time contexts. We find a recursive solution to the multiperiod problem; that is, we show that the lower bound today solves the one-period problem with the lower bound tomorrow as payoff. This formulation makes the multiperiod problem computationally feasible. In continuous time, it leads to a partial differential equation for the bounds. This is our central theoretical contribution, and it makes the technique relevant for serious option pricing applications, which are all inherently dynamic. In particular, we can handle continuous trading environments in which market incompleteness comes from nontraded state variables or options written on nontradable events.

In Section III, we explore several applications, in part to emphasize that the bounds are practical and easily computable and in part for the economic interest of the answers. We show the calculations behind the simple Black-Scholes example of figure 1; we show how to bound option deltas (derivative of the option price with respect

to the stock price); we calculate option price bounds when other options are used as hedge assets; we calculate multiperiod bounds using a discrete multinomial model, and we compare the results to index option prices; and we calculate bounds on an option subject to "basis risk"—an option on a nontraded underlying security or event that is imperfectly correlated with basis asset payoffs. Cochrane and Saá-Requejo (1999) present a more extended application to option pricing in continuous time with stochastic interest rate and stock volatility.

II. Calculating Good-Deal Bounds

A. *One Period*

We start with the simplest situation. There is one period and no intermediate trading until the payoff x^c is realized. Throughout, we assume that one of the basis payoffs is riskless, so we can write $E(m) = 1/R^f$. It is convenient to express the volatility constraint as a second moment. Thus, our problem (2) becomes

$$\begin{aligned} \underline{C} &= \min_{\{m\}} E(mx^c) \\ &\text{subject to } \mathbf{p} = E(m\mathbf{x}),\ E(m^2) \leq A^2,\ m \geq 0, \end{aligned} \tag{3}$$

where $A^2 \equiv (1 + h^2)/R^{f2}$. We presume a statistical model for the distribution of the payoffs $\mathbf{x}$, so that we can calculate moments. All expectations and probabilities refer to the true measure. For any solution to exist, of course, one must pick a sufficiently large bound A to price the basis assets:

$$\begin{aligned} A^2 &\geq \min_{\{m\}} E(m^2) \\ &\text{subject to } E(m\mathbf{x}) = \mathbf{p},\ m \geq 0. \end{aligned}$$

The problem has two inequality constraints. Hence we find a solution by trying all the combinations of binding and nonbinding constraints, in order of their ease of calculation, as follows: (1) We assume that the volatility constraint binds and the positivity constraint is slack. This is the easiest case since we have analytic formulas for the bounds and discount factor in this case. If the resulting discount factor $\underline{m}$ is nonnegative, this is the solution. If not, (2) we assume that the volatility constraint is slack and the positivity constraint binds. This configuration delivers the arbitrage bound on value. We find the minimum variance discount factor that generates the arbitrage bound. If this discount factor satisfies the volatility constraint,

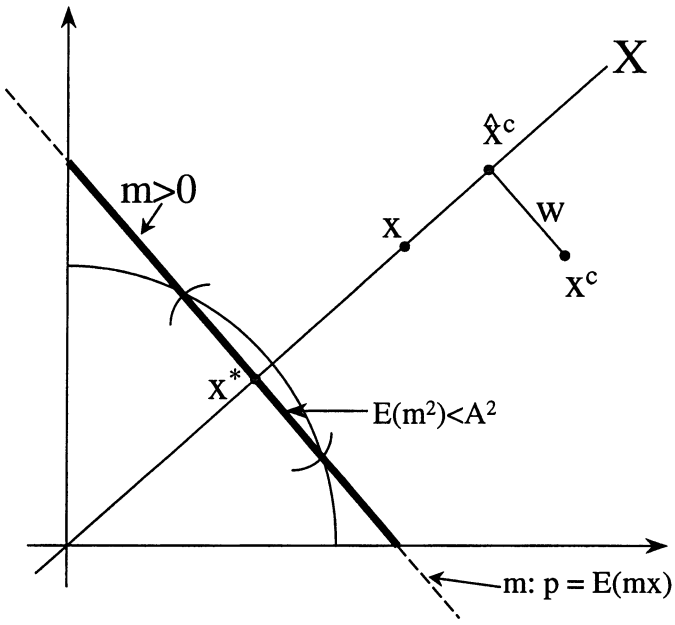

FIG. 2.—Notation and geometry: X is the space of portfolios of basis asset payoffs, $\mathbf{x}$ is the basis asset payoffs, x^c is the focus asset payoff, $\hat{x}^c$ is the best approximate hedge, and w is the residual. The straight dashed line m: $\mathbf{p} = E(m\mathbf{x})$ is the space of all discount factors that price the basis payoffs $\mathbf{x}$. The solid straight line marked $m > 0$ is the space of all positive discount factors that price the basis assets. The arc is the set of all random variables with second moment less than A^2.

this is the solution. If not, (3) we solve the problem with both constraints binding. Next we show how to handle each case.

1. Volatility Constraint Binds, Positivity Constraint Is Slack

If the positivity constraint is slack, the problem reduces to

$$\underline{C} = \min_{\{m\}} E(mx^c) \quad \text{subject to } \mathbf{p} = E(m\mathbf{x}),\ E(m^2) \leq A^2. \tag{4}$$

Rather than solve this problem directly with Lagrange multipliers on the constraints, we set up orthogonal decompositions of the focus payoff x^c and discount factor m. The solution then pops out. (The solution strategy is due to Hansen and Jagannathan [1991]. This problem is dual to theirs.) Figure 2 describes the idea. The symbol $X \equiv \{\mathbf{c}'\mathbf{x},\ \mathbf{c} \in R^N\}$ denotes the space of payoffs of portfolios of the

basis assets.[1] Though graphed as a line, X is typically an infinite-dimensional space. We know all prices in X, but the payoff x^c that we wish to value does not lie in X.

We orthogonally decompose the focus payoff x^c into an approximate hedge $\hat{x}^c$ and a residual w, by projection (ordinary least squares [OLS] regression):

$$\begin{aligned} x^c &= \hat{x}^c + w, \\ \hat{x}^c &\equiv \text{proj}(x^c|X) = E(x^c\mathbf{x}')E(\mathbf{x}\mathbf{x}')^{-1}\mathbf{x}, \\ w &\equiv x^c - \hat{x}^c. \end{aligned} \tag{5}$$

By construction, $E(w\hat{x}^c) = 0$ and $E(w\mathbf{x}) = 0$. As is standard in thinking about OLS regressions, figure 2 represents $E(w\mathbf{x}) = 0$ by plotting w at right angles to the X plane.

Discount factors m generate the prices of payoffs x in X by an inner product $p = E(mx)$. All points on the dashed line marked m in figure 2 have the same inner product $E(mx)$ with vectors $x \in X$ and hence generate the same prices on X. Such discount factors can be represented as $x^* = \text{proj}(m|X)$ plus orthogonal components. This fact allows us to impose the pricing constraint.

LEMMA 1. A discount factor m prices the basis assets, $\mathbf{p} = E(m\mathbf{x})$, if and only if it has the form

$$m = x^* + vw + \epsilon, \tag{6}$$

where

$$x^* = \mathbf{p}'E(\mathbf{x}\mathbf{x}')^{-1}\mathbf{x} \tag{7}$$

satisfies $p = E(x^*x)$ by construction, v is an arbitrary number, ϵ is any random variable with $E(\epsilon x) = 0$ and $E(\epsilon w) = 0$, and w is the residual defined in equation (5).

Proofs are in the Appendix. We shall quickly see that we want to choose $\epsilon = 0$, so we are reduced to choosing a single number v rather than choosing a random variable m.

Since the second moment defines distance in figure 2, the set of discount factors that satisfies the volatility constraint $E(m^2) \leq A^2$ lies inside the circle shown around the origin. This restricted range of discount factors will produce a restricted range of values for the residual w and hence for the focus payoff x^c. The maximum and minimum values will be generated when we pick v to exactly satisfy the

[1] Since we need second moments, they had better be defined. With $\mathbf{y} \equiv [\mathbf{x}' \; x^c]'$, we assume $E\mathbf{y}\mathbf{y}' < \infty$; i.e., $\mathbf{y}$ is in L^2, the space of all random variables with finite second moments. We assume that redundant assets have been pruned from the specification, so $E\mathbf{y}\mathbf{y}'$ is nonsingular. We also limit consideration to discount factors $m \in L^2$. Hansen and Jagannathan (1991) discuss these technical assumptions.

volatility constraint. Proposition 1 expresses this statement more formally.

PROPOSITION 1. The discount factor that generates the lower bound is

$$\underline{m} = x^* - \underline{v}w \tag{8}$$

and the bound is

$$\underline{C} = E(x^*x^c) - \underline{v}E(w^2), \tag{9}$$

where

$$\underline{v} = \sqrt{\frac{A^2 - E(x^{*2})}{E(w^2)}}. \tag{10}$$

The upper bound is given by $\overline{v} = -\underline{v}$.

The first term in equation (9) is the value of the approximate hedge portfolio,

$$E(x^*x^c) = E(x^*\hat{x}^c) = E(m\hat{x}^c) \tag{11}$$

for any discount factor m that prices basis assets. (Here and below we use the fact that $E(xy) = E[x \operatorname{proj}(y|X)]$.) The second term is the lowest possible valuation of the residual w consistent with the discount factor volatility bound:

$$\underline{v}E(w^2) = E(\underline{v}ww) = E[(x^* + \underline{v}w)w] = E(\underline{m}w).$$

The bounds are tighter if the volatility constraint A is smaller, if the residual is smaller, or if the approximate hedge is better. To see this precisely, we can write the size of the bounds as

$$\begin{aligned} \overline{C} - \underline{C} &= 2\sqrt{A^2 - E(x^{*2})}\,\sqrt{E(w^2)} \\ &= 2\sqrt{E(x^{c2})}\,\sqrt{A^2 - E(x^{*2})}\,\sqrt{1 - R^2}, \end{aligned} \tag{12}$$

where

$$R^2 \equiv \frac{E(\hat{x}^{c2})}{E(x^{c2})} = 1 - \frac{E(w^2)}{E(x^{c2})}.$$

In the first expression, we relate the size of the bounds to the size of the residual $\sqrt{E(w^2)}$ directly. In the second expression, we show that a higher R^2 of a regression of the focus asset on the basis assets leads to tighter bounds. Some of the discount factor volatility is "used up" in pricing the basis assets. Only that portion of the volatil-

ity bound A in excess of the discount factor volatility $E(x^{*2})$ required to price the basis assets can be applied to the price of the residual.

