J.R. Birge and V. Linetsky (Eds.), *Handbooks in OR & MS, Vol. 15*

DOI: 10.1016/S0927-0507(07)15013-1

Chapter 13

Option Pricing: Real and Risk-Neutral Distributions

George M. Constantinides
University of Chicago and NBER
E-mail: gmc@ChicagoGSB.edu

Jens Carsten Jackwerth
University of Konstanz
E-mail: Jens.Jackwerth@uni-konstanz.de

Stylianos Perrakis
Concordia University
E-mail: SPerrakis@jmsb.concordia.ca

Abstract

The central premise of the Black and Scholes [Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–659] and Merton [Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4, 141–184] option pricing theory is that there exists a self-financing *dynamic trading policy* of the stock and risk free accounts that renders the market *dynamically complete*. This requires that the market be *complete* and *perfect*. In this essay, we are concerned with cases in which dynamic trading breaks down either because the market is incomplete or because it is imperfect due to the presence of trading costs, or both. Market incompleteness renders the risk-neutral probability measure non unique and allows us to determine the option price only within a range. Recognition of trading costs requires a refinement in the definition and usage of the concept of a risk-neutral probability measure. Under these market conditions, a replicating dynamic trading policy does not exist. Nevertheless, we are able to impose restrictions on the pricing kernel and derive testable restrictions on the prices of options. We illustrate the theory in a series of market setups, beginning with the single period model, the two-period model and, finally, the general multiperiod model, with or without transaction costs. We also review related empirical results that document widespread violations of these restrictions.

Keywords: Derivative pricing; Risk-neutral distribution; Incomplete markets; Stochastic dominance bounds; Transaction costs; Index options; Volatility smile

1 Introduction

The Nobel-winning ingenious idea behind the classic option pricing model of Black and Scholes (1973) and Merton (1973), hereafter BSM, is that, in the absence of arbitrage, the price of an option equals the cost of setting up a judiciously managed portfolio with payoff that replicates the option payoff.

The central premise of the BSM theory is that there exists a self-financing *dynamic trading policy* of the stock and risk free accounts that renders the market *dynamically complete*. This requires that the market be *complete* and *perfect*. Two assumptions of the BSM model make the market complete. First, the price of the underlying security has continuous sample paths at the exclusion of jumps. Second, the stock return volatility is constant. These assumptions essentially imply that the price of the underlying security is a geometric Brownian motion. Finally, the assumption of the BSM model that renders the market perfect is that trading is frictionless. In the BSM model, the volume of trading over any finite time interval is infinite. The transaction costs associated with the replicating dynamic trading policy would be infinite for any given positive proportional transactions cost rate.

Formally, absence of arbitrage in a frictionless market implies the existence of a *risk-neutral probability measure*, not necessarily unique, such that the price of any asset equals the expectation of its payoff under the risk-neutral measure, discounted at the risk free rate. Furthermore, if the market is complete then the risk-neutral measure is unique and the option price is unique as well. In the BSM model, the price of the underlying security follows a geometric Brownian motion which renders the market complete and the option price unique as well.

The risk-neutral probability measure is the real probability measure with the expected rate of return on the underlying security replaced by the risk free rate. The real probability distribution of stock returns can be estimated from the time series of past returns. The risk-neutral probability distribution of stock returns can be estimated from the cross section of option prices. As discussed in detail in the empirical Section 10, this prediction of the BSM theory does not fare well and provides the motivation to reexamine the premises of the theory.

In this essay, we are concerned with cases in which dynamic trading breaks down either because the market is incomplete or because there are trading costs or both. Market incompleteness renders the risk-neutral probability measure non unique and allows us to determine the option price only within a range. Recognition of trading costs requires a refinement in the definition and usage of the concept of a risk-neutral probability measure.

In Section 2, we discuss the implications of the absence of arbitrage. We introduce the concept of the risk-neutral probability and the closely related concept of the *state price density* or *pricing kernel*. We apply the theory to price options under the assumption of the absence of arbitrage in complete and incomplete markets. In Section 3, we lay out the general framework for pricing options in a market that is incomplete and also imperfect due to trading costs. Under these market conditions, a replicating dynamic trading policy does not exist. Nevertheless, we are able to impose further restrictions on the pricing kernel and provide testable restrictions on the prices of options. In Sections 4–9, we illustrate the theory in a series of market setups, beginning with the single period model, the two-period model and finally the general multiperiod model, with or without transaction costs. In Section 10, we review related empirical results and, in Section 11, conclude.

2 Implications of the absence of arbitrage

2.1 General theory

Absence of arbitrage in a frictionless market implies the existence of a risk-neutral probability measure, not necessarily unique, such that the price of any asset equals the expectation of its payoff under the risk-neutral measure, discounted at the risk free rate. If a risk-neutral measure exists, the ratio of the risk-neutral probability density and the real probability density, discounted at the risk free rate, is referred to as the *pricing kernel* or *stochastic discount factor* (SDF). Thus, absence of arbitrage implies the existence of a strictly positive SDF. These ideas are implicit in the option pricing theory of Black and Scholes (1973) and Merton (1973) and are further developed by Ross (1976), Cox and Ross (1976), Constantinides (1978), Harrison and Kreps (1979), Harrison and Pliska (1981), and Delbaen and Schachermayer (1994).

To fix ideas, let there be J securities. Security j, $j = 1, \ldots, J$, has price P_j at the beginning of the period and payoff X_{ij} in state i, $i = 1, \ldots, I$, at the end of the period. An investor purchases θ_j securities of type j, $j = 1, \ldots, J$, with the objective to minimize the purchase cost, subject to the constraint that the portfolio payoff is strictly positive in all states of nature. The investor solves the following LP problem:

$$\inf_{\{\theta_j\}} \sum_{j=1}^{J} \theta_j P_j \tag{2.1}$$

subject to

$$\sum_{j=1}^{J} \theta_j X_{ij} > 0, \quad \forall i. \tag{2.2}$$

If the minimum purchase cost is negative, then there is an arbitrage opportunity.

Absence of arbitrage implies that the above problem, with the added condition

$$\sum_{j=1}^{J} \theta_j P_j < 0 \tag{2.3}$$

is infeasible. Then the *dual* of this LP problem is feasible. This implies the existence of strictly positive *state prices*, $\{\pi_i\}_{i=1,\dots,I}$, such that:

$$P_j = \sum_{i=1}^{I} \pi_i X_{ij}, \quad \forall j \tag{2.4}$$

and

$$\pi_i > 0, \quad \forall i. \tag{2.5}$$

If the number of states does not exceed the number of securities with linearly independent payoffs, the market is said to be *complete* and the state prices are unique. Otherwise, the market is *incomplete* and the state prices are not unique.

The normalized state prices $q_i \equiv \pi_i / \sum_{k=1}^{I} \pi_k$ can be thought of as probabilities because they are strictly positive and add up to one. The inverse of the sum of the state prices, $R \equiv 1/\sum_{k=1}^{I} \pi_k$, has the interpretation as one plus the risk free rate. Then we may write Eq. (2.4) as

$$P_j = R^{-1} \sum_{i=1}^{I} q_i X_{ij} = R^{-1} E^{Q}[X_j], \quad \forall j \tag{2.6}$$

with the interpretation that the price of security j is its expected payoff under the probability measure $Q = \{q_i\}$, discounted at the risk free rate. For this reason, the probability measure Q is referred to as a *risk-neutral* or *risk-adjusted* probability measure. Thus, absence of arbitrage implies the existence of a risk-neutral probability measure. This property of the absence of arbitrage is far more general than this simple illustration implies.

Let $P = \{p_i\}$ denote the real probability measure of the states. The ratio $m_i \equiv \pi_i / p_i$ is referred to as the *state price density* or *stochastic discount factor* or *pricing kernel* or *intertemporal marginal rate of substitution*. In terms of the pricing kernel, we may write Eq. (2.4) as

$$P_j = \sum_{i=1}^{I} p_i m_i X_{ij} = E^{P}[m_i X_j], \quad \forall j \tag{2.7}$$

where the expectation is with respect to the real probability measure P.

