



Optimal positioning in derivative securities

P. Carr & D. Madan

To cite this article: P. Carr & D. Madan (2001) Optimal positioning in derivative securities, , 1:1, 19-37, DOI: [10.1080/713665549](https://doi.org/10.1080/713665549)

To link to this article: <https://doi.org/10.1080/713665549>



Published online: 08 Apr 2010.



Submit your article to this journal [↗](#)



Article views: 451



View related articles [↗](#)



Citing articles: 46 View citing articles [↗](#)

Optimal positioning in derivative securities

Peter Carr¹ and Dilip Madan²

¹ Banc of America Securities, 40th Floor, 9 West 57th Street, New York, NY 10019, USA

² Robert H Smith School of Business, University of Maryland, College Park, MD 20742, USA

E-mail: pcarr@bofasecurities.com and dbm@rhsmith.umd.edu

Received 21 June 2000, in final form 13 November 2000

Abstract

We consider a simple single period economy in which agents invest so as to maximize expected utility of terminal wealth. We assume the existence of three asset classes, namely a riskless asset (the bond), a single risky asset (the stock), and European options of all strikes (derivatives). In this setting, the inability to trade continuously potentially induces investment in all three asset classes. We consider both a partial equilibrium where all asset prices are initially given, and a more general equilibrium where all asset prices are endogenously determined. By restricting investor beliefs and preferences in each case, we solve for the optimal position for each investor in the three asset classes. We find that in partial or general equilibrium, heterogeneity in preferences or beliefs induces investors to hold derivatives individually, even though derivatives are not held in aggregate.

1. Introduction

The portfolio selection problem pioneered by Markowitz [36] and Merton [37] generally does not formally consider derivative securities as potential investment vehicles. Similarly, the asset allocation approach favoured by practitioners does not³ consider derivatives as a distinct asset class. If derivatives (forwards, futures, swaps, options and exotics) are considered at all, they are only viewed as tactical⁴ vehicles for efficiently re-allocating funds across broad asset classes, such as cash, fixed-income, equity and alternative investments. While the question of optimal derivatives positioning has been addressed in the literature, few papers⁵ have focused directly on the demand for derivatives, especially in a general equilibrium.

³ For example, three recent comprehensive texts on asset allocation are Gibson [25], Vince [50], or Leibowitz *et al* [34], none of which cover derivatives.

⁴ See Evnine and Henriksson [21] and Tilley and Latainer [49] for discussions on the use of options in an asset allocation framework.

⁵ Some guidance may be gleaned from the standard literature in which dynamic trading in the underlying assets completes markets. For example, the Cox and Huang [14] solution to the Merton [37] problem can be used to obtain the payoff that is actually being replicated through dynamic trading in the underlying assets.

The absence of a direct focus on optimal derivatives positioning is partly due to the complexity of the problem and partly due to the overwhelming success of the arbitrage-based models for pricing derivatives. Since these models are dynamically complete in the underlying assets, these models are subject to the Hakansson [27, 28] ‘catch 22’: although derivatives can be perfectly priced in these economies, there is no justification for their existence. In these models, the optimal position in derivatives is usually either indeterminate or infinite, depending on whether an investor agrees or disagrees with the derivative’s market price. Since these models were developed for the purpose of pricing, they are ill-suited for the development of a normative theory, capable of guiding investors, regulators and other market participants about the efficacy of derivatives markets.

The purpose of this paper is to delineate the optimal derivatives positions for investors when they cannot trade continuously. In contrast to most of the literature on optimal allocation across derivatives, we pay particular attention to positioning in a general equilibrium setting. We show that under reasonable market conditions, derivatives comprise an important, interesting and separate asset class, imperfectly

correlated with other broad asset classes. Our paper extends a line of investigation initiated by Ross [43] in which derivatives complete the market. Working in a single period setting, we show how investors can determine their optimal investment in the riskless asset, a risky asset, and in options written on this risky asset. The optimal *positions* in these three asset classes are determined by the investors' *preferences* for risk, their *probabilities* of underlying returns, and by market *prices* of the three asset types. As in Pliska [40] and Cox [13], we solve for these optimal positions by dividing the optimal investment problem into two subproblems. The first subproblem is to determine the optimal wealth profile as a function of the market price of the single risky asset (called the stock). The second subproblem is to determine the positions needed in available instruments so as to achieve this optimal payoff function.