For calculations it is useful to substitute the definitions of x^* and w in equation (9) to obtain

$$\underline{C} = \mathbf{p}'E(\mathbf{x}\mathbf{x}')^{-1}E(\mathbf{x}x^c) - \sqrt{A^2 - \mathbf{p}'E(\mathbf{x}\mathbf{x}')^{-1}\mathbf{p}}\sqrt{E(x^{c2}) - E(x^c\mathbf{x}')E(\mathbf{x}\mathbf{x}')^{-1}E(\mathbf{x}x^c)}. \tag{13}$$

The upper bound $\overline{C}$ is the same formula with a plus sign in front of the square root. This formula is much less pretty, but it shows explicitly how to calculate the price bound from a statistical model for the second moments of $\mathbf{x}$, x^c.

Using (8), we check the assumption that the discount factor is in fact positive in every state of nature. If so, this is the good-deal bound. If not, we proceed to the next step.

2. Both Constraints Bind

Next we show how to find the bounds in a one-period model when both constraints bind. Though this is the third step in the procedure, it is easiest to describe this case now. In the geometry of figure 2, it is not necessarily true that the $E(m^2) \leq A^2$ set lies inside the $m \geq 0$ set. If it does for some states of nature and not for others, then both constraints bind.

When we introduce Lagrange multipliers, the problem is

$$\underline{C} = \min_{\{m>0\}} \max_{\{\lambda,\delta>0\}} E(mx^c) + \lambda'[E(m\mathbf{x}) - \mathbf{p}] + \frac{\delta}{2}[E(m^2) - A^2].$$

The first-order conditions to this problem yield a discount factor that is a truncated linear combination of the payoffs:

$$m = \max\left(-\frac{x^c + \lambda'\mathbf{x}}{\delta}, 0\right) = \left[-\frac{x^c + \lambda'\mathbf{x}}{\delta}\right]^+. \tag{14}$$

The last equality defines the $[\cdot]^+$ notation for truncation. To derive this expression, take partial derivatives with respect to m in each state. We could plug expression (14) into the constraints and solve numerically for Lagrange multipliers λ and δ that enforce the constraints. Alas, this procedure requires the solution of a system of nonlinear equations in (λ, δ), which is often a numerically difficult or unstable problem.

Hansen, Heaton, and Luttmer (1995) show how to recast the

problem as a maximization, which is numerically much easier. Interchanging min and max, we get

$$\underline{C} = \max_{\{\lambda,\delta>0\}} \min_{\{m>0\}} E(mx^c) + \lambda'[E(m\mathbf{x}) - \mathbf{p}] + \frac{\delta}{2}[E(m^2) - A^2]. \quad (15)$$

The inner minimization yields the same first-order conditions (14). Plugging those first-order conditions into the outer maximization of (15) and simplifying, we obtain

$$\underline{C} = \max_{\{\lambda,\delta>0\}} E\left\{-\frac{\delta}{2}\left[-\frac{x^c + \lambda'\mathbf{x}}{\delta}\right]^{+2}\right\} - \lambda'\mathbf{p} - \frac{\delta}{2}A^2. \quad (16)$$

We search numerically over (λ, δ) to find the solution to this problem. The upper bound is found by replacing max with min and replacing $\delta > 0$ with $\delta < 0$.

3. Positivity Binds, Volatility Is Slack

If the volatility constraint is slack and the positivity constraint binds, the problem reduces to

$$\underline{C} = \min_{\{m\}} E(mx^c)$$

$$\text{subject to } \mathbf{p} = E(m\mathbf{x}),\ m > 0.$$

These are the *arbitrage bounds.* One can often deduce the arbitrage bounds for specific problems without explicitly solving a minimization. It is a linear program otherwise. It remains, however, to check whether the discount factor volatility constraint can be satisfied at the arbitrage bound.

Denote the lower arbitrage bound by C_l. The minimum variance (second-moment) discount factor that generates the arbitrage bound C_l solves

$$E(m^2)_{\min} = \min_{\{m\}} E(m^2)$$

$$\text{subject to } \begin{bmatrix} \mathbf{p} \\ C_l \end{bmatrix} = E\left(m\begin{bmatrix} \mathbf{x} \\ x^c \end{bmatrix}\right),\ m > 0.$$

With the same conjugate method, this problem is equivalent to

$$E(m^2)_{\min} = \max_{\{\nu,\mu\}} -E\{[-(\mu x^c + \mathbf{v}'\mathbf{x})]^{+2}\} - 2\mathbf{v}'\mathbf{p} - 2\mu C_l.$$

Again, we search numerically for $(\mathbf{v}, \mu)$ to solve this problem. If $E(m^2)_{\min} \leq A$, C_l is the solution to the good-deal bound; if not, we proceed with the case in which both constraints are binding described above.

B. *Multiple Periods: A Recursive Solution*

Consider the same problem, but allow one intermediate period at which the investor can change his portfolio of the basis assets. To keep the notation simple, we do not include intermediate payoffs, but this is a simple extension. We use time periods 1 and 2 for the rebalancing date and final date, respectively. The objective function for the lower bound is then

$$\underline{C}_0 = \min_{\{m_1, m_2\}} E_0(m_1 m_2 x_2^c),$$

where $E_t(\cdot) = E(\cdot | I_t)$ denotes conditional expectation.

Next, the constraints. The discount factors m_1 and m_2 must price the basis assets, so we require

$$\mathbf{p}_0 = E_0(m_1 \mathbf{p}_1),$$
$$\mathbf{p}_1 = E_1(m_2 \mathbf{x}_2) \quad \text{for all information sets } I_1.$$

We generalize the discount factor volatility constraint to the requirement that the volatility *in each period* respect a bound[2]

$$E_0(m_1^2) \leq A_0^2; \quad E_1(m_2^2) \leq A_1^2. \tag{17}$$

The volatility constraints may be functions of information (state) and thus allow time-varying risk premia. Finally, we require discount factors to be nonnegative: $m_1 \geq 0$ and $m_2 \geq 0$. Now we can state the two-period problem and its recursive solution.

[2] An alternative constraint might be that the *two-period* Sharpe ratio is below a given value

$$E_0[(m_1 m_2)^2] \leq A^2.$$

This constraint does not lead to a recursive solution. The constraint (17) implies a constraint of the form

$$E_0[(m_1 m_2)^2] = E_0[m_1^2 E_1(m_2^2)] \leq E_0(m_1^2 A_1^2) \leq A_0^2 A_1^2,$$

but the converse is not true. Therefore, individual constraints give sharper bounds than a two-period constraint. The Sharpe ratio scales with the square root of the time period. An appropriate numerical value for the bound using a six-month trading interval is the square root of the appropriate value for a one-year trading interval.

PROPOSITION 2. The two-period problem

$$\underline{C}_0 = \min_{\{m_1, m_2\}} E_0(m_1 m_2 x_2^c)$$

subject to

$$\mathbf{p}_0 = E_0(m_1 \mathbf{p}_1), \quad E_0(m_1^2) \leq A_0^2, \quad m_1 \geq 0, \tag{18}$$

$$\mathbf{p}_1 = E_1(m_2 \mathbf{x}_2), \quad E_1(m_2^2) \leq A_1^2 \quad \forall\, I_1,\ m_2 \geq 0 \tag{19}$$

has the same solution as the sequence of one-period problems:

$$\begin{aligned} &\underline{C}_1 = \min_{\{m_2\}} E_1(m_2 x_2^c) \\ &\text{subject to } \mathbf{p}_1 = E_1(m_2 \mathbf{x}_2), \quad E_1(m_2^2) \leq A_1^2, \quad m_2 \geq 0, \\ &\underline{C}_0 = \min_{\{m_1\}} E_0(m_1 \underline{C}_1) \\ &\text{subject to } \mathbf{p}_0 = E_0(m_1 \mathbf{p}_1), \quad E_0(m_1^2) \leq A_0^2, \quad m_1 \geq 0. \end{aligned} \tag{20}$$

The proof is in the Appendix. Basically, since $m_1 \geq 0$ and the constraint sets on m_1 and m_2 do not affect each other, the solution to the two-period problem min $E_0[m_1 E_1(m_2 x^c)]$ must minimize $E_1(m_2 x^c)$ in each state of nature at time 1. If $m_1 < 0$ were possible, we would want to maximize $E_1(m_2 x^c)$ in some states of nature; the $m_1 \geq 0$ assumption rules this out. Since the bounds are not prices or payoffs of traded securities, it is not initially obvious that one can compute the date 0 bound using the date 1 bound as a "payoff," but the proposition verifies that it is correct to do so.

The recursive statement extends in an obvious way to multiple periods. To compute multiperiod bounds, we can now work backward: for a payoff at T, compute one-period bounds as before at $T - 1$. Then compute bounds at $T - 2$ using the $T - 1$ bounds as payoffs, and so forth.

C. *Continuous Time*

1. Notation

Passing to continuous time is conceptually straightforward, but the notation is unavoidably a bit different.

Consider an asset with price S_t that gives a stream of payoffs or dividends $D_t dt$. A discount factor is a process Λ_t that generates the price by

$$S_t \Lambda_t = E_t \int_{s=0}^{\infty} \Lambda_{t+s} D_{t+s}\, ds. \tag{21}$$

The continuous-time and discrete-time discount factor concepts are related by $m_{t+1} = \Lambda_{t+1}/\Lambda_t$. Again, we suppress time subscripts where they are not necessary; for example, we write $\Lambda = \Lambda_t$.

Using (21) at t and $t + \Delta t$, we can find the continuous-time equivalent to $p = E(mx)$:

$$0 = \frac{E_t[d(\Lambda S)]}{\Lambda S} + \frac{D}{S}dt. \tag{22}$$

If there is an instantaneously risk-free rate r_t (a security with price one that pays $D_t = r_t$, or a money market account whose value grows at $dB/B = r_t dt$), then equation (22) implies

$$E_t\left(\frac{d\Lambda}{\Lambda}\right) = -r_t dt.$$

This equation is the continuous-time counterpart to $R^f = 1/E(m)$.