2.2 Application to the pricing of options

Let the stock market index have price S_0 at the beginning of the period; *ex dividend* price S_i with probability p_i in state i, $i = 1, \ldots, I$, at the end of the period; and *cum dividend* price $(1 + \delta)S_i$ at the end of the period. The jth derivative, $j = 1, \ldots, J$, has price P_j at the beginning period, and its cash payoff X_{ij} is $G_j(S_i)$, a given function of the terminal stock price, at the end of the period in state i. In this context, absence of arbitrage implies the existence of a strictly positive pricing kernel $m : m_i$, $i = 1, \ldots, I$, such that:

$$1 = R \sum_{i=1}^{I} p_i m_i, \tag{2.8}$$

$$S_0 = \sum_{i=1}^{I} p_i m_i (1 + \delta) S_i \tag{2.9}$$

and

$$P_j = \sum_{i=1}^{I} p_i m_i G_j(S_i), \quad j = 1, \ldots, J. \tag{2.10}$$

Non-existence of a strictly positive pricing kernel implies arbitrage such as violations of the Merton (1973) no-arbitrage restrictions on the prices of options.

In practice, it is always possible to estimate the real probability measure P from time series data on past index returns. A *derivatives pricing model* is then a theory that associates the appropriate pricing kernel $m : m_i > 0$, $i = 1, \ldots, I$, with the estimated probability measure P.

In the absence of arbitrage, a *unique* pricing kernel may be derived in terms of the prices of J securities with linearly independent payoffs, if the market is complete, $J \geqslant I$. Then any derivative is uniquely priced in terms of the prices of I securities. This is the essence of derivatives pricing when the market is complete. An example of a complete market is the binomial model, described next.

In a *single-period binomial* model, there are just two states and the pricing kernel is derived in terms of the prices of the risk free asset and the stock or index on which options are written. Then any derivative is uniquely priced in terms of the risk free rate and the stock or index price. The natural extension of the single period binomial model is the widely used *multiperiod binomial* model developed by Cox and Ross (1976), Cox et al. (1979), and Rendleman and Bartter (1979). The stock price evolves on a multi-stage binomial tree over the life of the option so that the stock price assumes a wide range of values. Yet the market is complete because in each subperiod there are only two states. An option can be hedged or replicated on the binomial tree by adjusting the amounts held in the stock and the risk free asset at each stage of the binomial process. This type of trading is called *dynamic trading* and renders the market *dynamically complete*. These fundamental ideas underlie the original option pricing model of Black and Scholes (1973) and Merton (1973). The binomial model is often used as a pedagogical tool to illustrate these ideas as in the textbook treatments by Hull (2006) and McDonald (2005). The binomial model is also a powerful tool in its own right in numerically pricing American and exotic options.

In this essay, we are concerned with cases in which dynamic trading or hedging breaks down either because the market is incomplete or because there are trading costs or both. In these cases, we impose further restrictions on the pricing kernel by taking into account the economic environment in which the derivatives are traded.

3 Additional restrictions implied by utility maximization

3.1 Multiperiod investment behavior with proportional transaction costs

We consider a market with heterogeneous agents and investigate the restrictions on option prices imposed by a particular class of utility-maximizing traders that we simply refer to as *traders*. We do not make the restrictive assumption that all agents belong to the class of the utility-maximizing traders. Thus our results are unaffected by the presence of agents with beliefs, endowments, preferences, trading restrictions, and transaction cost schedules that differ from those of the utility-maximizing traders.

As in Constantinides (1979), trading occurs at a finite number of trading dates, $t = 0, 1, \ldots, T, \ldots, T'$.[1] The utility-maximizing traders are allowed to hold only two primary securities in the market, a bond and a stock. The stock has the natural interpretation as the market index. Derivatives are introduced in the next section. The bond is risk free and pays constant interest $R - 1$ each period. The traders may buy and sell the bond without incurring transaction costs. At date t, the *cum dividend* stock price is $(1 + \delta_t)S_t$, the cash dividend is $\delta_t S_t$, and the *ex dividend* stock price is S_t, where δ_t is the dividend yield. We assume that the rate of return on the stock, $(1+\delta_{t+1})S_{t+1}/S_t$, is identically and independently distributed over time.

The assumption of i.i.d. returns is not innocuous and, in particular, rules out state variables such as stochastic volatility, stochastic risk aversion, and stochastic conditional mean of the growth rate in dividends and consumption. In this essay, we deliberately rule out such state variables in order to explore the extent to which market incompleteness and market imperfections (trading costs) alone explain the prices of index options. We discuss models with such state variables in Section 10.

[1] The calendar length of the trading horizon is N years and the calendar length between trading dates is N/T' years. Later on we vary T' and consider the mispricing of options under different assumptions regarding the calendar length between trading dates.

Stock trades incur proportional transaction costs charged to the bond account as follows. At each date t, the trader pays $(1+k)S_t$ out of the bond account to purchase one *ex dividend* share of stock and is credited $(1-k)S_t$ in the bond account to sell (or, sell short) one *ex dividend* share of stock. We assume that the transactions cost rate satisfies the restriction $0 \leqslant k < 1$. Note that there is no presumption that all agents in the economy face the same schedule of transaction costs as the traders do.

A trader enters the market at date t with dollar holdings x_t in the bond account and y_t/S_t *ex dividend* shares of stock. The endowments are stated net of any dividend payable on the stock at time t.[2] The trader increases (or, decreases) the dollar holdings in the stock account from y_t to $y_t' = y_t + v_t$ by decreasing (or, increasing) the bond account from x_t to $x_t' = x_t - v_t - k|v_t|$. The decision variable v_t is constrained to be measurable with respect to the information at date t. The bond account dynamics are

$$x_{t+1} = \{x_t - v_t - k|v_t|\}R + (y_t + v_t)\frac{\delta_t S_{t+1}}{S_t}, \quad t \leqslant T' - 1 \tag{3.1}$$

and the stock account dynamics are

$$y_{t+1} = (y_t + v_t)\frac{S_{t+1}}{S_t}, \quad t \leqslant T' - 1. \tag{3.2}$$

At the terminal date, the stock account is liquidated, $v_{T'} = -y_{T'}$, and the net worth is $x_{T'} + y_{T'} - k|y_{T'}|$. At each date t, the trader chooses investment v_t to maximize the expected utility of net worth, $E[u(x_{T'} + y_{T'} - k|y_{T'}|)|S_t]$.[3] We make the plausible assumption that the utility function, $u(\cdot)$, is increasing and concave, and is defined for both positive and negative terminal net worth.[4] Note that even this weak assumption of monotonicity and concavity of preferences is not imposed on all agents in the economy but only on the subset of agents that we refer to as traders.

We recursively define the value function $V(t) \equiv V(x_t, y_t, t)$ as

$$V(x_t, y_t, t) = \max_v E\bigg[V\bigg(\{x_t - v - k|v|\}R$$

[2] We elaborate on the precise sequence of events. The trader enters the market at date t with dollar holdings $x_t - \delta_t y_t$ in the bond account and y_t/S_t *cum dividend* shares of stock. Then the stock pays cash dividend $\delta_t y_t$ and the dollar holdings in the bond account become x_t. Thus, the trader has dollar holdings x_t in the bond account and y_t/S_t *ex dividend* shares of stock.

[3] The results extend routinely to the case that consumption occurs at each trading date and utility is defined over consumption at each of the trading dates and over the net worth at the terminal date. See Constantinides (1979) for details. The model with utility defined over terminal net worth alone is a more realistic representation of the objective function of financial institutions.

[4] If utility is defined only for non-negative net worth, then the decision variable is constrained to be a member of a convex set that ensures the non-negativity of net worth. See Constantinides (1979) for details. However, the derivation of bounds on the prices of derivatives requires an entirely different approach and yields weaker bounds. This problem is studied in Constantinides and Zariphopoulou (1999, 2001).

$$+ (y_t + v)\frac{\delta_t S_{t+1}}{S_t}, (y_t + v)\frac{S_{t+1}}{S_t}, t+1\bigg) \bigg| S_t\bigg] \tag{3.3}$$

for $t \leqslant T' - 1$, and

$$V(x_{T'}, y_{T'}, T') = u\big(x_{T'} + y_{T'} - k|y_{T'}|\big). \tag{3.4}$$

We assume that the parameters satisfy appropriate technical conditions such that the value function exists and is once differentiable.

Equations (3.1)–(3.4) define a dynamic program that can be numerically solved for given utility function and stock return distribution. We shall not solve this dynamic program because our goal is to derive restrictions on the prices of options that are independent of the specific functional form of the utility function but solely depend on the plausible assumption that the traders' utility function is *monotone increasing* and *concave* in the terminal wealth.