Under certain conditions such as continuously open markets, continuous price processes, zero transactions costs, no leverage or short-selling constraints, symmetric information, and simple volatility structures, the introduction of derivatives does not enhance the investment possibility frontier. Thus, under some fairly stringent conditions, investors will be indifferent to the introduction of derivatives markets. As a result, the optimal investment decision can be analysed in a setting in which derivatives are absent. This is the viewpoint taken in the continuous-time analyses by Merton [37], Brennan *et al* [8], Pliska [41] and Cox and Huang [14]. Merton [37] derives the optimal consumption and portfolio policies by solving a nonlinear partial differential equation (pde) governing the utility of optimized wealth. Brennan *et al* [8] numerically solve the Merton pde in a three-state variable Markov context assuming that investors have proportional risk tolerance. Pliska [41] showed that martingale methods can be used to fruitfully decompose the investment problem as described above. Cox and Huang [14] also use this approach to reduce Merton's problem to one of solving a linear pde, simplifying the analysis significantly.

A violation of any of the perfect market conditions described above can introduce a demand by investors for derivatives. The literature on this demand can be dichotomized into at least three streams. An early stream focused on the welfare gains achieved by introducing options into an economy. This literature was initiated by Ross [43], and received important contributions from Hakansson [27, 28], Breeden and Litzenberger [6], Friesen [24], Kreps [33] and Arditti and John [1]. We contribute to this literature stream by demonstrating sufficient conditions under which positions in derivatives improve the welfare of individuals.

A more recent stream discusses the optimal design of derivatives contracts. In particular, Johnston and McConnell [32] and Duffie and Jackson [17] study the optimal design of futures contracts. Brennan and Solanki [9] and Shimko [47] study the optimal design of nonlinear payoffs in a one period model for a single investor. Draper and Shimko [16] discuss how a discrete-time dynamic strategy in a portfolio of options can be used to help hedge a nonlinear liability. While the papers on optimal security design represent important contributions to the literature, most of these analyses are conducted in a partial equilibrium context. In contrast, this paper considers optimal

positioning and security design when the risk-neutral density is endogeneously determined by market clearing conditions. We also contribute to this stream by presenting sufficient conditions under which each investor's portfolio optimally separates into a finite number of derivative funds, where each fund has a payoff which is independent of the investor's attributes.

A third stream of the literature takes payoffs with certain properties as given, and then determines the characteristics of investors who would consider these payoffs as optimal. For example, Benninga and Blume [2] consider the kinked increasing convex payoff associated with the standard form of portfolio insurance. They show that the utility function of an investor for whom this payoff is optimal is highly non-standard. In a seminal work, Leland [35] focuses on optimal payoffs which are globally convex. He shows that when the investor shares the market's beliefs, then the optimal payoff is globally convex if and only if the investor's risk tolerance increases with his wealth more rapidly than the aggregate investor's risk tolerance increases with market wealth. Assuming linear risk tolerance with identical cautiousness and log-normal beliefs with identical volatility, he also shows that the investor's optimal payoff is globally convex if and only if the investor has a higher expected return than the market. As in our work, Leland generalizes the work of Brennan and Solanki [9] to heterogeneous beliefs. Thus, the Leland paper both pre-dates Brennan and Solanki and is more general.

Our paper differs from Benninga and Blume [2] and Leland [35] primarily in its focus. We address the inverse problem of determining what the optimal payoffs are, given prices, preferences and probability beliefs. For example, we show when an investor believes volatility will be high, then we show that this belief induces the sale of at-the-money options, while his risk aversion induces the purchase of out-of-the-money options. Leland also assumes conditions which lead to the existence of a representative agent, whereas we assume greater heterogeneity among our agents.