Equation (22) implies

$$E_t\frac{dS}{S} + \left(\frac{D}{S} - r\right)dt = -E_t\left(\frac{d\Lambda}{\Lambda}\frac{dS}{S}\right). \tag{23}$$

This equation is the continuous-time counterpart to $E(R^e) = -\text{cov}(m, R^e)/E(m)$ for excess returns R^e. Thus a security with diffusion dz should have an expected excess return equal to

$$-\frac{1}{dt}E_t\left(\frac{d\Lambda}{\Lambda}dz\right).$$

This quantity is the "market price" of the risk dz, and this equation allows us to translate between market price of risk restrictions and discount factor restrictions.

The continuous-time equivalent to the link between Sharpe ratios and discount factor volatility, $|E(R^e)|/\sigma(R^e) \leq \sigma(m)/E(m)$, can be found from (22) as

$$\frac{(\tilde{\mu}_S - r)^2}{\sigma_S^2}dt \leq E_t\frac{d\Lambda^2}{\Lambda^2},$$

where $\tilde{\mu}_S \equiv E_t(dS/S) + (D/S)\,dt$ is the conditional expected return and $\sigma_S^2 \equiv E_t(dS^2/S^2)$ is the conditional variance of return.

2. Statistical Model

We need a statistical model, the equivalent of the moments of the basis and focus assets $E(\mathbf{x})$, $E(\mathbf{x}\mathbf{x}')$, $E(x^c)$, and so forth in the discrete-time formulas. The statistical model must describe the conditional moments at every point in time.

We model the price processes of an n_S-dimensional vector of basis assets by a diffusion[3]

$$\frac{dS}{S} = \mu_S(S, V, t)\,dt + \sigma_S(S, V, t)\,dz, \tag{24}$$

where z is an n_S-dimensional vector of independent Brownian motions, and $E(dz\,dz') = I$. The basis assets may pay dividends at rate $D(S, V, t)\,dt$. The term V represents an n_V-dimensional vector of additional state variables that follow

$$dV = \mu_V(S, V, t)\,dt + \sigma_{Vz}(S, V, t)\,dz + \sigma_{Vw}(S, V, t)\,dw, \tag{25}$$

where dw is an n_V-dimensional vector of Brownian increments orthogonal to dz: $E(dw\,dw') = I$; $E(dw\,dz') = 0$. We again assume that there is an instantaneously risk-free rate $r(S, V, t)\,dt$. This assumption is not essential, but it simplifies the algebra and presentation. When r varies stochastically, it is one of the state variables in V.

3. The Problem

We want to value an asset that pays continuous dividends at rate $x^c(S, V, t)\,dt$ and with a terminal payment $x_T^c(S, V, T)$. We might not be able to perfectly hedge or replicate this asset for two reasons: The risks associated with n_V nontraded shocks cannot be hedged, and the payoff x^c may depend on the nontraded variables V.

The problem is now to choose a discount factor *process* to minimize the asset value

$$\underline{C}_t = \min_{\{\Lambda_s, t<s\le T\}} E_t\int_{s=t}^{T} \frac{\Lambda_s}{\Lambda_t} x_s^c\,ds + E_t\left(\frac{\Lambda_T}{\Lambda_t} x_T^c\right) \tag{26}$$

[3] Understand division to operate element by element on the vectors, e.g., $dS/S = [dS_1/S_1\ dS_2/S_2\ \ldots]$. When explicit enumeration of arguments is not necessary, we write S for $S(t)$ and μ_S and σ_S for $\mu_S(S, V, t)$ and $\sigma_S(S, V, t)$. We assume that all diffusion parameters $\mu_S(S, V, t)$, $\sigma_S(S, V, t)$, $\mu_V(S, V, t)$, $\sigma_V(S, V, t)$, etc. are continuous in all their arguments. We assume that all variance-covariance matrices such as $\sigma_S(S, V, t)\sigma_S(S, V, t)'$, $\sigma_V(S, V, t)\sigma_V(S, V, t)'$ are nonsingular for all $S \in \mathbb{R}^{n_S}$, $V \in \mathbb{R}^{n_V}$, $t \in [0, T]$.

subject to the constraints that (1) the discount factor prices the basis assets S, r at each moment in time, (2) the instantaneous volatility of the discount factor process is less than a prespecified value A^2 (or, more generally, less than a process $A(S, V, t)^2$ at each moment in time), and (3) the discount factor is positive, $\Lambda_s > 0$, $t \leq s \leq T$. The upper bound follows by replacing min with max. Since there are no jumps in news variables, we add to our list of economic assumptions on the discount factor that it also follows a diffusion process without jumps.

4. Differential Statements

The problem can be solved recursively, by proposition 2. Thus we can study how to move one step back in time,

$$\underline{C}_t\Lambda_t = \min_{\{\Lambda_s\}} E_t \int_{s=t}^{t+\Delta t} \Lambda_s x_s^c ds + E_t(\Lambda_{t+\Delta t}\underline{C}_{t+\Delta t}),$$

or, for small time intervals,

$$\underline{C}_t\Lambda_t = \min_{\{\Delta\Lambda\}} E_t[x_t^c\Delta t + (\underline{C}_t + \Delta\underline{C})(\Lambda_t + \Delta\Lambda)].$$

Letting $\Delta t \to 0$, we can write the objective in differential form,

$$0 = \frac{x_t^c}{\underline{C}}dt + \min_{\{d\Lambda\}} \frac{E_t[d(\Lambda\underline{C})]}{\Lambda\underline{C}}, \tag{27}$$

subject to the constraints. We can also write (27) as

$$E_t\frac{d\underline{C}}{\underline{C}} + \frac{x_t^c}{\underline{C}}dt - r^f dt = -\min_{\{d\Lambda\}} E_t\left(\frac{d\Lambda}{\Lambda}\frac{d\underline{C}}{\underline{C}}\right).$$

Since the second and third terms on the left-hand side are fixed, the condition sensibly tells us to find the lowest value $\underline{C}$ by maximizing the drift $E_t d\underline{C}$ at each date.

5. Constraints

Now we express the constraints. As in the discrete-time case, we orthogonalize the discount factor and then the solution pops out.

LEMMA 2. The term Λ_t is a discount factor driven by dz, dw that prices the basis assets S, r if and only if it can be represented as

$$\frac{d\Lambda}{\Lambda} = \frac{d\Lambda^*}{\Lambda^*} - v dw, \tag{28}$$

where

$$\frac{d\Lambda^*}{\Lambda^*} \equiv -rdt - \tilde{\mu}_S' \Sigma_S^{-1} \sigma_S dz,$$

$$\tilde{\mu}_S \equiv \mu_S + \frac{D}{S} - r;$$

$$\Sigma_S = \sigma_S \sigma_S',$$

and v is an arbitrary $1 \times n_V$ matrix.[4]

This proposition is the obvious continuous-time counterpart to lemma 1 and has the same geometric interpretation as in figure 2. We can let $d\Lambda/\Lambda$ load on additional shocks, orthogonal to both V and S, with no effect on its ability to price focus or basis assets. For this reason the proposition qualifies "driven by dz, dw." However, our minimization or maximization of asset values will again put such loading to zero. (This proposition and proof are similar to proposition 3.1 in He and Pearson [1992].)

The volatility constraint is

$$\frac{1}{dt} E_t \frac{d\Lambda^2}{\Lambda^2} \leq A^2,$$

and hence, from (28),

$$vv' \leq A^2 - \frac{1}{dt} E_t \frac{d\Lambda^{*2}}{\Lambda^{*2}} = A^2 - \tilde{\mu}_S' \Sigma_S^{-1} \tilde{\mu}_S. \tag{29}$$

By expressing the constraints via (28) and (29), we have again reduced the problem of choosing the stochastic process for Λ to the choice of loadings v on the noises dw with unknown values, subject to a quadratic constraint on vv'. Since we are picking differentials

[4] We require

$$E\left[\exp\left(\frac{1}{2}\int_0^T |\tilde{\mu}_S' \Sigma_S^{-1} \sigma_S|^2 dt\right)\right] < \infty$$

and

$$E\left[\exp\left(\frac{1}{2}\int_0^T |v|^2 dt\right)\right] < \infty$$

to ensure that the stochastic integrals that describe the dynamics of Λ are well defined.

and have ruled out jumps, the positivity constraint on the choice of $d\Lambda$ is slack as long as $\Lambda > 0$.

From equation (28), v is the vector of market prices of risks of the dw shocks:

$$-\frac{1}{dt}E\left(\frac{d\Lambda}{\Lambda}dw\right) = v.$$

Thus the problem is equivalent to finding at each date the assignment of market prices of risk to the dw shocks that minimizes (maximizes) the focus payoff value, subject to the constraint that the total (sum of squared) market price of risk is bounded by A^2.

6. Solutions: A Differential Characterization

At each moment, the bound calculation is now exactly the same as in the one-period discrete-time case with a slack positivity constraint. However, except for the moment just before the terminal date, the focus payoff is the next period's lower bound. Thus we obtain a lower bound at each moment that depends on the distribution of the lower bound at the next moment. These results must be strung together in order to obtain the lower bound at each moment in terms of the underlying assumed stochastic properties of the focus and basis assets.

To be specific, we assume that the lower bound $\underline{C}$ follows a diffusion process, so we write

$$\frac{d\underline{C}}{\underline{C}} + \mu_{\underline{C}}(S, V, t)\,dt + \sigma_{\underline{C}z}(S, V, t)\,dz + \sigma_{\underline{C}w}(S, V, t)\,dw. \qquad (30)$$

The terms $\sigma_{\underline{C}z}$ and $\sigma_{\underline{C}w}$ capture the stochastic evolution of the bound over the next instant. Therefore, a differential or moment-to-moment characterization of the bound will tell us $\mu_{\underline{C}}$ in terms of $\sigma_{\underline{C}z}$ and $\sigma_{\underline{C}w}$.