The value function is increasing and concave in (x_t, y_t), properties that it inherits from the assumed monotonicity and concavity of the utility function, as shown in Constantinides (1979):

$$V_x(t) > 0, \qquad V_y(t) > 0, \quad t = 0, \ldots, T, \ldots, T' \tag{3.5}$$

and

$$\begin{aligned} &V\big(\alpha x_t + (1-\alpha)x'_t, \alpha y_t + (1-\alpha)y'_t, t\big) \\ &\quad \geqslant \alpha V(x_t, y_t, t) + (1-\alpha)V(x'_t, y'_t, t), \\ &\quad 0 < \alpha < 1,\ t = 0, \ldots, T, \ldots, T'. \end{aligned} \tag{3.6}$$

On each date, the trader may transfer funds between the bond and stock accounts and incur transaction costs. Therefore, the marginal rate of substitution between the bond and stock accounts differs from unity by, at most, the transaction costs rate:

$$(1-k)V_x(t) \leqslant V_y(t) \leqslant (1+k)V_x(t), \quad t = 0, \ldots, T, \ldots, T'. \tag{3.7}$$

Marginal analysis on the bond holdings leads to the following condition on the marginal rate of substitution between the bond holdings at dates t and $t+1$:

$$V_x(t) = RE_t\big[V_x(t+1)\big], \quad t = 0, \ldots, T, \ldots, T' - 1. \tag{3.8}$$

Finally, marginal analysis on the stock holdings leads to the following condition on the marginal rate of substitution between the stock holdings at date t and the bond and stock holdings at date $t+1$:

$$\begin{aligned} &V_y(t) = E_t\bigg[\frac{S_{t+1}}{S_t}V_y(t+1) + \frac{\delta_t S_{t+1}}{S_t}V_x(t+1)\bigg], \\ &\quad t = 0, \ldots, T, \ldots, T' - 1. \end{aligned} \tag{3.9}$$

Below we employ these restrictions on the value function to derive restrictions on the prices of options.

3.2 Application to the pricing of options

We consider J European-style derivatives on the index, with random cash payoff $G_j(S_T)$, $j = 1, 2, \ldots, J$, at their common expiration date T, $T \leqslant T'$. At time zero, the trader can buy the jth derivative at price $P_j + k_j$ and sell it at price $P_j - k_j$, net of transaction costs. Thus $2k_j$ is the bid–ask spread plus the round-trip transaction costs that the trader incurs in trading the jth derivative. Note that there is no presumption that all agents in the economy face the same bid–ask spreads and transaction costs as the traders do.

We assume that the traders are marginal in all J derivatives. Furthermore, we assume that, if a trader holds a finite (positive or negative) number of derivatives, these positions are sufficiently small relative to her holdings in the bond and stock that the monotonicity and concavity conditions (3.5) and (3.6) on the value function remain valid.[5]

Marginal analysis leads to the following restrictions on the prices of options:

$$(P_j - k_j)V_x(0) \leqslant E_0\big[G_j(S_T)V_x(T)\big] \leqslant (P_j + k_j)V_x(0), \quad j = 1, 2, \ldots, J. \tag{3.10}$$

Similar restrictions apply to the prices of options at dates $t = 1, \ldots, T-1$.

Below, we illustrate the implementation of the restrictions on the prices of options in a number of important special cases. First, we consider the case $T = 1$ which rules out trading between the bond and stock accounts over the lifetime of the options. We refer to this case as the *single-period case*. Note that the single-period case does not rule out trading over the trader's horizon after the options expire; it just rules out trading over the lifetime of the options. We discuss the single-period case both with and without transaction costs.

A useful way to identify the options that cause infeasibility or near-infeasibility of the problem is to single out a "test" option, say the Jth option, and solve the problem

$$\min_{\{V_x(t),V_y(t)\}_{t=0,\ldots,T}} E_0\left[G_J(S_T)\frac{V_x(T)}{V_x(0)}\right], \tag{3.11}$$

subject to conditions (3.5)–(3.10), where in Eq. (3.10) the subscript j runs from 1 to $J-1$. If this problem is feasible, then the attained minimum has the following interpretation. If one can buy the test option for less than the minimum attained in this problem, then at least one investor, *but not necessarily all investors*, increases her expected utility by trading the test option.

Likewise, we may solve the problem

$$\max_{\{V_x(t),V_y(t)\}_{t=0,\ldots,T}} E_0\left[G_J(S_T)\frac{V_x(T)}{V_x(0)}\right], \tag{3.12}$$

[5] Conditions (3.7)–(3.9) remain valid even if the holdings of the derivatives are not small.

subject to conditions (3.5)–(3.10), where in Eq. (3.10) the subscript j runs from 1 to $J-1$. If this problem is feasible, then the attained maximum has the following interpretation. If one can write the test option for more than the maximum attained in this problem, then at least one investor, *but not necessarily all investors*, increases her expected utility by trading the test option.

As the number of trading dates T increases, the computational burden rapidly increases. One way to reduce computational complexity is to limit attention to the case $J = 1$ (one option) and convex payoff (as, for example, the payoff of a call or put option). In this special case, we present closed-form solutions with and without transaction costs and, in many cases, present limiting forms of the option prices, as the number of intermediate trading dates becomes infinitely large.

4 Special case: one period without transaction costs

4.1 Results for general payoffs

The stock market index has price S_0 at the beginning of the period; *ex dividend* price S_i with probability p_i in state i, $i = 1, \dots, I$, at the end of the period; *cum dividend* price $(1+\delta)S_i$ at the end of the period; and return $(1+\delta)S_i/S_0$. We define by $z_i \equiv S_i/S_0$ the *ex dividend* price ratio. We order the states such that S_i is increasing in i. The jth derivative, $j = 1, \dots, J$, has price P_j at the beginning period and cash payoff $G_j(z_i)$ at the end of the period in state i. We denote by $V^i(t)$ the value function at date t and state i.

Since the transaction costs rate is assumed to be zero, we have $V_x(0) = V_y(0)$ and $V_x^i(1) = V_y^i(1)$. We identify the previously defined stochastic discount factor or pricing kernel m_i with the intertemporal marginal rate of substitution in state i, $m_i \equiv V_x^i(1)/V_x(0)$. Conditions (3.8)–(3.10) become:

$$1 = R\sum_{i=1}^{I} p_i m_i, \tag{4.1}$$

$$1 = \sum_{i=1}^{I} p_i m_i (1+\delta) z_i \tag{4.2}$$

and

$$P_j = \sum_{i=1}^{I} p_i m_i G_j(z_i), \quad j = 1, \dots, J. \tag{4.3}$$

The concavity relation (3.6) of the value function implies additional restrictions on the pricing kernel. Historically, the expected premium of the return on the stock over the bond is positive. Under the assumption of positive expected premium, the trader is long in the stock. Since the assumption in the single-period model is that there is no trading between the bond and stock accounts over the life of the option, the trader's wealth at the end of the period is increasing in the stock return. Note that this conclusion critically depends on the assumption that there is no intermediate trading in the bond and stock. Since we employed the convention that the stock return is increasing in the state i, the trader's wealth on date T is increasing in the state i. Then the concavity of the value function implies that the marginal rate of substitution is decreasing in the state i:

$$m_1 \geqslant m_2 \geqslant \cdots \geqslant m_I > 0. \tag{4.4}$$

A pricing kernel satisfying restrictions (4.1)–(4.4) defines the intertemporal marginal rate of substitution of a trader who maximizes her increasing and concave utility and is marginal in the options, the index and the risk free rate. If there does not exist a pricing kernel satisfying restrictions (4.1)–(4.4), then any trader with increasing and concave utility can increase her expected utility by trading in the options, the index, and the risk free rate – hence equilibrium does not exist. These strategies are termed *stochastically dominant* for the purposes of this essay, insofar as they would be adopted by all traders with utility possessing the required properties, in the same way that all risk averse investors would choose a dominant portfolio over a dominated one in conventional second degree stochastic dominance comparisons. Thus, the existence of a pricing kernel that satisfies restrictions (4.1)–(4.4) is said to rule out *stochastic dominance* between the observed prices.

We emphasize that the restriction on option prices imposed by the criterion of the absence of stochastic dominance is motivated by the economically plausible assumption that there exists at least *one* agent in the economy with the properties that we assign to a trader. This is a substantially weaker assumption than requiring that *all* agents have the properties that we assign to traders. Stochastic dominance then implies that at least one agent, *but not necessarily all agents*, increases her expected utility by trading.[6]

As before, we single out a "test" option, say the Jth option, and derive bounds that signify infeasibility if the price of the test option lies outside the bounds. The general form of this problem was stated in expressions (3.11) and (3.12). In the special case of no trading over the life of the option and zero transactions costs, the bounds on the test option with payoff $G_J(z_i)$ in state i are given by

$$\max_{\{m_i\}}\Bigl(\text{or}, \min_{\{m_i\}}\Bigr) \sum_{i=1}^{I} p_i m_i G_J(z_i), \tag{4.5}$$

[6] We also emphasize that the restriction of the absence of stochastic dominance is weaker than the restriction that the capital asset pricing model (CAPM) holds. The CAPM requires that the pricing kernel be linearly decreasing in the index price. The absence of stochastic dominance merely imposes that the pricing kernel be monotone decreasing in the index price.

subject to conditions (4.1)–(4.4), where in Eq. (4.3) the subscript j runs from 1 to $J-1$.