Our paper is also related to some recent contributions by Benninga and Mayshar [3], Franke *et al* [22, 23] and Cassano [12]. All of these papers focus on the demand for options in a general equilibrium setting by agents with HARA (hyperbolic absolute risk aversion) utility functions. In this context, Benninga and Mayshar show that heterogeneity in the degree of CRRA across investors induces a representative agent with decreasing relative risk aversion. In this economy, the Black–Scholes formula does not hold and all option prices are overpriced compared to Black–Scholes. Furthermore, implied volatility varies with strike prices in a manner consistent with their behaviour in practice. Assuming that the representative investor has decreasing relative risk aversion as derived in Benninga and Mayshar [3], Franke *et al* [23] derive necessary conditions for an implied volatility 'smile'. In another paper, Franke *et al* [22] emphasize background risk rather than heterogeneity in preferences as an explainer of option demand. Their analysis shows that high levels of this risk can lead to a demand for options, even if there are no differences in preferences. Since the presence of heterogeneous background

risk alters the distribution of individual wealth in the context of homogeneous beliefs, we suspect that these results could be related to the consequences of heterogeneous beliefs, which are investigated here. Also, the three-fund separation results which we obtain in general equilibrium are very similar to those in Franke *et al* [22]. Cassano [12] considers the consequences of our setup for explaining the determinants of option volume. Just as heterogeneity in preferences and/or beliefs induces valuation results that are more consistent with observed prices, these heterogeneous models can induce option volume results that are also more consonant with reality.

Our first set of results pertain to a Markowitz-style [36] partial equilibrium setting in which prices of the bond, stock, and options are taken as given. However, since variance is an inadequate risk measure when analysing derivatives positions (Dybvig and Ingersoll [20]), we instead consider an investor who maximizes the expected value of an increasing, concave utility function⁶. For each such investor and each traded asset, we derive the concept of a personalized price and show that optimal positions are determined so that personalized prices equate to market prices for all traded assets. When options of all strikes trade, these conditions determine optimal individual derivatives positions. In general, the optimal financial payoff is determined by comparing the investor's probability density with the market's 'risk-neutral' density. In particular, a given investor has larger payoffs in states which the investor perceives as more likely than the market suggests, as reflected in the prices of state-contingent claims. This propensity to move payoffs into states with greater relative probability is tempered by the degree of the investor's risk aversion.

Specializing preferences to linear risk tolerance, we observe that zero cautiousness investors and positive cautiousness investors behave quite differently. In particular, zero cautiousness investors fix their investment in the optimal customized fund at their risk tolerance and place all wealth above this level in the riskless asset. In contrast, investors with positive cautiousness fix their investment in the riskless asset and invest all wealth above this threshold in a customized derivative.

We alternatively specialize beliefs by restricting attention to log-normal personal beliefs and to log-normal risk-neutral densities, as in the Black–Scholes model. Under this log-normality assumption, we determine the optimal payoff for an investor given that his personal density differs from the risk-neutral density in either the mean or variance or both. When we focus on an investor who differs from the market only on expected return, we obtain the predictable result that his optimal payoff is increasing with the stock price if he is bullish, and is decreasing otherwise. However, we find that the indifference point between long and short is when the expected return is the risk-free rate. Consequently, an investor whose expected return is above the risk-free rate, but below that required for the risk borne, should actually have an increasing payoff. For a bullish investor, we also find that even though

the investor agrees with the market on volatility, risk aversion causes the payoff to be concave for a highly risk-averse investor and convex for a less risk-averse investor. Thus, a highly risk-averse investor who believes that the stock has a positive risk premium will prefer an increasing concave payoff similar to that in a buy-write strategy. While there is a combination of beliefs and preferences for which the optimal payoff is linear, this combination is a very special case. Consequently, in most cases, options enhance efficiency even if the investor's views diverge from the market only on the mean.