PROPOSITION 3. The lower bound discount factor $\underline{\Lambda}_t$ follows

$$\frac{d\underline{\Lambda}}{\underline{\Lambda}} = \frac{d\Lambda^*}{\Lambda^*} - \underline{v}dw \qquad (31)$$

and $\mu_{\underline{C}}$, $\sigma_{\underline{C}z}$, and $\sigma_{\underline{C}w}$ satisfy the restriction

$$\mu_{\underline{C}} + \frac{x^c}{\underline{C}} - r = -\frac{1}{dt}E_t\left(\frac{d\Lambda^*}{\Lambda^*}\sigma_{\underline{C}z}dz\right) + \underline{v}\sigma'_{\underline{C}w}, \qquad (32)$$

where

$$\underline{v} = \sqrt{A^2 - \frac{1}{dt}E_t\frac{d\Lambda^{*2}}{\Lambda^{*2}}}\frac{\sigma_{\underline{C}w}}{\sqrt{\sigma_{\underline{C}w}\sigma'_{\underline{C}w}}}. \tag{33}$$

The upper bound process $\overline{C}_t$ and discount factor $\overline{\Lambda}_t$ have the same representation with $\overline{v} = -\underline{v}$.

The statement and proof of this proposition are straightforward analogues to proposition 1 and have the same geometric interpretation as shown in figure 2. The term $d\Lambda^*/\Lambda^*$ is the combination of basis asset shocks that prices the basis assets by construction, in analogy to x^*. The term $\sigma_{\underline{C}w}dw$ corresponds to the error w, and $\sigma_{\underline{C}w}\sigma'_{\underline{C}w}$ corresponds to $E(w^2)$. The proposition looks a little different because now we choose a vector v rather than a number. We could define a residual $\sigma_{\underline{C}w}dw$, and then the problem would reduce to choosing a number, the loading of $d\Lambda$ on this residual. It is not convenient to do so in this case since $\sigma_{\underline{C}w}$ potentially changes over time.

As in the discrete-time case, we can plug in the definition of Λ^* to obtain explicit, if less intuitive, expressions for the optimal discount factor and the resulting lower bound:

$$\frac{d\underline{\Lambda}}{\underline{\Lambda}} = -rdt + \tilde{\mu}'_S\Sigma_S^{-1}\sigma_S dz - \sqrt{A^2 - \tilde{\mu}'_S\Sigma_S^{-1}\tilde{\mu}_S}\frac{\sigma_{\underline{C}w}}{\sqrt{\sigma_{\underline{C}w}\sigma'_{\underline{C}w}}}dw \tag{34}$$

and

$$\mu_{\underline{C}} + \frac{x^e}{\underline{C}} - r = \tilde{\mu}'_S\Sigma_S^{-1}\sigma_S\sigma_{\underline{C}z} + \sqrt{A^2 - \tilde{\mu}'_S\Sigma_S^{-1}\tilde{\mu}_S}\sqrt{\sigma_{\underline{C}w}\sigma'_{\underline{C}w}}. \tag{35}$$

7. A Partial Differential Equation for the Bounds

We now have a differential characterization of the lower bound and the discount factor that generates the lower bound. We have to chain together those differential characterizations.

First, we find a partial differential equation for the bounds. We hypothesize a solution $\underline{C}(S, V, t)$. We use Ito's lemma to derive expressions for $\mu_{\underline{C}}$ and $\sigma_{\underline{C}z}$, $\sigma_{\underline{C}w}$ in terms of the partial derivatives of $\underline{C}(S, V, t)$. We substitute these expressions into restriction (35). The mechanics are relegated to an algebraic appendix (Cochrane and Saá-Requejo 1998). The result is ugly but straightforward to evaluate numerically, and analytically in special cases. It expresses the time derivative $\partial\underline{C}/\partial t$ in terms of derivatives with respect to state variables, and thus it can be used to work back from a terminal period.

PROPOSITION 4. The lower bound $\underline{C}(S, V, t)$ is the solution to the partial differential equation

$$x^c - r\underline{C} + \frac{\partial \underline{C}}{\partial t}$$

$$+ \frac{1}{2}\sum_{i,j} \frac{\partial^2 \underline{C}}{\partial S_i \partial S_j} S_i S_j \sigma_{S_i} \sigma'_{S_j} + \frac{1}{2}\sum_{i,j} \frac{\partial^2 \underline{C}}{\partial V_i \partial V_j} (\sigma_{Vz_i} \sigma'_{Vz_j} + \sigma_{Vw_i} \sigma'_{Vw_j})$$

$$+ \sum_{i,j} \frac{\partial^2 \underline{C}}{\partial S_i \partial V_j} S_i \sigma_{S_i} \sigma'_{Vz_j}$$

$$= \left(\frac{D}{S} - r\right)' (S\underline{C}_S) + (\tilde{\mu}'_S \Sigma_S^{-1} \sigma_S \sigma'_{Vz} - \mu'_V)\underline{C}_V$$

$$+ \sqrt{A^2 - \tilde{\mu}'_S \Sigma_S^{-1} \tilde{\mu}_S} \sqrt{\underline{C}'_V \sigma_{Vw} \sigma'_{Vw} \underline{C}_V}$$

subject to the boundary conditions provided by the focus asset payoff x_T^c. The term $\underline{C}_V$ denotes the vector with typical element $\partial \underline{C}/\partial V_j$ and $(S\underline{C}_S)$ denotes the vector with typical element $S_i \partial \underline{C}/\partial S_i$. Replacing the plus sign with a minus sign before the square root gives the partial differential equation satisfied by the upper bound.

Note that the drift of the basis assets μ_S enters into the partial differential equations for the bounds. Again, actual and not just risk-neutral probabilities matter.

8. A Special Case in Which We Know the Discount Factor

In general, the $\underline{\Lambda}$ process depends on the parameters $\sigma_{\underline{C}w}$. Hence, without solving the partial differential equation above, we do not know how to spread the loading of $d\underline{\Lambda}$ across the multiple sources of risk dw whose risk prices we do not observe. Equivalently, we do not know how to optimally spread the total market price of risk across the elements of dw. Thus, in general, we cannot use an integration approach to find the bound; that is, we cannot characterize $\underline{\Lambda}$ enough simply to calculate

$$\underline{C}_t = E_t\left[\int_{s=t}^{T} \frac{\underline{\Lambda}_s}{\underline{\Lambda}_t} x_s^c \, ds\right] + E_t\left(\frac{\underline{\Lambda}_T}{\underline{\Lambda}_t} x_T^c\right).$$

However, if there is only one shock dw, then we do not have to worry about how the loading of $d\underline{\Lambda}$ spreads across multiple sources of risk. The vector v can be determined simply by the volatility constraint. In this special case, dw and $\sigma_{\underline{C}w}$ are scalars. Hence equation (31) simplifies as follows.

PROPOSITION 5. In the special case in which there is only one extra noise dw driving the V process, we can find the optimum discount factor $\underline{\Lambda}$ directly as

$$\frac{d\underline{\Lambda}}{\underline{\Lambda}} = -rdt + \tilde{\mu}_S' \Sigma_S^{-1} \sigma_S dz - \sqrt{A^2 - \tilde{\mu}_S' \Sigma_S^{-1} \tilde{\mu}_S} dw.$$

In some applications, the loading of $d\Lambda$ on multiple shocks dw may be constant over time. In such cases, one can again construct the discount factor and solve for bounds by (possibly numerical) integration, avoiding the solution of a partial differential equation.

III. Applications

A. *Black-Scholes with No Intermediate Trading*

We start by calculating bounds on call option values in the Black-Scholes setup with no trading until expiration. The results were presented in figure 1. The call option payoff is

$$x^c = \max(S_T - K, 0),$$

where S_T is the stock price at expiration, and K is the strike price. The hedge assets are the underlying stock with current price S and a risk-free bond with return R^f. The stock return $R = S_T/S$ is lognormally distributed. We use the lognormal density to calculate all the first and second moments of the stock and option payoffs. For example,

$$E(x^c) = S \int_{K/S}^{\infty} \left(R - \frac{K}{S} \right) f(R)\, dR.$$

The derivation and expression of these moments are long and unenlightening, so we relegate them to the algebraic appendix (Cochrane and Saá-Requejo 1998). With these moments in hand, we follow the procedure described in Section II*A*.

B. *Options as Hedge Assets*

The Black-Scholes setup is famous for historical reasons and because it delivers a closed-form solution with continuous trading. However,

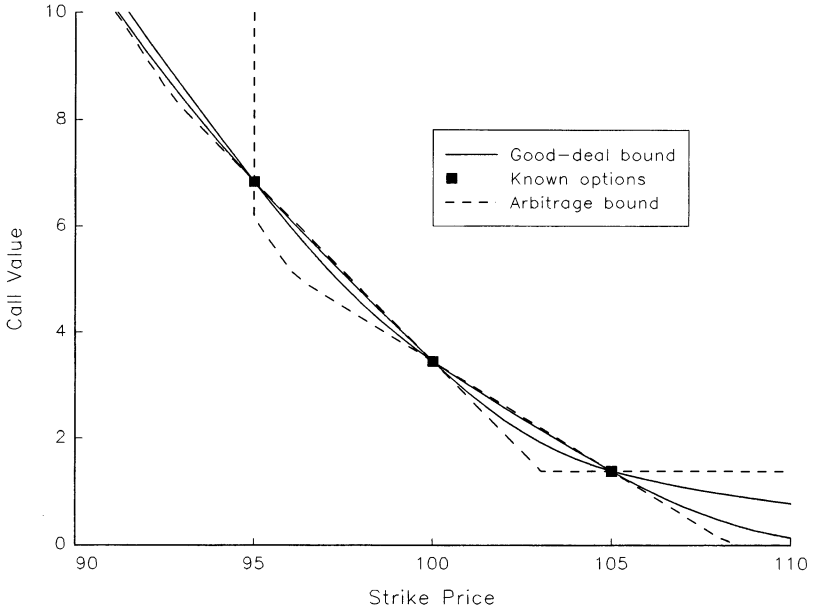

FIG. 3.—Bounds on option values with three months to expiration, no intermediate trading. Options are hedged with stock, interest rate, and three options whose prices and strikes are marked with squares.

other assets may provide better approximate hedges in an incomplete market. In particular, options with different strike prices have payoffs more similar to that of the option at hand, so they may provide sharper information about an option's value than the underlying stock. Therefore, we examine how one can fill in option prices across strikes with no dynamic hedging.

Mechanically, we just include other options with observed prices in the payoff space X along with the underlying stock and a risk-free bond, and we use integrals against the lognormal density to calculate the required moments. We again relegate the lengthy evaluation of the integrals to the algebraic appendix (Cochrane and Saá-Requejo 1998).