4.2 Results for convex payoffs

The feasibility of relations (4.1)–(4.4) can be expressed in closed form in the special case where the options are puts and calls, with payoff $G_j(z_i)$ that is a convex function of the end-of-period return (or stock price). Ryan (2000, 2003) provided inequalities that define an admissible range of prices for each option by considering the prices of the two options with immediately adjacent strike prices and Huang (2005) tightened these inequalities. In practice, this means that (4.1)–(4.4) become infeasible in most realistic problems with a large enough set of traded options.

Perrakis and Ryan (1984), Levy (1985), and Ritchken (1985) expressed the upper and lower bounds in (4.5) in closed form in the special case $J=1$ (one option) where the option is a put or call, with payoff $G_1(z_i)$ that is a convex function of the end-of-period stock price. Consider a European call option with strike price K, payoff $G_1(z_i)=[S_0 z_i(1+\delta)-K]^+ \equiv c_i$ and price $P_1=c$. Define $\hat{z} \equiv \sum_{i=1}^{I} p_i z_i$ and assume $(1+\delta)\hat{z} \geqslant R$. Equations (4.1)–(4.5) become

$$\max_{\{m_i\}}(\text{or}, \min) \sum_{i=1}^{I} p_i m_i c_i \tag{4.6}$$

subject to

$$\begin{aligned} &\sum_{i=1}^{I} p_i m_i (1+\delta) z_i = 1, \quad \text{and} \\ &R\sum_{i=1}^{I} p_i m_i = 1, \quad m_1 \geqslant \cdots \geqslant m_I > 0. \end{aligned} \tag{4.7}$$

The solution to (4.6)–(4.7) crucially depends on the minimum value $z_{\min} \equiv z_1$. If $z_{\min} > 0$, the upper and lower bounds $\bar{c}_0$ and $\underline{c}_0$ on the call option price are given by

$$\begin{aligned} \bar{c}_0 &= \frac{1}{R}\left[\frac{R-(1+\delta)z_{\min}}{(1+\delta)(\hat{z}-z_{\min})}\hat{c}_I + \frac{(1+\delta)\hat{z}-R}{(1+\delta)(\hat{z}-z_{\min})}c_1\right], \\ \underline{c}_0 &= \frac{1}{R}\left[\frac{R-(1+\delta)\hat{z}_h}{(1+\delta)(\hat{z}_{h+1}-\hat{z}_h)}\hat{c}_{h+1} + \frac{(1+\delta)\hat{z}_{h+1}-R}{(1+\delta)(\hat{z}_{h+1}-\hat{z}_h)}\hat{c}_h\right]. \end{aligned} \tag{4.8}$$

In the above equations, h is a state index such that $(1+\delta)\hat{z}_h \leqslant R \leqslant (1+\delta)\hat{z}_{h+1}$ and we have used the following notation for conditional expectations for $k = 1, \ldots, I$:

$$\hat{c}_k = \frac{\sum_{i=1}^{k} c_i p_i}{\sum_{i=1}^{k} p_i} = E\big[c_T \mid S_T \leqslant S_0(1+\delta)z_k\big],$$

$$\hat{z}_k = \frac{\sum_{i=1}^{k} z_i p_i}{\sum_{i=1}^{k} p_i} = E\left[z_T \mid z_T \leqslant z_k\right]. \tag{4.9}$$

Inspection of Eqs. (4.8) and (4.9) reveals that both the upper and lower bounds of the call option are discounted expectations with two different distributions, $U = \{u_i\}$ and $L = \{l_i\}$. These distributions are both *risk neutral*, since it can be easily verified that $R^{-1} \sum_{i=1}^{i=I} u_i(1+\delta)z_i = R^{-1} \sum_{i=1}^{i=I} l_i(1+\delta)z_i = 1$. These distributions are:

$$\begin{aligned}
u_1 &= \frac{R-(1+\delta)z_{\min}}{(1+\delta)(\hat{z}-z_{\min})} p_1 + \frac{(1+\delta)\hat{z}-R}{(1+\delta)(\hat{z}-z_{\min})},\\
u_i &= \frac{R-(1+\delta)z_{\min}}{(1+\delta)(\hat{z}-z_{\min})} p_i, \quad i=2,\dots,I,\\
l_i &= \frac{(1+\delta)\hat{z}_{h+1}-R}{(1+\delta)(\hat{z}_{h+1}-\hat{z}_h)} \frac{p_i}{\sum_{k=1}^{h} p_k} + \frac{R-(1+\delta)\hat{z}_h}{(1+\delta)(\hat{z}_{h+1}-\hat{z}_h)} \frac{p_i}{\sum_{k=1}^{h+1} p_k},\\
&\quad i=1,\dots,h,\\
l_{h+1} &= \frac{R-(1+\delta)\hat{z}_h}{(1+\delta)(\hat{z}_{h+1}-\hat{z}_h)} \frac{p_{h+1}}{\sum_{k=1}^{h+1} p_k}.
\end{aligned} \tag{4.10}$$

As the states increase, the distribution of z becomes continuous over the interval $[z_{\min}, \infty)$, with actual distribution $P(z)$ and expectation $E(z)$. Then, U and L become

$$\begin{aligned}
U(z) &= \begin{cases} P(z) & \text{with probability } \frac{R-(1+\delta)z_{\min}}{(1+\delta)(E(z)-z_{\min})},\\ 1_{z_{\min}} & \text{with probability } \frac{(1+\delta)E(z)-R}{(1+\delta)(E(z)-z_{\min})}. \end{cases}\\
L(z) &= P\big(z \mid (1+\delta)E(z) \leqslant R\big).
\end{aligned} \tag{4.11}$$

We note that the two call option bounds become two increasing and *convex* functions $\bar{c}(S_0)$ and $\underline{c}(S_0)$ given by

$$\begin{aligned}
\bar{c}(S_0) &= \frac{1}{R} E^U\left[\left(S_0(1+\delta)z - K\right)^+\right],\\
\underline{c}(S_0) &= \frac{1}{R} E^L\left[\left(S_0(1+\delta)z - K\right)^+\right].
\end{aligned} \tag{4.12}$$

In the important special case $z_{\min} = 0$, the upper bound in (4.12) becomes

$$\bar{c}(S_0) = \frac{1}{(1+\delta)E[z]} E^P\left[\left(S_0(1+\delta)z - K\right)^+\right]. \tag{4.13}$$

Similar results are available for put options. We have thus shown that under the *no intermediate trading* assumption the option price is bound by two values given as the expectation of discounted payoff under two limiting distributions. Oancea and Perrakis (2006) provided corresponding bounds when $(1+\delta)\hat{z} \leqslant R$.

5 Special case: one period with transaction costs and general payoffs

In a one-period model with transaction costs and general payoffs, conditions (3.8)–(3.10) become

$$V_x(0) = R\sum_{i=1}^{I} p_i V_x^i(1), \tag{5.1}$$

$$V_y(0) = \sum_{i=1}^{I} p_i\left[\frac{S_i}{S_0}V_y^i(1) + \frac{\delta S_i}{S_0}V_x^i(1)\right] \tag{5.2}$$

and

$$(P_j - k_j)V_x(0) \leqslant \sum_{i=1}^{I} p_i G_j(S_i)V_x^i(1) \leqslant (P_j + k_j)V_x(0),$$
$$j = 1, \ldots, J. \tag{5.3}$$

Conditions (3.5)–(3.7) become[7]

$$V_x(0) > 0, V_y(0) > 0, V_x^i(1) > 0, V_y^i(1) > 0, \quad i = 1, \ldots, I, \tag{5.4}$$

$$V_y^1(1) \geqslant V_y^2(1) \geqslant \cdots \geqslant V_y^I(1) > 0 \tag{5.5}$$

and

$$(1-k)V_x^i(1) \leqslant V_y^i(1) \leqslant (1+k)V_x^i(1), \quad i = 1, \ldots, I. \tag{5.6}$$

As before, a useful way to pinpoint options that cause infeasibility or near-infeasibility of the problem is to single out a "test" option and solve the problems (3.11) and (3.12) subject to restrictions (5.1)–(5.6).