We also derive optimal payoffs for the complementary case where an investor agrees with the market on expected return, but disagrees with the market on volatility. We obtain the predictable result that the optimal payoff is U-shaped if the investor believes that implied volatilities are low. However, if the investor believes that implied volatilities are high, then the optimal payoff is not that of an inverted U over the whole domain, but in general resembles that of a unimodal probability density. Thus, the volatility beliefs induce a concave payoff for smaller price moves, but risk aversion convexifies the payoff for larger moves, in order to dampen the dispersion of payoffs. Thus, our analysis allows us to determine the optimal payoff when beliefs and preferences conflict in their implications for positions in options.

For log-normal beliefs and linear risk tolerance preferences, we find that positive cautiousness investors who agree with the market on volatility want a customized derivative whose payoff is that of the stock price raised to a power. The power is in fact given by the Sharpe ratio of the asset so that Sharpe ratios above unity induce convex payoffs, while Sharpe ratios between zero and one induce concave payoffs. In all cases, the payoff is bounded below. In contrast, zero cautiousness investors who agree on volatility invest in the log contract, which has positive probability of infinite loss. When these investors disagree with the market on volatility, then zero cautiousness investors prefer payoffs which are quadratic in the log of the stock price, while positive cautiousness investors prefer that the log of their payoff be quadratic in the log of the stock price.

While our partial equilibrium analysis suggests that options are optimally held at the individual level, it must be recognized that options are in zero net supply in aggregate. Thus, a natural question that arises is whether our conclusions regarding the optimality of options positions survive when we shift attention to a general equilibrium. Furthermore, it is interesting to consider who buys and who sells options, especially when beliefs and preferences conflict. To consider the effects of these supply conditions, we consider an n -person general equilibrium. Here, instead of taking asset prices as given, we solve for the state pricing density that simultaneously clears the stock, bond and options markets.

We first show that under homogeneous beliefs, differences in risk aversion across investors can induce a demand for derivatives on the part of all investors. Importantly, we also observe that when all investors agree on the probability distribution of returns, the risk-neutral distribution can differ from this common distribution in interesting ways. For example, in many cases the risk-neutral distribution is often

⁶ Jarrow and Madan [31] show that mean-variance preferences lead to negative personalized state prices. Consequently, investors with mean-variance preference will refuse free calls, since the perceived risk outweighs the expected return.

not in the same parametric class of densities as the common prior belief. Furthermore, when beliefs are homogeneous and symmetric about their mean, the risk-neutral density is often skewed to the left. This skew arises from strict risk aversion on the part of all individuals in the economy and the requirement that just the stock be held in aggregate.

For the special case of homogeneous beliefs, linear risk tolerance, and identical cautiousness, the resulting two-fund separation implies that investors in the economy do not hold derivatives positions. On the other hand, so long as an investor has beliefs that differ from a risk tolerance weighted average of individual beliefs, then the investor optimally holds derivatives. Furthermore, when each investor's beliefs can be represented by a set of basis functions, then these basis functions comprise the derivatives funds into which optimal portfolios separate. For example, when investors have constant risk tolerance and heterogeneous log-normal beliefs, we show that a four-fund separation holds in which besides the stock and the bond, investors also take positions in a claim which pays the log of the stock price and a second claim which pays the square of the log.

The outline of the paper is as follows. In section 2, we review the well-known relationship between option prices and the risk-neutral distribution in a complete market. We also exhibit for the first time an explicit decomposition of an arbitrary payoff into a portfolio of bonds, stocks and options. We then derive the investor's optimal payoff, given arbitrary preferences, beliefs and risk-neutral distribution. These results are specialized to utility functions displaying linear risk tolerance in section 3. Section 4 develops our partial equilibrium model in which the personal and risk-neutral distributions are assumed to be log-normal, as in the Black–Scholes model. Our focus is on the positions taken in derivatives, with particular interest paid to the case when preferences and beliefs conflict⁷ in their implications when analysed individually. Section 5 details our general equilibrium model in which the risk-neutral distribution is not specified *ex ante*, but rather is determined by the requirement that markets clear in all assets. We derive explicit formulae for optimal payoffs and the risk-neutral distribution under heterogeneous log-normal beliefs and heterogeneous logarithmic or exponential preferences. The final section summarizes the paper and provides directions for future research. The appendices contain proofs of our technical results.