Figure 3 shows good-deal bounds as a function of the strike price, using the same three-month horizon and parameters as before. The black squares plot the prices of three options whose prices are observed. The curves give bounds on the value of an additional option. The arbitrage bounds ($m > 0$ constraint) in this case state that the option price must be a concave function of the strike price and must obey the standard call arbitrage bounds.

The good-deal bounds improve on the arbitrage bounds through-

out. Their small size gives some justification to the common practice of drawing a smooth line through observed option prices, but it also warns that small differences in how one draws such a line can have a dramatic effect on the Sharpe ratios of option-based portfolios. The good-deal bounds are much tighter than arbitrage bounds beyond the last trade options, where concavity places fewer restrictions on value. Overall, the bounds are much tighter than those of figure 1, verifying the intuition that other options are a better approximate hedge for a given option than the underlying security is.

C. *Deltas*

Option pricing theory is used extensively to quantify risk exposure by measuring how much an option value would change if an underlying state variable such as the stock price changed. This sensitivity is known as the option's "delta." Here we expect (at best) a range for deltas rather than a number. (Deltas are also used to construct hedges but are less interesting for that purpose in an incomplete markets context. For example, we already know how to construct a hedge that minimizes the variance of residual risk, $\hat{x}^e = \text{proj}(x^e | X)$.)

Suppose that we observe an option price, which we write as $C(S, V, K, T)$, where S is the stock price, V is additional state variables (if any), K is the strike price, and T is the time to expiration. We want to know, how would the option price change if the stock price changed a bit, $C(S + \Delta S, V, K, T)$? Alas, our methods give bounds on the prices of other *securities*, such as $C(S, V, K + \Delta K, T)$, not the same security in a different state of the world.

To infer prices in a different state from the prices of other securities, we assume that the option value is homogeneous of degree one in the stock price and the strike price, S and K:

$$C(\alpha S, V, \alpha K, T) = \alpha C(S, V, K, T).$$

This assumption basically says that the units (dollars or cents) of the underlying price are irrelevant. Merton (1973) shows that this assumption holds when the distribution of returns is independent of the level of the asset price. Homogeneity implies

$$\frac{\partial C}{\partial S} S + \frac{\partial C}{\partial K} K = C.$$

Hence we can evaluate deltas with

$$\frac{\partial C}{\partial S} = \frac{C}{S} - \frac{\partial C}{\partial K} \frac{K}{S}. \tag{36}$$

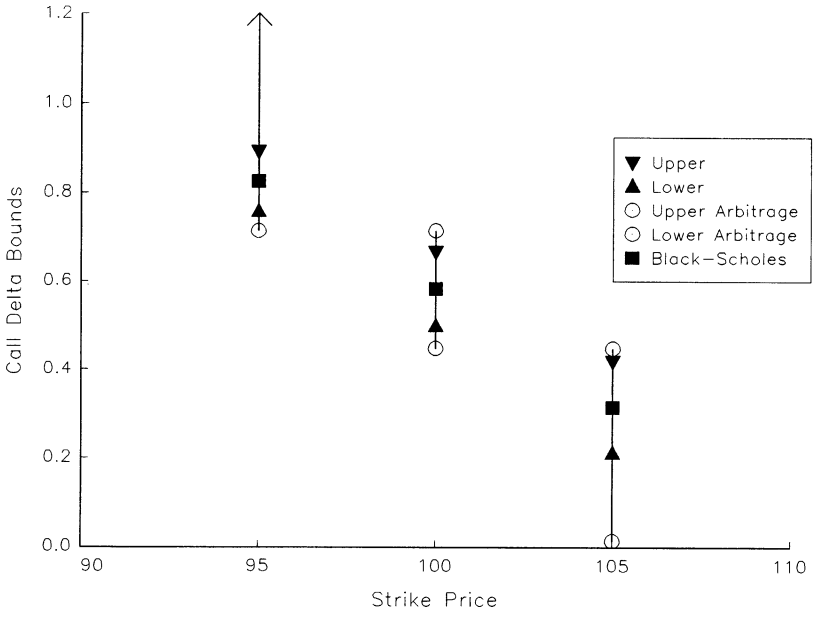

FIG. 4.—Bounds on option price deltas $\partial C/\partial S$ with three months to expiration, no intermediate trading.

To calculate a derivative, we have to start at a point with a known or hypothesized option value. (Otherwise all we know is that the value is in a bound now and in another bound at $t + \Delta t$; the time and hence state derivative can be infinite.) We calculate bounds at the black squares at which option prices are observed. At such a point, the maximum and minimum slopes of the bounds $\partial \underline{C}/\partial K$ and $\partial \overline{C}/\partial K$ determine maximum and minimum deltas via equation (36).

Figure 4 presents a graph of upper and lower delta bounds computed in this way, using the same setup as figure 3. Since the slopes of the good-deal bounds are smaller than those of the arbitrage bounds at all the observed option prices, the good-deal delta bounds are tighter than delta bounds based on arbitrage bounds alone. In this numerical example, the gain is not that large for the central option value since the slopes of the good-deal bounds in figure 3 were not that much less than the slopes of the arbitrage bounds. The gain is quite large for the first and last options, where arbitrage bounds widen.

In sum, the good-deal bounds can be used in this way to quantify risk exposure and to assess the uncertainty in risk exposure calculations that assume market prices for untraded risks.

D. *A Multiperiod, Multinomial Approach to Index Options*

So far we have used the Black-Scholes lognormal environment to understand how our bounds work in a well-understood setup. Here we pursue a more serious application to S&P 500 index option pricing. We calculate bounds on a three-month index call option. We use a weekly trading interval, so we iterate the bound calculation back 12 times from expiration. The lack of continuous hedging is still the source of market incompleteness. We model the stock return as a freely specified multinomial rather than a lognormal. Discrete but frequent trading and multinomial statistical models give a convenient and commonly used environment for the numerical application of option pricing techniques. We again choose a discount factor volatility target equal to twice the market Sharpe ratio.

1. Modeling the Return Distribution

The actual stock index return distribution features fatter tails than the lognormal distribution. Therefore, we use a kernel estimate of the empirical return distribution. We set up a grid of return values R_i at the weekly hedging interval with associated probability values π_i. Since we are not pricing by arbitrage, this distribution need not be binomial or generated by a binomial tree. Instead, we use a large number of grid points to accurately capture the empirical return distribution. We use return data R_t to assign probabilities π_i via a Gaussian kernel,

$$\pi_i = k \times \sum_{t=1}^{T} \exp\left[-\left(\frac{R_i - R_t}{w}\right)^2\right]$$

for each i, where w is a weight chosen visually at 0.01 and k is a constant chosen so that the probabilities sum to one. This procedure assumes that the return data R_t at each date t are drawn from the same unconditional distribution. This assumption is more reasonable for excess or real returns than for nominal returns, so we use for R_t the S&P 500 return less the Treasury bill rate plus the constant 5 percent interest rate that we use in the option pricing calculations. We use the full Center for Research in Security Prices (CRSP) daily S&P 500 data sample, July 1962–December 1996.

Figure 5 presents the resulting distribution of weekly returns. Each triangle represents one grid point of the multinomial distribution. The plus signs plot tail values of the actual return data; their height on the graph is arbitrary. One can see the fat tails of the

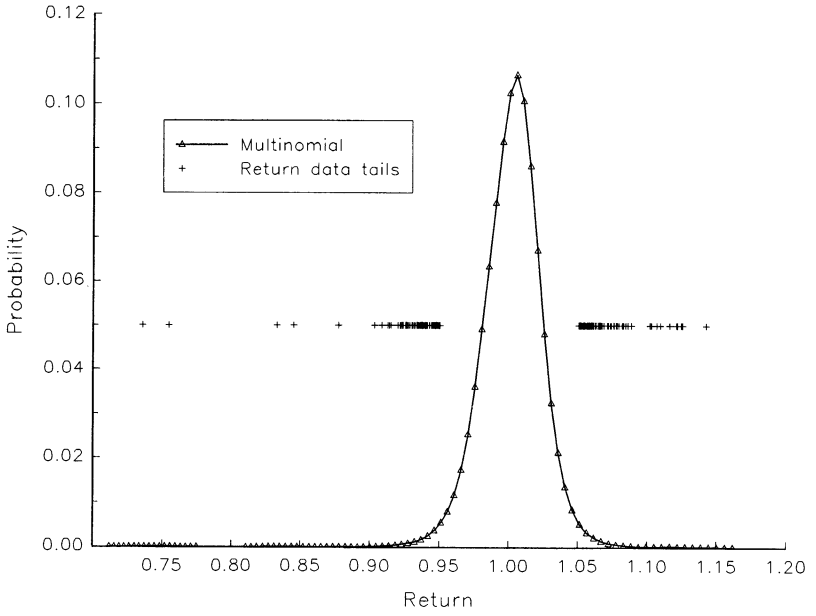

FIG. 5.—Multinomial return distribution fitted from CRSP S&P 500 weekly return data. The plus signs mark return observations larger than ±5 percent, as measured on the x-axis. Their height is arbitrary.

distribution driven by the occasional spectacular outlier returns, especially crashes.

We assume that this multinomial is the conditional return distribution and that it is independent over time. We checked that the long-horizon distributions implied by chaining together empirical distributions as in figure 5 are not bad approximations to the actual long-horizon distributions. Adding conditioning information (stochastic volatility, serial correlation in returns, etc.) is an important extension, which will allow for variation in the bounds over time but will add state variables.

2. Calculating Bounds

We set up a grid of stock prices $\{S_j\}$. This grid may be finer or coarser than the return distribution grid. At each date, we represent the lower (upper) bound function $\underline{C}(S, t)$ by its value on the grid of stock prices. To evaluate moments, we interpolate $\underline{C}(S, t)$ on the stock price grid. For example, suppose that we know $\underline{C}(S, t + \Delta t)$ and we wish to find $\underline{C}(S, t)$. Start at stock price grid index j, time t. For each element of the return grid R_i, we find the corresponding

value of the stock price at time $t + \Delta t$, $S_j R_i$. By linear interpolation, we find the bound at this point, $\underline{C}(S_j R_i, t + \Delta t)$. Now, by summing over states i weighted by the associated probabilities, we can find all the required moments for the time t bound calculation.