In order to highlight the difference in the formulation brought about by transaction costs, we adopt a notation similar to that in (4.1)–(4.5). We define $m_i \equiv V_x^i(1)/V_x(0)$, the marginal rate of substitution between the *bond* account at time one and the bond account at time zero and state i; and $\lambda_i \equiv V_y^i(1)/V_x(0)$, the marginal rate of substitution between the *stock* account at time one and the bond account at time zero and state i. Then (5.1)–(5.6) become

$$1 = R\sum_{i=1}^{I} p_i m_i, \tag{5.7}$$

[7] Since the value of the bond account at the end of the period is independent of the state i, the concavity conditions $V_{xx}(t) < 0$ and $V_{xx}(1)V_{yy}(1) - (V_{xy}(1))^2 > 0$ cannot be imposed. Only the concavity condition $V_{yy}(t) < 0$ is imposed as in Eq. (5.5).

$$(1-k) \leqslant \sum_{i=1}^{I} p_i z_i(\lambda_i + \delta m_i) \leqslant (1+k), \tag{5.8}$$

$$(P_j - k_j) \leqslant \sum_{i=1}^{I} p_i m_i G_j(z_i) \leqslant (P_j + k_j), \quad j = 1, \ldots, J, \tag{5.9}$$

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_I > 0 \tag{5.10}$$

and

$$(1-k)m_i \leqslant \lambda_i \leqslant (1+k)m_i, \quad i = 1, \ldots, I. \tag{5.11}$$

The bounds on the nth option with payoff $G_n(z_i)$ in state i are given by

$$\max_{m_i,\lambda_i}\Big(\text{or, } \min_{m_i,\lambda_i}\Big) \sum_{i=1}^{I} p_i m_i G_n(z_i). \tag{5.12}$$

Transaction costs double the number of variables that must be determined by the solution of the program. Furthermore, transaction costs expand the feasible region of the pricing kernel for any given set of option prices. Indeed, it is easy to see that for $k = 0$, $k_j = 0$, $j = 1, \ldots, J$ the problem (5.7)–(5.12) becomes identical to (4.1)–(4.5). Therefore, if a feasible solution to (4.1)–(4.4) exists, then this solution is feasible for (5.7)–(5.11) with $m_i = \lambda_i$, $i = 1, \ldots, I$. This implies that the spread between the two objective functions of (4.5) lies within the spread of the objective functions of (5.12).

6 Special case: two periods without transaction costs and general payoffs

The single-period model without transaction costs implies that the wealth at the end of the period is an increasing function of the stock price at the end of the period and, therefore, the pricing kernel is a decreasing function of the stock price at the end of the period. Likewise, the single period model with transaction costs implies that the value of the stock account at the end of the period is an increasing function of the stock price at the end of the period and, therefore, the marginal utility of wealth out of the stock account is a decreasing function of the stock price at the end of the period.

Constantinides and Zariphopoulou (1999) pointed out that intermediate trading invalidates the above implications with or without transaction costs, because the wealth at the end of the period (or, the value of the stock account at the end of the period) becomes a function not only of the stock price at the option's expiration but also of the entire sample path of the stock price.[8]

[8] In the special case of i.i.d. returns, power utility, and zero transaction costs, the wealth at the end of the period is a function only of the stock price. However, this assumption would considerably diminish the generality of the model.

Constantinides and Perrakis (2002) recognized that it is possible to recursively apply the single-period approach with or without transaction costs and derive stochastic dominance bounds on option prices in a market with intermediate trading over the life of the options.

In this section, we study a two-period model without transaction costs and, in the next section, a two-period model with transaction costs. In the absence of transaction costs, the value function $V(t) \equiv V(x_t, y_t, t)$ defined in (3.1)–(3.4) becomes a function of the aggregate trader wealth, $V(x_t + y_t, t)$. Therefore, we have $V_x(t) = V_y(t)$, $t = 0, 1, 2$. As before, we define the first period pricing kernel as $m_{1i} \equiv V_x^i(1)/V_x(0)$. For the second period, we define the pricing kernel as $m_{2ik} \equiv V_y^{ik}(2)/V_x(0)$, $i, k = 1, \dots, I$. Then conditions (3.5)–(3.11) become

$$1 = R\sum_{i=1}^{I} p_i m_{1i}, \qquad 1 = R\sum_{k=1}^{I} p_k \frac{m_{2ik}}{m_{1i}}, \quad i = 1, \dots, I, \tag{6.1}$$

$$1 = \sum_{i=1}^{I} p_i m_{1i}(1+\delta)z_i, \qquad 1 = \sum_{k=1}^{I} p_k \frac{m_{2ik}}{m_{1i}}(1+\delta)z_k, \quad i = 1, \dots, I, \tag{6.2}$$

$$P_j = \sum_{i=1}^{I}\sum_{k=1}^{I} p_i p_k m_{2ik} G_j(z_i z_k), \quad j = 1, \dots, J \tag{6.3}$$

and

$$m_{11} \geqslant m_{12} \geqslant \cdots \geqslant m_{1I} > 0, \qquad m_{2i1} \geqslant m_{2i2} \geqslant \cdots \geqslant m_{2iI} > 0, \quad i = 1, \dots, I. \tag{6.4}$$

We test for feasibility by solving the program

$$\max_{m_{1i}, m_{2ik}} \left(\text{or}, \min_{m_{1i}, m_{2ik}}\right) \sum_{i=1}^{I}\sum_{k=1}^{I} p_i p_k m_{2ik} G_n(z_{1i} z_{2k}). \tag{6.5}$$

The extension of the program (6.1)–(6.5) to more than two periods becomes potentially explosive. In Section 8, we present closed form expressions for the bounds on the prices of European options in the special case where the payoff $G_j(S_T)$ is convex (call or put) and $J = 1$, by using the expressions developed in Section 4.2.

7 Special case: two periods with transaction costs and general payoffs

We now allow for transaction costs in the two-period model with general payoffs. Unlike Section 6, we have $V_x(t) \neq V_y(t)$, $t = 0, 1, 2$. We define the first period marginal rates of substitution as $m_{1i} \equiv V_x^i(1)/V_x(0)$ and

$\lambda_{1i} \equiv V_y^i(1)/V_x(0)$, $i = 1, \ldots, I$. We define the two-period marginal rates of substitution as $m_{2ik} \equiv V_x^{ik}(2)/V_x(0)$ and $\lambda_{2ik} \equiv V_y^{ik}(2)/V_x(0)$, $i, k = 1, \ldots, I$. Then conditions (3.5)–(3.11) become

$$1 = R\sum_{i=1}^{I} p_i m_{1i}, \qquad 1 = R\sum_{k=1}^{I} p_k \frac{m_{2ik}}{m_{1i}}, \quad i = 1, \ldots, I, \tag{7.1}$$

$$(1-k) \leqslant \sum_{i=1}^{I} p_i z_{1i}(\lambda_{1i} + \delta m_{1i}) \leqslant (1+k),$$

$$\lambda_{1i} = \sum_{k=1}^{I} p_k z_{2k}(\lambda_{2ik} + \delta m_{2ik}), \quad i = 1, \ldots, I, \tag{7.2}$$

$$P_j - k_j \leqslant \sum_{i=1}^{I}\sum_{k=1}^{I} p_i p_k m_{2ik} G_j(z_i z_k) \leqslant P_j + k_j, \quad j = 1, \ldots, J, \tag{7.3}$$

$$\lambda_{11} \geqslant \lambda_{12} \geqslant \cdots \geqslant \lambda_{1I} > 0, \qquad \lambda_{2i1} \geqslant \lambda_{2i2} \geqslant \cdots \geqslant \lambda_{2iI} > 0, \quad i = 1, \ldots, I \tag{7.4}$$

and

$$(1-k)m_{1i} \leqslant \lambda_{1i} \leqslant (1+k)m_{1i},$$
$$(1-k)m_{2ik} \leqslant \lambda_{2ik} \leqslant (1+k)m_{2ik}, \quad i = 1, \ldots, I, \ k = 1, \ldots, I. \tag{7.5}$$

As before, we test for feasibility by solving the program

$$\max_{m_{1i},\lambda_{1i},m_{2ik},\lambda_{2ik}} \Big(\text{or}, \min_{m_{1i},\lambda_{1i},m_{2ik},\lambda_{2ik}}\Big) \sum_{i=1}^{I}\sum_{k=1}^{I} p_i p_k m_{2ik} G_n(z_{1i} z_{2k}) \tag{7.6}$$

subject to (7.1)–(7.5). Constantinides et al. (2007) tested for violations of the stochastic dominance conditions (7.1)–(7.6).

In Section 9, we present closed form expressions for the bounds on the prices of European options for $T \geqslant 2$ in the special case where the payoff $G_j(S_T)$ is convex (call or put) and $J = 1$, by using the expressions developed in Section 4.2.