2. The optimal investment problem

When derivatives are considered as potential investment vehicles, the problem is complicated by the fact that over-the-counter markets permit investment in a virtually unlimited array of alternatives. In our single period setting with a single underlying asset, the problem can be reduced to that of determining an optimal piecewise linear payoff made up of a portfolio of options. However, this simplification comes at

the expense of not being able to appreciate the fundamental principles driving optimal positioning, as well as the inability to derive closed form solutions for the optimal payoffs of potential interest to the investing community. For these reasons, we consider a setting in which there is an entire continuum of option strikes. This setup shares the analytical advantages of the Merton [37] continuous-time economy, except that we now deal with infinitely many assets at a single point in time.

To deal with the infinite asset problem, we break the investment problem down into a subproblem of determining the optimal terminal wealth profile, and a second problem of determining how to span this payoff with the available assets. The next subsection reviews how any payoff can be spanned by a linear combination of the payoffs from a riskless asset, a single risky asset and options of all strikes. The following subsection shows how preferences, probabilities and prices interact to determine the optimal payoff, and consequently, the optimal position.

2.1. Spanning

Consider a one period model in which investments are made at time 0 with all payoffs being received at time 1. There is a riskless asset costing B_0 initially and paying unity at time 1, which we call the bond. There is also a single risky asset, costing S_0 initially and paying the random amount S at time 1, which we call the stock. In addition to markets in the bond and stock, we assume that markets exist for out-of-the-money European puts and calls of all strikes. While this assumption is not standard, it allows us to examine the question of optimal positioning in a complete market without requiring the heavy machinery of continuous-time mathematics. We note that the assumption of a continuum of strikes is essentially the counterpart of the standard assumption of continuous trading. Just as the latter assumption is frequently made as a reasonable approximation to an environment where investors can trade frequently, we take our assumption as a reasonable approximation when there are a large but finite number of option strikes (e.g. for the S&P 500). In each case, the assumption adds analytic tractability without representing a large departure from reality.

For this market structure, it is known from the literature that investors may purchase any smooth function of the underlying stock price by taking a static position at time 0 in these markets⁸. It is also known, (see Breeden and Litzenberger [6]), that there is a unique risk-neutral density that may be identified from the prices of options. The market completeness results of the literature (Green and Jarrow [26] and Nachman [38]), however, are not constructive and do not describe the static position that needs to be taken in the various markets.

We show here that any twice continuously differentiable function, $f(S)$, of the terminal stock price S , can be replicated by a unique initial position of $f(S_0) - f'(S_0)S_0$ unit discount

⁷ For example, we examine whether a highly risk averse individual who thinks implied volatility is high should buy options or sell them.

⁸ This observation was first noted in Breeden and Litzenberger [6] and established formally in Green and Jarrow [26] and Nachman [38].

bonds, $f'(S_0)$ shares, and $f''(K)dK$ out-of-the-money options of all strikes K :

$$f(S) = [f(S_0) - f'(S_0)S_0] + f'(S_0)S + \int_0^{S_0} f''(K)(K - S)^+ dK + \int_{S_0}^{\infty} f''(K)(S - K)^+ dK. \quad (1)$$

The positions in the bond and the stock create a tangent to the payoff at the initial stock price. The positions in the out-of-the-money options are used to bend the tangent line so as to match the payoff at all price levels. The derivation leading to equation (1) is contained in appendix 1.

An important special case of (1) is put-call-parity, where we take $f(S)$ to be the payoff of an in-the-money call $(S - K_0)^+$ for $K_0 < S_0$. In this case $f'(S_0) = 1$, while $f''(K_0)$ is a delta function⁹ centred at K_0 . Equation (1) may also be used to identify the risk-neutral density. Since the payoff $f(S)$ is linear in the payoffs from the available assets, the same linear relationship must prevail among the initial values. Specifically, letting $V_0[f]$ denote the initial value of the arbitrary¹⁰ payoff $f(\cdot)$, and letting B_0 , $P_0(K)$ and $C_0(K)$ denote the initial prices of the bond, put, and call respectively, then it follows from (1) and the no arbitrage condition that:

$$V_0[f] = [f(S_0) - f'(S_0)S_0]B_0 + f'(S_0)S_0 + \int_0^{S_0} f''(K)P_0(K)dK + \int_{S_0}^{\infty} f''(K)C_0(K)dK. \quad (2)$$