The discount factor is the choice variable, and it is now a vector with one element for every point on the return grid. At each stock price grid point, we first compute the price bounds assuming that only the volatility constraint binds. If the resulting discount factor is positive (at every point on the return grid), this is the solution and the positivity constraint is slack. If not, we compute arbitrage bounds,[5] and we compute the minimum second-moment discount factor needed to attain the arbitrage bounds. If that second moment is less than its bound A^2, the arbitrage bounds are the solution and the discount factor volatility constraint is slack. If not, we compute the bounds with both volatility and positivity constraints binding as described above.

3. Results

Figure 6 presents bounds, translated to implied volatilities. The middle line imposes a zero market price, or risk-neutral valuation, of the residual risk by simply setting $v = 0$ in the bound formula (9). The upper bound and zero residual risk price calculations show

[5] Ritchken (1985) gives a solution for the arbitrage bounds in a multinomial environment, the solution to the linear programming problem

$$\min_{\{m\}} E(mx^c) \quad \text{subject to } \mathbf{p} = E(m\mathbf{x}),\ m \geq 0.$$

The upper bound has a nonzero discount factor in the smallest (1) and largest (N) return states:

$$m_1 = \frac{1}{\pi_1 R^f} \frac{R_N - R^f}{R_N - R_1},$$

$$m_N = \frac{1}{\pi_N R^f} \frac{R^f - R_N}{R_N - R_1}.$$

The lower bound discount factor is nonzero only in the two states (i, j) with return just above and below the risk-free rate:

$$m_i = \frac{1}{\pi_i R^f} \frac{R_j - R^f}{R_j - R_i};$$

$$m_j = \frac{1}{\pi_j R^f} \frac{R^f - R_i}{R_j - R_i}.$$

These discount factors are not unique; other discount factors can also generate the arbitrage bounds. Therefore, to test whether this arbitrage bound solution is the solution to the good-deal bound, we use these discount factors to compute the arbitrage bound, and then we search for the minimum variance discount factor that generates the arbitrage bound.

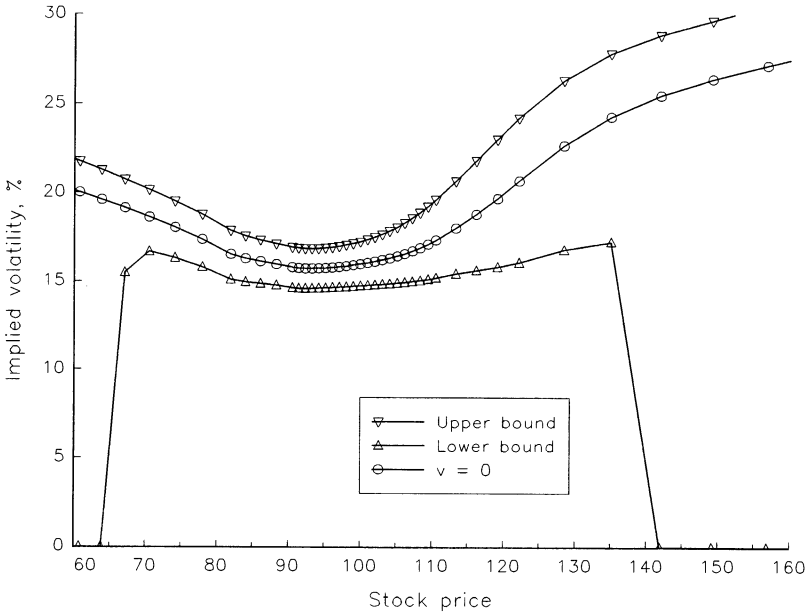

FIG. 6.—Upper and lower good-deal bounds for S&P 500 index options with three months to expiration and strike price $K = 100$. The stock return follows a multinomial distribution fitted to S&P 500 return data. Each option is hedged with the index and a 5 percent risk-free rate. Hedge portfolios may be changed weekly.

a pronounced volatility smile. This smile is driven by the fat tails in the return distribution. However, the lower bound does not show much, if any, smile, and it reduces to the lower arbitrage bound past a certain point. Therefore, a robust volatility smile depends on tighter bounds on market prices of residual risk than we have imposed, or more frequent than weekly rebalancing.

While interpreting these results, keep in mind that implied volatility is a highly nonlinear function, and very small price changes for far-from-the-money options translate to very large changes in implied volatility. For example, at a stock price of \$142, where the lower good-deal bound meets the lower arbitrage bound, option prices are driven by the probability of a $42/142 = 30$ percent decline in stock prices. At this stock price, the Black-Scholes price with 15 percent volatility is only 1.79×10^{-6} dollars above the arbitrage bound! Even the upper good-deal bound is only 1.5¢ above the arbitrage bound. Furthermore, the results in this range are driven by the extreme tail estimates in the probability distribution and so are subject to substantial sampling uncertainty. One more crash in the sample could double these option values.

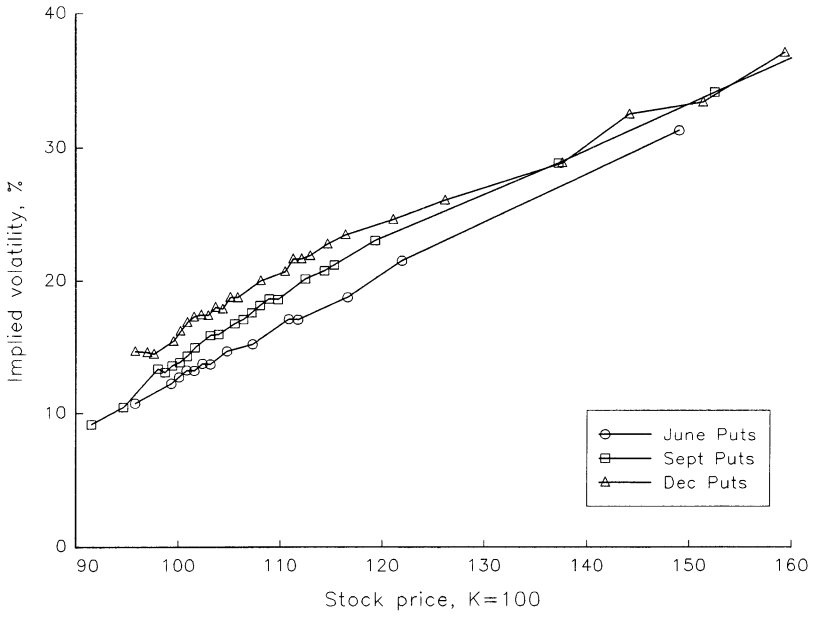

FIG. 7.—Implied volatilities for CBOE S&P 500 index put options with three months to expiration, 1996.

To compare these bounds with actual option prices, figure 7 presents data on daily closing Chicago Board Options Exchange (CBOE) S&P 500 index option prices for three months in 1996.[6]

The data show the well-known volatility smile: implied volatilities (prices) are higher for farther-in-the-money calls and equivalent out-of-the-money puts on the right-hand side of the graph. This rise in volatility is a little steeper than that predicted by our calculation of the bounds, but not by much. (Note that the horizontal scale on fig. 6 extends farther to the left than that of fig. 7, so you see only one side of the predicted smile.) The calculation replicates the implied volatilities of the far-from-the-money options. Since the lower bound is never below 15 percent and the option data go as low as 10 percent, the puzzle in the data is the *low* implied volatility of the near-money options.

There is some variability across strike prices (graphed as variability

[6] We present volatility curves for the end of June, September, and December, when the fullest set of options across strike prices is available. We convert put prices to call equivalent using put-call parity, $C = P + S - (K/R^f)$. We choose the interest rate each month between 1 and 5 percent to best match put and call implied volatility curves. To conform with other figures, we transform (implied) call prices with a common stock price S^* and varying strikes K to call prices with a common strike price $K^* = 100$ and varying stock price $S = S^*K^*/K$ by assuming linear homogeneity, $C(S^*K^*/K, K^*) = (K^*/K) \times C(S^*, K)$.

across stock prices) for any given date, and the variation between puts and calls (not shown) is about the same size. This variability is much smaller than the size of the bounds. One is tempted to conclude that variability across strikes is economically unimportant. To take this conclusion seriously, however, we should consider hedging each option with all the other options but also take seriously the illiquidity and bid/ask spread of all the other options, especially those far from the money.

There is also a certain amount of variability across time: the curves shift up or down by about three to four points of implied volatility. This is also less than the size of the bounds. This observation suggests that an important fraction of observed variation in implied volatility over time may simply be due to variation within good-deal bounds that is difficult to profitably hedge away, *not* due to variation in the conditional variance of stock returns. In this case, the ability to hedge one option with another will not change the picture since the whole curve wanders up or down.

We have of course conducted many more experiments than we have space to show. As the time to expiration increases, the predicted implied volatility curves flatten out, and the region in which the lower bound lies above the arbitrage bound increases. For shorter times to expiration, the predicted volatility smiles steepen, but the bounds become wider. With a monthly hedging interval, the upper bound is about the same but the lower bound is looser. The monthly interval also produces a steeper volatility smile to the right and a shallower one to the left. The monthly return distribution has larger crashes and smaller positive outliers. This fact suggests that one may better match empirical volatility smiles if one calibrates the weekly return distribution to more accurately reflect the implied three-month distribution.

E. Basis Risk and Real Options

Suppose that we want to value a European call option on an event V that is not a traded asset, but is correlated with a traded asset that can be used as an approximate hedge. There is an unavoidable "basis risk" between the option and the securities that can be used to hedge the option, even if we can trade continuously. This situation is common with real options, with nonfinancial options (options on airplane purchases), and in the application of option pricing techniques to real investments. For example, V could represent the price of a nondurable good, and one wants to value the right to build an investment project that can produce one unit of the good at cost K. Even purely financial options are increasingly written on events that are not traded assets, such as catastrophe insurance options. For

options on traded but illiquid underlying assets, it is often better to hedge an option with an imperfectly correlated but more liquid asset.

The terminal payoff is

$$x_T^c = \max(V_T - K, 0).$$

We model the joint evolution of the traded asset S and the event V on which the option is written as

$$\frac{dS}{S} = \mu_S dt + \sigma_S dz,$$

$$\frac{dV}{V} = \mu_V dt + \sigma_{Vz} dz + \sigma_{Vw} dw,$$

$$E(dz^2) = E(dw^2) = 1, \quad E(dzdw) = 0.$$

There is also a constant risk-free rate r.