8 Multiple periods without transaction costs and with convex payoffs

For the case $J = 1$ and with convex payoffs, it is possible to use the special structure of the closed-form solution (4.8)–(4.12), in order to decompose the general problem into a series of one-period problems for any value of T. Indeed, consider the U and L distributions defined in (4.10) or (4.11) and define the following recursive functions:

$$\begin{aligned}
\bar{c}_t(S_t) &= \frac{1}{R} E^U\big[\bar{c}_{t+1}\big(S_t(1+\delta)z_{t+1}\big) \mid S_t\big],\\
\underline{c}_t(S_t) &= \frac{1}{R} E^L\big[\underline{c}_{t+1}\big(S_t(1+\delta)z_{t+1}\big) \mid S_t\big],\\
\bar{c}_T(S_T) &= \underline{c}_T(S_T) = \big(S_{T-1}z_T(1+\delta) - K\big)^+.
\end{aligned} \tag{8.1}$$

In (8.1), the P, U and L distributions of the successive price ratios $z_{t+1} \equiv S_{t+1}/S_t$ are allowed to depend on the current index value S_t, *provided* such dependence preserves the convexity of the option value $c_t(S_t)$ at any time t with respect to S_t.

Assume that z_{t+1} takes I ordered values $z_{t+1,i}$, $i = 1, \ldots, I$ that determine the states at time $t+1$, set $c_{t+1,i} \equiv c_t(S_t(1+\delta)z_{t+1,i})$ and define at time t the variables $m_{t+1} : m_{t+1,i} \equiv V_y^i(t+1)/V_x(t)$, $i = 1, \ldots, I$. We can then show by induction that the expressions (8.1) define upper and lower bounds on the option value $c_t(S_t)$ at any time $t < T$.[9] We clearly have[10]

$$c_t(S_t) = \sum_{i=1}^{i=I} p_{t+1,i} m_{t+1,i} c_{t+1,i} = E^P\big[m_{t+1} c_t\big(S_t(1+\delta)z_{t+1}\big) \mid S_t\big]. \tag{8.2}$$

With these definitions consider now the program

$$\begin{aligned}
&\min(\text{or}, \max)_{\{m_{t+1,i}\}} c_t = \sum_{i=1}^{I} c_{t+1,i} p_{t+1,i} m_{t+1,i},\\
&\text{subject to:}\quad 1 = \sum_{i=1}^{I} (1+\delta) z_{t+1,i} p_{t+1,i} m_{t+1,i},\\
&\qquad\qquad\qquad 1 = R \sum_{i=1}^{I} p_{t+1,j} m_{t+1,i},\\
&\qquad\qquad\qquad m_{t+1,1} \geqslant m_{t+1,2} \geqslant \cdots \geqslant m_{t+1,I} > 0.
\end{aligned} \tag{8.3}$$

Given the assumed convexity of $c_{t+1} = c_t(S_t(1+\delta)z_{t+1})$, the solution of (8.3) produces upper and lower bounds on $c_t(S_t)$ that are discounted expectations of $c_t(S_t(1+\delta)z_{t+1})$ under the U and L distributions given by (4.10) or (4.11),

[9] The multiperiod upper bound in (8.1) was initially developed in Perrakis (1986). The lower bound was derived in Ritchken and Kuo (1988).

[10] In (8.2) the expectations are conditional on the stock price at time t. In fact the model is more general and the P-distribution may be allowed to depend on other variables such as, for instance, the current volatility of the stock price *provided* convexity is preserved and these other variables do not affect independently the trader's utility function.

conditional on S_t. The bounds on c_t are still given by the recursive expressions in (8.1).

Oancea and Perrakis (2006) addressed the asymptotic behavior of the multiperiod bounds in (8.1) as the number of trading dates increases. They considered specific cases of convergence of the P distribution to a particular stochastic process at the limit of continuous time. They showed that both the U and L distributions defined in (4.11) converge to a *single* risk-neutral stochastic process whenever the P distribution converges to a generalized diffusion, possibly a two-dimensional one, that preserves convexity of the option with respect to the underlying asset price.[11] A necessary and sufficient condition for the convergence of a discrete process to a diffusion is the *Lindeberg condition*, which was used by Merton (1982) to develop criteria for the convergence of binomial and, more generally, multinomial discrete time processes. This condition is applicable to multidimensional diffusion processes.

With minor reformulation, Oancea and Perrakis (2006) extended the validity of the bounds to stochastic volatility and GARCH models of the stock price. They also demonstrated that U and L converge to distinct limits when the limit of the P distribution is a mixed jump-diffusion process. They applied the stochastic dominance bounds to a discrete time process that converges to a mixed jump-diffusion process, in which the logarithm of the jump size amplitude G converges to a distribution with support $G \in [G_{\min}, G_{\max}]$, with $G_{\min} < 0 < G_{\max}$. The fact that the two option bounds converge to two different values is not particularly surprising. Recall that the bounds derived in earlier studies are also dependent either on the special assumption of fully diversifiable jump risk as in Merton (1976), or on the risk aversion parameter of the power utility function of the representative investor, as in Bates (1991) and Amin (1993). The option prices derived in these earlier studies are special cases located within the continuous time limits of the stochastic dominance bounds derived by (8.1).

9 Multiple periods with transaction costs and with convex payoffs

Constantinides and Perrakis (2002) recognized that it is possible to recursively apply the single-period approach with transaction costs and derive stochastic dominance bounds on option prices in a market with intermediate trading over the life of the options. The task of computing these bounds is easy compared to the full-fledged investigation of the feasibility of conditions (3.5)–(3.10) for large T for two reasons. As with the no transaction costs case, the derivation of the bounds takes advantage of the special structure of the payoff

[11] The conditions for the preservation of convexity were first presented by Bergman et al. (1996). Convexity is preserved in all one-dimensional diffusions and in most two-dimensional diffusions that have been used in the option pricing literature.

of a call or put option, specifically the convexity of the payoff as a function of the stock price. Second, the set of assets is limited to three assets: the bond, stock, and one option, the test option. Below, we state these bounds without proof.

At any time t prior to expiration, the following is an upper bound on the price of a call:

$$\bar{c}(S_t, t) = \frac{(1+k)}{(1-k)\{(1+\delta)\hat{z}\}^{T-t}} E\big[\big[(1+\delta)S_T - K\big]^+ \mid S_t\big], \tag{9.1}$$

where $(1+\delta)\hat{z}$ is the expected return on the stock per unit time. Observe that (9.1) is the same as the upper bound given in (4.13) for $z_{\min} = 0$ times the roundtrip transaction cost. The tighter upper bound given in (4.8), (4.11), and (8.1) does not survive the introduction of transaction costs and is eventually dominated by (9.1).

A partition-independent lower bound for a call option can also be found, but only if it is additionally assumed that there exists at least one trader for whom the investment horizon coincides with the option expiration, $T = T'$. In such a case, transaction costs become irrelevant in the put-call parity and the following is a lower bound[12]:

$$\begin{aligned} \underline{c}(S_t, t) = {} & (1+\delta)^{t-T} S_t - K/R^{T-t} \\ & + E\big[(K - S_T)^+ \mid S_t\big] \big/ \big\{(1+\delta)\hat{z}\big\}^{T-t}, \end{aligned} \tag{9.2}$$

where R is one plus the risk free interest rate per unit time.

Put option upper and lower bounds also exist that are independent of the frequency of trading. They are given as follows:

$$\overline{p}(S_t, t) = \frac{K}{R^{T-t}} + \frac{1-k}{1+k}\big((1+\delta)\hat{z}\big)^{t-T}\big[E\big[[K - S_T]^+\big] - K \mid S_t\big], \tag{9.3}$$

and

$$\underline{p}(S_t, t) \begin{cases} \big((1+\delta)\hat{z}\big)^{t-T} \dfrac{1-k}{1+k} E\big[[K - S_T]^+ \mid S_t\big], & t \leqslant T-1, \\ [K - S_T]^+, & t = T. \end{cases} \tag{9.4}$$

The bounds presented in (9.1)–(9.4) may not be the tightest possible bounds for any given frequency of trading. Nonetheless, they have the property that they do not depend on the frequency of trading over the life of the option. For a comprehensive discussion and derivation of these and other possibly tighter bounds that are specific to the allowed frequency of trading, see Constantinides and Perrakis (2002). See also Constantinides and Perrakis (2007) for extensions to American-style options and futures options.

[12] In the special case of zero transaction costs, the assumption $T = T'$ is redundant because the put-call parity holds.

10 Empirical results

A robust prediction of the BSM option pricing model is that the volatility implied by market prices of options is constant across strike prices. Rubinstein (1994) tested this prediction on the S&P 500 index options (SPX), traded on the Chicago Board Options Exchange, an exchange that comes close to the dynamically complete and perfect market assumptions underlying the BSM model. From the start of the exchange-based trading in April 1986 until the October 1987 stock market crash, the implied volatility is a moderately downward-sloping function of the strike price, a pattern referred to as the "volatility smile", also observed in international markets and to a lesser extent on individual-stock options. Following the crash, the volatility smile is typically more pronounced.[13]

An equivalent statement of the above prediction of the BSM model, that the volatility implied by market prices of options is constant across strike prices, is that the *risk-neutral* stock price distribution is lognormal. Aït-Sahalia and Lo (1998), Jackwerth and Rubinstein (1996), and Jackwerth (2000) estimated the risk-neutral stock price distribution from the cross section of option prices.[14] Jackwerth and Rubinstein (1996) confirmed that, prior to the October 1987 crash, the risk-neutral stock price distribution is close to lognormal, consistent with a moderate implied volatility smile. Thereafter, the distribution is systematically skewed to the left, consistent with a more pronounced smile.

Several no-arbitrage models have been proposed and tested that generalize the BSM model. These models explore the effects of generalized stock price processes including stock price jumps and stochastic volatility and typically generate a volatility smile. The textbooks by Hull (2006) and McDonald (2005) provide excellent discussions of these models.

Economic theory imposes restrictions on equilibrium models beyond merely ruling out arbitrage. As we have demonstrated in Section 3, if prices are set by a utility-maximizing trader in a frictionless market, the pricing kernel must be a monotonically decreasing function of the market index price. To see this, the pricing kernel equals the representative agent's intertemporal marginal rate of substitution over each trading period. If the representative agent has *state independent* (derived) utility of wealth, then the concavity of the utility function implies that the pricing kernel is a decreasing function of wealth. Under the two maintained hypotheses that the marginal investor's (derived) utility of wealth is state independent *and* wealth is monotone increasing in the market index level, the pricing kernel is a decreasing function of the market index level.

[13] Brown and Jackwerth (2004), Jackwerth (2004), Shefrin (2005), and Whaley (2003) review the literature and potential explanations.

[14] Jackwerth (2004) reviews the parametric and non-parametric methods for estimating the risk-neutral distribution.

In a frictionless representative-agent economy, Aït-Sahalia and Lo (2000), Jackwerth (2000), and Rosenberg and Engle (2002) estimated the pricing kernel implied by the observed cross section of prices of S&P 500 index options as a function of wealth, where wealth is proxied by the S&P 500 index level. Jackwerth (2000) reported that the pricing kernel is everywhere decreasing during the pre-crash period 1986–1987 but widespread violations occur over the post-crash period 1987–1995. Aït-Sahalia and Lo (2000) reported violations in 1993 and Rosenberg and Engle (2002) reported violations over the period 1991–1995.[15] On the other hand, Bliss and Panigirtzoglou (2004) estimated plausible values for the risk aversion coefficient of the representative agent, albeit under the assumption of power utility, thus restricting the shape of the pricing kernel to be monotone decreasing in wealth.

Several theories have been suggested to explain the inconsistencies with the BSM model and the violations of monotonicity of the pricing kernel. Bollen and Whaley (2004) suggested that *buying pressure* drives the volatility smile while Han (2004) and Shefrin (2005) provided behavioral explanations based on *sentiment*. However, most of the discussion has focused on the *risk premia* associated with *stock market crashes* and *state dependence of the pricing kernel*.

Bates (2001) introduced heterogeneous agents with utility functions that explicitly depend on the number of stock market crashes, over and above their dependence on the agent's terminal wealth. The calibrated economy exhibits the inconsistencies with the BSM model but fails to generate the non-monotonicity of the pricing kernel. Brown and Jackwerth (2004) suggested that the reported violations of the monotonicity of the pricing kernel may be an artifact of the maintained hypothesis that the pricing kernel is state independent but concluded that volatility cannot be the sole omitted state variable in the pricing kernel.

Pan (2002), Garcia et al. (2003), and Santa-Clara and Yan (2004), among others, obtained plausible parameter estimates in models in which the pricing kernel is state dependent, using panel data on S&P 500 options. Others calibrated equilibrium models that generate a volatility smile pattern observed in option prices. Liu et al. (2005) investigated rare-event premia driven by uncertainty aversion in the context of a calibrated equilibrium model and demonstrated that the model generates a volatility smile pattern observed in option prices. Benzoni et al. (2005) extended the above approach to show that uncertainty aversion is not a necessary ingredient of the model. More significantly, they demonstrated that the model can generate the stark regime shift that occurred at the time of the 1987 crash. While not all of the above papers deal explicitly with the monotonicity of the pricing kernel, they do address the problem of reconciling the option prices with the historical index record.

[15] Rosenberg and Engle (2002) found violations when they used an orthogonal polynomial pricing kernel but not when they used a power pricing kernel.

These results are encouraging but stop short of demonstrating absence of stochastic dominance violations on a month-by-month basis in the cross section of S&P 500 options. This inquiry is the focus in Constantinides et al. (2007), hereafter CJP. CJP empirically investigated whether the observed cross sections of S&P 500 index option prices are consistent with various economic models that explicitly allow for a dynamically incomplete market and also recognize trading costs and bid–ask spreads. In the first part of their paper, CJP introduced transaction costs (trading fees and bid–ask spreads) in trading the index and options and investigated the extent to which violations of stochastic dominance, gross of transaction costs, are explained by transactions costs. They found that transaction costs decrease the frequency of violations but violations persist in several months both before and after the October 1987 crash.

Then CJP explored the second maintained hypothesis that every economic agent's wealth on the expiration date of the options is monotone increasing in the S&P 500 index price on that date. This assumption is unwarranted once we recognize that trading occurs over the (one-month) life of the options. With intermediate trading, a trader's wealth on the expiration date of the options is generally a function not only of the price of the market index on that date but also of the entire path of the index level. Thus the pricing kernel is a function not only of the index level but also of the *entire path* of the index level. CJP explored the month-by-month violations of stochastic dominance while allowing the pricing kernel to depend on the path of the index level.

In estimating the real distribution of the S&P 500 index returns, CJP refrained from adopting a particular parametric form of the distribution and proceeded in four different ways. In the first approach, they estimated the *unconditional* distribution as the histograms extracted from two different *historical* index data samples covering the periods 1928–1986 and 1972–1986. In the second approach, they estimated the *unconditional* distribution as the histograms extracted from two different *forward-looking* samples, one that includes the October 1987 crash (1987–2002) and one that excludes it (1988–2002). In the third approach, CJP modeled the variance of the index return as a GARCH (1, 1) process and estimated the *conditional* variance over the period 1972–2002 by the semiparametric method of Engle and Gonzalez-Rivera (1991) that does not impose the restriction that conditional returns are normally distributed. In the fourth approach, CJP used the VIX-implied volatility as an estimate of the *conditional* variance.

Based on the index return distributions extracted in the above four approaches, CJP tested the compliance of option prices with the predictions of models that sequentially introduce market incompleteness, transactions costs, and intermediate trading over the life of the options.

CJP's empirical design allows for at least three implications associated with state dependence. First, each month they searched for a pricing kernel to price the cross section of one-month options without imposing restrictions on the time series properties of the pricing kernel month by month. Thus they allowed the pricing kernel to be state dependent. Second, in the second part of their investigation, CJP allowed for intermediate trading; a trader's wealth on the expiration date of the options is generally a function not only of the price of the market index on that date but also of the entire path of the index level thereby rendering the pricing kernel state dependent. Third, CJP allowed the variance of the index return to be state dependent and employed the estimated conditional variance.

A novel finding is that, even though pre-crash option prices follow the BSM model reasonably well, it does not follow that these options are correctly priced. Pre-crash option prices are incorrectly priced, if index return expectations are formed based on the historical experience. Furthermore, some of these prices lie below the theoretical bounds, contrary to received wisdom that historical volatility generally underprices options in the BSM model.

Another novel finding dispels the common misconception that the observed smile is too steep after the crash. Most of the violations post-crash are due to the option smile not being steep enough relative to expectations on the index price formed post-crash. Even though the BSM model assumes that there is no smile, an investor who properly understood the post-crash distribution of index returns should have priced the options with a steeper smile than the smile reflected in the actual option prices.