Since the setup is so close to the Black-Scholes setup, we can give closed-form solutions to the good-deal bounds in this case:

$$\underline{C}_t \text{ or } \overline{C} = V_0 e^{\eta T} \phi(d + \tfrac{1}{2}\sigma_V \sqrt{T}) - K e^{-rT} \phi(d - \tfrac{1}{2}\sigma_V \sqrt{T}), \quad (37)$$

where $\phi(\cdot)$ denotes the left tail of the normal distribution and

$$\sigma_V^2 \equiv E_t \frac{dV^2}{V^2} = \sigma_{Vz}^2 + \sigma_{Vw}^2,$$

$$d \equiv \frac{\ln(V_0/K) + (\eta + r)T}{\sigma_V \sqrt{T}},$$

$$\eta \equiv \left[h_V - h_S \left(\rho - a \sqrt{\frac{A^2}{h_S^2} - 1} \sqrt{1 - \rho^2} \right) \right] \sigma_V,$$

$$h_S \equiv \frac{\mu_S - r}{\sigma_S}; \quad h_V \equiv \frac{\mu_V - r}{\sigma_V},$$

$$\rho \equiv \text{corr}\left(\frac{dV}{V}, \frac{dS}{S} \right) = \frac{\sigma_{Vz}}{\sigma_V},$$

$$a = \begin{cases} +1 \text{ upper bound} \\ -1 \text{ lower bound.} \end{cases}$$

This expression is exactly the Black-Scholes formula with the addition of the η term. The Black-Scholes formula appears as a special case, $\eta = 0$ when $V = S$. The term μ_V enters the formula because the event V may not grow at the same rate as the asset S. Obviously, the correlation ρ between V shocks and asset shocks enters the formula, and as this correlation declines, the bounds widen. The bounds also widen as the volatility constraint A becomes larger relative to the asset Sharpe ratios h_S.

We derive equation (37) by applying the discount factor integration technique described in proposition 5. When we apply that proposition, the lower bound is generated by the discount factor

$$\frac{d\underline{\Lambda}}{\underline{\Lambda}} = -rdt - h_S dz - \sqrt{A^2 - h_S^2}\, dw,$$

and the bound is then given by

$$\underline{C}_t = E_t\left[\frac{\underline{\Lambda}_T}{\underline{\Lambda}_t}\max(V_T - K)\right].$$

The upper bound is generated with a plus sign in front of the square root. The terms S_T, V_T, and Λ_T are jointly lognormally distributed, so the double integral defining the expectation is straightforward to perform. Once again, we relegate the algebra to the algebraic appendix (Cochrane and Saá-Requejo 1998).

As a numerical example, we once again consider a call option with three months to expiration and strike value 100. To keep the example as simple as possible, we specify that V has the same drift and volatility as the stock index ($\mu_V = \mu_S$, $\sigma_V = \sigma_S$). We specify the correlation between dV and the stock index $\rho = 0.9$, and we again use a discount factor volatility bound equivalent to twice the market Sharpe ratio. Figure 8 presents the upper and lower price bounds.

As before, the dollar size of the bounds is smaller for out-of-the-money options, but the proportional size of the bounds shown in the figure is smaller for in-the-money options. The bounds are substantial and wider above the Black-Scholes value than below it[7] with 10–30 percent errors. The errors increase as time to expiration in-

[7] We do not report implied volatilities because the highly nonlinear implied volatility metric is not appropriate in this case. For example, the lower price bound can fall below the lower arbitrage bound, with no violation of arbitrage, since the correlation is not perfect. We cannot even calculate implied volatilities of such prices.

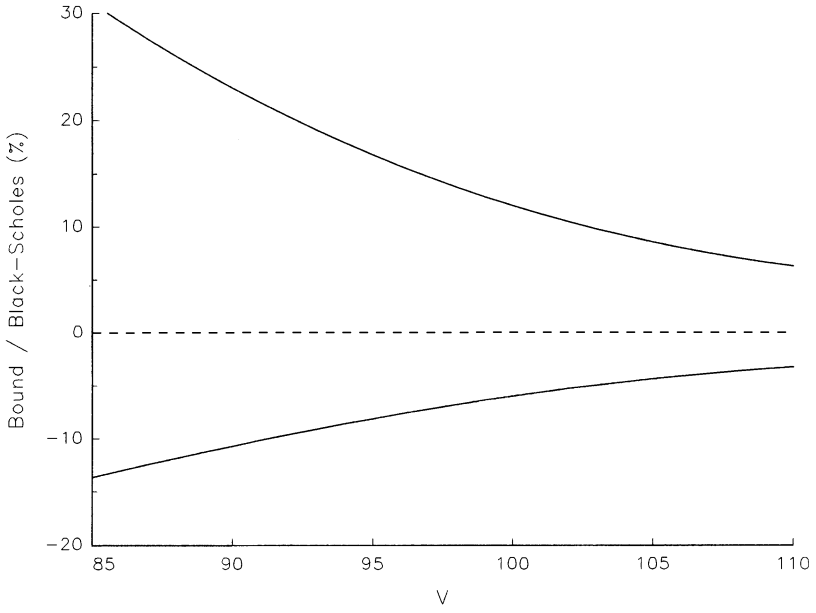

FIG. 8.—Good-deal bounds for a European call option written on an event V whose shocks are correlated .9 with the stock index shocks. Continuous trading is allowed. There are three months to expiration, and other parameters match the index.

creases. The figure helps us to evaluate the common perception that Black-Scholes is a good approximation even when its assumptions are violated, as in this case.

IV. Extensions

A. Picking the Basis Assets and Volatility Constraint

The size of the bounds is directly related to the size of the residual, or the R^2 in a regression of option payoffs on basis asset payoffs. It is therefore tempting to use lots of basis assets, but fishing for R^2 is dangerous since it is hard to build a believable and stable model of the joint distribution of a large number of assets. This is not just an issue for econometricians: Large hedge fund losses in the summer of 1998 were partly attributed to too-low estimates of the correlation between yield spreads on different bond classes (Henriques and Kahn 1998). Many basis assets can be more easily accommodated when a theoretical structure governs their joint distribution, as was the case when we used other options as basis assets.

The size of our bounds is also directly related to the difference between the maximum discount factor volatility and the volatility used up in pricing the basis assets, $A^2 - E(x^{*2})$. This fact suggests

that one include high–Sharpe ratio basis assets to obtain tighter bounds, even if those assets are not well correlated with the focus payoff. However, it may be better to lower the volatility constraint A^2 instead. This procedure gives up whatever use the extra assets had as hedges, but it avoids the construction of a large joint distribution model.

For example, when pricing bond options, one could use a volatility constraint equal to the market Sharpe ratio and exclude the market from the set of basis assets rather than use a constraint equal to twice the market Sharpe ratio and build a model of the joint distribution of stocks and bonds. Similarly, we used a constraint of twice the market Sharpe ratio and ignored the possibility of high–Sharpe ratio, low-β strategies such as Fama and French's (1993) value portfolio, rather than include them, model them, and use a perhaps more realistic constraint (given their presence) of four times the market Sharpe ratio.

B. Transactions Costs

Specifying frictionless trading at discrete intervals is a crude way to handle transactions costs. Our one-period calculations could include explicit transactions costs, using Luttmer's (1996) solution for Hansen-Jagannathan bounds with transactions costs. But this does not solve the central problem, namely, to find the optimal times at which to trade in a continuous-time framework with transactions costs. For this reason, most of the literature on option pricing bounds with explicit transactions costs also considers two-period analysis or ad hoc trading rules, for example, Leland (1985), Constantinides and Zariphopoulou (1997), and Constantinides (1998).

C. More Discount Factor Restrictions

Adding further discount factor restrictions is a natural way to tighten the bounds.

Levy (1985) and Constantinides (1998) calculate option price bounds by assuming that the discount factor declines monotonically with a state variable. We have a related condition: the covariance between stock returns and the discount factor must be negative to generate positive expected returns (the pricing constraint). But global monotonicity is more stringent than covariance, and this constraint may tighten the bounds.

Our bounds allow the worst case that marginal utility growth is perfectly correlated with a portfolio of basis and focus assets. In many cases one could credibly impose a sharper limit than $-1 \leq \rho \leq 1$ on this correlation to obtain tighter bounds.

Bernardo and Ledoit (2000, this issue) use the restriction $a \geq m \geq b$ to sharpen the no-arbitrage restriction $\infty \geq m > 0$ and cleverly relate this bound taken alone to portfolio "gain-loss ratios." This restriction is related to the volatility restriction: upper and lower bounds on m imply a bound on its variance, and by Chebyshev's inequality, a variance bound implies that the probability that m exceeds upper and lower bounds can be made vanishingly small. Our central proposition 2 that one can solve multiperiod problems recursively goes through in this case.

Bernardo and Ledoit also suggest that one impose $a \geq m/y \geq b$, where y is an explicit discount factor model such as the consumption-based model or CAPM, as a way of imposing a "weak implication" of that particular model. Limits on $\sigma(m/y)$ or $\sigma(m - y)$ might be technically easier to impose, especially in continuous time, but this idea can also be added.

Until one has recovered a true discount factor $m_{t+1} = \beta u'(c_{t+1}) \div u'(c_t)$, further restrictions will always tighten bounds. Also, and despite our extended motivation, one should not be too dogmatic about which weak restriction is best. By their nature, weak restrictions will admit values that are strange according to stronger or different restrictions. Therefore, the various restrictions are not competitors. One should impose as many discount factor restrictions that apply to a given situation, and, perhaps most important, with which one can still conveniently calculate bounds for a given application.

D. Multiple Focus Assets

We have emphasized a single focus asset. For some applications, such as examining the set of index options across the strike price, it may make more sense to find a bound on the values $(p_1, p_2, p_3, \ldots)$ of several focus payoffs simultaneously rather than consider each focus payoff in isolation.

Appendix

Proofs of Lemmas and Propositions

Proof of Lemma 1

If: $E(w\mathbf{x}) = E(\epsilon\mathbf{x}) = 0$, so $E[(x^* + vw + \epsilon)\mathbf{x}] = E(x^*\mathbf{x}) = \mathbf{p}$.

Only if: For any m that satisfies $\mathbf{p} = E(m\mathbf{x})$, we have $\text{proj}(m|\mathbf{x}) = E(m\mathbf{x}')E(\mathbf{x}\mathbf{x}')^{-1}\mathbf{x}$ by the OLS projection formula, and then $\text{proj}(m|\mathbf{x}) = \mathbf{p}'E(\mathbf{x}\mathbf{x}')^{-1}\mathbf{x} = x^*$ by the assumption that $\mathbf{p} = E(m\mathbf{x})$. Define a residual $\delta = m - x^*$. By construction, $E(\delta\mathbf{x}) = 0$. Thus any discount factor m can

be represented as $m = x^* + \delta$. Projecting δ on w and defining ϵ as the residual again, we obtain $m = x^* + vw + \epsilon$.

Proof of Proposition 1

We substitute for m from equation (6) to express the volatility constraint as

$$E(m^2) = E(x^{*2}) + v^2E(w^2) + E(\epsilon^2) \leq A^2$$

and to express the objective as

$$E[mx^c] = E(x^*x^c) + vE(wx^c) + E(\epsilon x^c) = E(x^*x^c) + vE(w^2). \quad \text{(A1)}$$

The problem now reduces to the choice of v and ϵ. Since ϵ appears only in the inequality constraint, we choose $\epsilon = 0$ to weaken that constraint as much as possible. Now we have a linear objective subject to a quadratic constraint and one parameter v to choose. As long as $w \neq 0$ so $E(w^2) \neq 0$, the volatility constraint binds and we simply pick v to satisfy the constraint, leading to equation (10). Equations (8) and (9) follow by substituting $\underline{v}$ into equation (6) and into the objective, equation (A1).

The Sharpe ratio bound must be greater than or equal to that generated by the basis assets, $A^2 \geq E(x^{*2}) = \mathbf{p}'E(\mathbf{x}\mathbf{x}')^{-1}\mathbf{p}$, of course, or there is no discount factor that even satisfies the constraints. If $w = 0$, then the volatility constraint does not bind. Any discount factor m that satisfies the pricing constraint produces the same price for x^c, so $\underline{C} = \overline{C} = E(x^*x^c)$.

Proof of Proposition 2

Denote by Ω_1 and $\Omega_2(I_1)$ the constraint sets on m_1, m_2 defined by (18) and (19), respectively. Using the law of iterated expectations, we can write

$$\underline{C}_0 = \min_{\{m_1 \in \Omega_1\}} \min_{\{m_2 \in \Omega_2(I_1)\}} E_0[m_1E_1(m_2x_2^c)].$$

Suppose that the solution to this problem gives $E_1(m_2x_2^c) > \underline{C}_1$ in some state of nature at time 1. If this were true, one could lower $E_1(m_2x_2^c)$ in that state without violating any of the constraints, by definition of $\underline{C}_1$. Since $m_1 > 0$, this action would lower $E_0[m_1E_1(m_2x_2^c)]$, contrary to the assumption that we started at a solution to the two-period problem.

Proof of Lemma 2

We verify that Λ satisfies the basic pricing equation (22) for the risk-free rate and basis assets. Liberally using $E(dSdw) = 0$ and $E(dzdw) = 0$, we get

$$E_t\left(\frac{d\Lambda}{\Lambda}\right) = -rdt,$$

$$E_t\left(\frac{d\Lambda S}{\Lambda S}\right) + \frac{D}{S}dt = E_t\left(\frac{dS}{S}\right) + E_t\left(\frac{d\Lambda^*}{\Lambda^*}\right) + E_t\left(\frac{d\Lambda^* dS}{\Lambda^* S}\right) + \frac{D}{S}dt$$

$$= \left(\mu_S - r - \tilde{\mu}_S'\Sigma_S^{-1}\sigma_S\sigma_S' + \frac{D}{S}\right)dt = 0.$$

Only if holds by projection, as in lemma 1.

Proof of Proposition 3

If we substitute equation (28) into the problem (27) in order to impose the pricing constraint, the problem is

$$0 = \frac{x^c}{\underline{C}}dt + E_t\left[\frac{d(\Lambda^*\underline{C})}{\Lambda^*\underline{C}}\right] - \min_{\{v\}} v\,E_t\left(dw\frac{d\underline{C}}{\underline{C}}\right)$$

$$\text{subject to } vv' \le A^2 - \frac{1}{dt}E_t\left(\frac{d\Lambda^{*2}}{\Lambda^{*2}}\right).$$

When we use equation (30) for $d\underline{C}/\underline{C}$ in the last term, the problem is

$$0 = \frac{x^c}{\underline{C}} + \frac{1}{dt}E_t\left[\frac{d(\Lambda^*\underline{C})}{\Lambda^*\underline{C}}\right] - \min_{\{v\}} v\sigma'_{\underline{C}w}$$

$$\text{subject to } vv' \le A^2 - \frac{1}{dt}E_t\left(\frac{d\Lambda^{*2}}{\Lambda^{*2}}\right). \tag{A2}$$

This is a linear objective in v with a quadratic constraint. Therefore, as long as $\sigma_{\underline{C}w} \neq 0$, the constraint binds and the optimal v is given by (33). The equality $\bar{v} = -\underline{v}$ gives the maximum since $\sigma_{\underline{C}w}\sigma'_{\underline{C}w} > 0$. Plugging the optimal value for v in (A2) gives

$$0 = \frac{x^c}{\underline{C}} + \frac{1}{dt}E_t\left[\frac{d(\Lambda^*\underline{C})}{\Lambda^*\underline{C}}\right] - v\sigma'_{\underline{C}w}.$$

For clarity, and exploiting the fact that $d\Lambda^*$ does not load on dw, we write the middle term as

$$\frac{1}{dt}E_t\left[\frac{d(\Lambda^*\underline{C})}{\Lambda^*\underline{C}}\right] = \mu_{\underline{C}} - r + \frac{1}{dt}E_t\left(\frac{d\Lambda^*}{\Lambda}\sigma_{\underline{C}z}dz\right).$$

If $\sigma_{\underline{C}w} = 0$, any v leads to the same price bound. In this case we can most simply take $v = 0$.

References

Abel, Andrew B., and Eberly, Janice C. "Optimal Investment with Costly Reversibility." *Rev. Econ. Studies* 63 (October 1996): 581–93.

Bernardo, Antonio E., and Ledoit, Olivier. "Gain, Loss, and Asset Pricing." *J.P.E.* 108 (February 2000): 144–72.

Black, Fischer, and Scholes, Myron S. "The Pricing of Options and Corporate Liabilities." *J.P.E.* 81 (May/June 1973): 637–54.

Cochrane, John H., and Saá-Requejo, Jesús. "Algebraic Appendix to 'Beyond Arbitrage.'" Manuscript. Chicago: Univ. Chicago, 1998. http://www-gsb.uchicago.edu/fac/john.cochrane/.

———. "Good-Deal Option Price Bounds with Stochastic Volatility and Stochastic Interest Rate." Manuscript. Chicago: Univ. Chicago, 1999.

Constantinides, George M. "Transactions Costs and the Volatility Implied

by Option Prices." Manuscript. Chicago: Univ. Chicago, Grad. School Bus., 1998.

Constantinides, George M., and Zariphopoulou, Thaleia. "Bounds on Option Prices in an Intertemporal Setting with Proportional Transaction Costs and Multiple Securities." Manuscript. Chicago: Univ. Chicago, Grad. School Bus., 1997.

Dixit, Avinash K., and Pindyck, Robert S. *Investment under Uncertainty*. Princeton, N.J.: Princeton Univ. Press, 1994.

Fama, Eugene F., and French, Kenneth R. "Common Risk Factors in the Returns on Stocks and Bonds." *J. Financial Econ.* 33 (February 1993): 3–56.

Hansen, Lars Peter; Heaton, John; and Luttmer, Erzo G. J. "Econometric Evaluation of Asset Pricing Models." *Rev. Financial Studies* 8 (Summer 1995): 237–74.

Hansen, Lars Peter, and Jagannathan, Ravi. "Implications of Security Market Data for Models of Dynamic Economies." *J.P.E.* 99 (April 1991): 225–62.

Hansen, Lars Peter, and Richard, Scott F. "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models." *Econometrica* 55 (May 1987): 587–613.

He, Hua, and Pearson, Neil D. "Consumption and Portfolio Policies with Incomplete Markets and Short-Sale Constraints: The Infinite Dimensional Case." *J. Econ. Theory* 54 (August 1991): 259–304.

Henriques, Diana B., and Kahn, Joseph. "Lessons of a Long, Hot Summer." *New York Times* (December 6, 1998).

Kandel, Shmuel, and Stambaugh, Robert F. "Portfolio Inefficiency and the Cross-Section of Expected Returns." *J. Finance* 50 (March 1995): 157–84.

Ledoit, Olivier. "Essays on Risk and Return in the Stock Market." Ph.D. dissertation, Massachusetts Inst. Tech., 1995.

Leland, Hayne E. "Option Pricing and Replication with Transactions Costs." *J. Finance* 40 (December 1985): 1283–1301.

Levy, Haim. "Upper and Lower Bounds of Put and Call Option Value: Stochastic Dominance Approach." *J. Finance* 40 (September 1985): 1197–1217.

Luttmer, Erzo G. J. "Asset Pricing in Economies with Frictions." *Econometrica* 64 (November 1996): 1439–67.

MacKinlay, A. Craig. "Multifactor Models Do Not Explain Deviations from the CAPM." *J. Financial Econ.* 38 (May 1995): 3–28.

Merton, Robert C. "Theory of Rational Option Pricing." *Bell J. Econ. and Management Sci.* 4 (Spring 1973): 141–83.

Ritchken, Peter H. "On Option Pricing Bounds." *J. Finance* 40 (September 1985): 1219–33.

Ross, Stephen A. "The Arbitrage Theory of Capital Asset Pricing." *J. Econ. Theory* 13 (December 1976): 341–60. (*a*)

———. "Risk, Return and Arbitrage." In *Risk and Return in Finance*, vol. 1, edited by Irwin Friend and James L. Bicksler. Cambridge, Mass.: Ballinger, 1976. (*b*)

Shanken, Jay. "The Current State of the Arbitrage Pricing Theory." *J. Finance* 47 (September 1992): 1569–74.