In all cases, there is a higher percentage of months with stochastic dominance violations by out-of-the-money calls (or, equivalently, in-the-money puts) than by in-the-money calls, suggesting that the mispricing is caused by the right-hand tail of the index return distribution and not by the left-hand tail. This observation is novel and contradicts the common inference drawn from the observed implied volatility smile that the problem lies with the left-hand tail of the index return distribution.

Finally, CJP found that the effect of allowing for one intermediate trading date over the life of the one-month options is to uniformly decrease the number of feasible months in each subperiod. They concluded that intermediate trading strengthens the single-period evidence of systematic stochastic dominance violations.

Constantinides et al. (2007) extended the results in CJP to American options on S&P 500 index *futures*. They demonstrated corresponding violations and implemented trading strategies that exploit the violations.

11 Concluding remarks

We presented an integrated approach to the pricing of options that allows for incomplete and imperfect markets. The BSM option pricing model is the nested case of complete and perfect markets. When the market is incomplete, imperfect, or both, the principle of no-arbitrage by itself implies restrictions on the prices of options that are too weak to be useful to either price options or confront the data with a testable hypothesis.

Instead of the principle of the absence of arbitrage that underlies the BSM model, we introduced the economic restriction that at least one risk-averse trader is a marginal investor in the options and the underlying security. Given the cross section of the prices of options and the real probability distribution of the return of the underlying security, the implied restrictions may be tested by merely solving a linear program. We also showed that the economic restrictions may be expressed in the form of upper and lower bounds on the price of an option, given the prices of the stock and the other outstanding options.

By providing an integrated approach to the pricing of options that allows for incomplete and imperfect markets, we provided testable restrictions on option prices that include the BSM model as a special case. We reviewed the empirical evidence on the prices of S&P 500 index options. The economic restrictions are violated surprisingly often, suggesting that the mispricing of these options cannot be entirely attributed to the fact that the BSM model does not allow for market incompleteness and realistic transaction costs. These are indeed exciting developments and are bound to stimulate further theoretical and empirical work to address the month-by-month pattern of option price violations.

Acknowledgements

Constantinides acknowledges financial support from the Center for Research in Security Prices of the University of Chicago and Perrakis from the Social Sciences and Humanities Research Council of Canada.

References

Aït-Sahalia, Y., Lo, A. (1998). Nonparametric estimation of state price densities implicit in financial asset prices. *Journal of Finance* 53, 499–547.

Aït-Sahalia, Y., Lo, A. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics* 94, 9–51.

Amin, K.I. (1993). Jump diffusion option valuation in discrete time. *Journal of Finance* 48, 1833–1863.

Bates, D.S. (1991). The crash of '87: Was it expected? The evidence from option markets. *Journal of Finance* 46, 1009–1044.

Bates, D.S. (2001). The market for crash risk. Working paper. University of Iowa, Iowa City.

Benzoni, L., Collin-Dufresne, P., Goldstein, R.S. (2005). Can standard preferences explain the prices of out-of-the-money S&P 500 options? Working paper. University of Minnesota.

Bergman, Y.Z., Grundy, B.D., Wiener, Z. (1996). General properties of option prices. *The Journal of Finance* 51, 1573–1610.

Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–659.

Bliss, R.R., Panigirtzoglou, N. (2004). Option-implied risk aversion estimates. *Journal of Finance* 59, 407–446.

Bollen, N., Whaley, R. (2004). Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* 59, 711–753.

Brown, D.P., Jackwerth, J. (2004). The kernel puzzle: Reconciling index option data and economic theory. Working paper. University of Wisconsin, Madison.

Constantinides, G.M. (1978). Market risk adjustment in portfolio valuation. *Journal of Finance* 33, 603–616.

Constantinides, G.M. (1979). Multiperiod consumption and investment behavior with convex transactions costs. *Management Science* 25, 1127–1137.

Constantinides, G.M., Perrakis, S. (2002). Stochastic dominance bounds on derivatives prices in a multiperiod economy with proportional transaction costs. *Journal of Economic Dynamics and Control* 26, 1323–1352.

Constantinides, G.M., Perrakis, S. (2007). Stochastic dominance bounds on American option prices in markets with frictions. *Review of Finance* 11, 71–115.

Constantinides, G.M., Zariphopoulou, T. (1999). Bounds on prices of contingent claims in an intertemporal economy with proportional transaction costs and general preferences. *Finance and Stochastics* 3, 345–369.

Constantinides, G.M., Zariphopoulou, T. (2001). Bounds on derivative prices in an intertemporal setting with proportional transaction costs and multiple securities. *Mathematical Finance* 11, 331–346.

Constantinides G.M., Czerwonko M., Jackwerth J., Perrakis S. (2007). Are options on index futures profitable for risk averse investors? Empirical Evidence. *Working paper*. University of Chicago, Chicago.

Constantinides, G.M., Jackwerth, J., Perrakis, S. (2007). Mispricing of S&P 500 index options. *Review of Financial Studies*, in press.

Cox, J., Ross, S.A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3, 145–166.

Cox, J., Ross, S.A., Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics* 7, 229–263.

Delbaen, F., Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* 300, 463–520.

Engle, R.F., Gonzalez-Rivera, G. (1991). Semiparametric ARCH models. *Journal of Business and Economic Statistics* 9/4, 345–359.

Garcia, R., Luger, R., Renault, E. (2003). Empirical assessment of an intertemporal option pricing model with latent variables. *Journal of Econometrics* 116, 49–83.

Han, B. (2004). Limits of arbitrage, sentiment and index option smile. Working paper. Ohio State University.

Harrison, J.M., Kreps, D.M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20, 381–408.

Harrison, J.M., Pliska, S.R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications* 11, 215–260.

Huang, J. (2005). Option bounds and second order arbitrage opportunities. Working paper. Lancaster University.

Hull, J.C. (2006). *Options, Futures, and Other Derivatives*. Prentice Hall.

Jackwerth, J. (2000). Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* 13, 433–451.

Jackwerth, J. (2004). Option-implied risk-neutral distributions and risk aversion. ISBN 0-943205-66-2, Research Foundation of AIMR, Charlotteville, USA.

Jackwerth, J., Rubinstein, M. (1996). Recovering probability distributions from option prices. *Journal of Finance* 51, 1611–1631.

Levy, H. (1985). Upper and lower bounds of put and call option value: Stochastic dominance approach. *Journal of Finance* 40, 1197–1217.

Liu, J., Pan, J., Wang, T. (2005). An equilibrium model of rare-event premia and its implications for option smirks. *Review of Financial Studies* 18, 131–164.

McDonald, R.L. (2005). *Derivatives Markets*. Addison–Wesley.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4, 141–184.

Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3, 125–144.

Merton, R.C. (1982). On the mathematics and economics assumptions of continuous-time models. In: *Essays in Honor of Paul Cootner*. Prentice Hall, Englewood Cliffs, NJ.

Oancea, I., Perrakis, S. (2006). Stochastic dominance and option pricing: An alternative paradigm. Working paper. Concordia University.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.

Perrakis, S. (1986). Option bounds in discrete time: Extensions and the pricing of the American put. *Journal of Business* 59, 119–141.

Perrakis, S., Ryan, P.J. (1984). Option pricing bounds in discrete time. *Journal of Finance* 39, 519–525.

Rendleman, R., Bartter, B. (1979). Two-state option pricing. *Journal of Finance* 34, 1092–1110.

Ritchken, P.H. (1985). On option pricing bounds. *Journal of Finance* 40, 1219–1233.

Ritchken, P.H., Kuo, S. (1988). Option bounds with finite revision opportunities. *Journal of Finance* 43, 301–308.

Rosenberg, J., Engle, R. (2002). Empirical pricing kernels. *Journal of Financial Economics* 64, 341–372.

Ross, S.A. (1976). Options and efficiency. *Quarterly Journal of Economics* 90, 75–89.

Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance* 3, 771–818.

Ryan, P.J. (2000). Tighter option bounds from multiple exercise prices. *Review of Derivatives Research* 4 (2), 155–188.

Ryan, P.J. (2003). Progressive option bounds from the sequence of concurrently expiring options. *European Journal of Operational Research* 151, 193–223.

Santa-Clara, P., Yan, S. (2004). Jump and volatility risk and risk premia: A new model and lessons from S&P 500 options. Working paper. UCLA.

Shefrin, H. (2005). *A Behavioral Approach to Asset Pricing*. Elsevier/North-Holland, Amsterdam.

Whaley, R.E. (2003). Derivatives. In: Constantinides, G.M., Harris, M., Stulz, R. (Eds.), *Financial Markets and Asset Pricing: Handbook of the Economics of Finance, vol. 1B*. In: *Handbooks in Economics*, vol. 21. Elsevier/North-Holland, Amsterdam.