Appendix 2 shows that (2) directly implies that the initial value of an arbitrary payoff $f(\cdot)$ is:

$$V_0[f] = B_0 \int_0^{\infty} f(K)q(K)dK, \quad (3)$$

where the state pricing density $B_0q(K)$ may be recovered from option prices by the relation:

$$B_0q(K) = \begin{cases} \frac{\partial^2 P_0(K)}{\partial K^2} & \text{for } K \leq S_0; \\ \frac{\partial^2 C_0(K)}{\partial K^2} & \text{for } K > S_0. \end{cases} \quad (4)$$

The result (4) is of course well known from Breeden and Litzenberger [6] and is here seen as a simple consequence of the replication strategy (1).

2.2. The individual investor's problem

We suppose that there are n investors in the economy, indexed by $i = 1, \dots, n$. Each investor has an endowment of β_i

⁹ See Richards and Youn [42] for an accessible introduction to generalized functions such as delta functions.

¹⁰ We require that the payoff be twice differentiable and that the integrals in (2) not diverge.

shares, where $\sum_{i=1}^n \beta_i = 1$. For an initial stock price of S_0 , the initial value of the endowment is the investor's initial wealth $W_0^i \equiv \beta_i S_0$. Each investor's preferences are characterized by an increasing concave utility function U_i defined over their random terminal wealth W_i . Each investor's beliefs are characterized by a probability density function $p_i(S)$, defined on the entire positive half line with $p_i(S) > 0$ for $S > 0$. Each investor is assumed to maximize expected utility $\int_0^{\infty} U(W_i)p_i(S)dS$. Since all terminal wealth is consumed, W_i can be replaced by a function $f(S)$ relating terminal wealth to the terminal stock price S . The completeness of the market allows investors to maximize expected utility by choice of any function $f(S)$ that they can afford:

$$\max_{f(\cdot)} \int_0^{\infty} U_i[f(S)]p_i(S)dS. \quad (5)$$

The affordability of the payoff $f(\cdot)$ is captured by requiring that the initial value, $V_0[f]$, of the payoff, $f(S)$, must be less than or equal to the investor's initial wealth, W_0^i . From (3), the initial value of the payoff is simply its discounted expected value, where expectations are calculated using the risk-neutral density $q(S)$. Thus, the budget constraint is:

$$B_0 \int_0^{\infty} f(S)q(S)dS \leq W_0^i. \quad (6)$$

To solve this constrained optimization problem, we follow Brennan and Solanki [9] and consider the Lagrangian for this problem, given by:

$$\mathcal{L} = \int_0^{\infty} U_i[f_i(S)]p_i(S)dS - \lambda_i \left[\int_0^{\infty} f_i(S)B_0q(S)dS - W_0^i \right]. \quad (7)$$

Differentiating with respect to the payoff function $f(\cdot)$ and setting the result to zero yields the first order condition determining the optimal payoff function $\phi_i(S)$:

$$\frac{p_i(S)}{B_0q(S)} U_i'[\phi_i(S)] = \lambda_i. \quad (8)$$

Before analysing the structure of optimal payoffs, we offer two interpretations of the condition (8) defining them. The first is a Marshallian view, while the second is in terms of personalized prices.

2.2.1. A Marshallian view of payoff design. First note that all terms on the left hand side (LHS) of (8) are positive, and so is the Lagrange multiplier λ_i . From (7), λ_i is the increase in expected utility at the optimum if initial wealth is raised by a dollar. To interpret the LHS, note that the expected payoff from buying an Arrow–Debreu security paying off in state S is $p_i(S)$. The initial cost of this security is $B_0q(S)$. Consequently, the fraction on the LHS of (8) is the expected return from the Arrow–Debreu security. Since the investor is risk-averse,

the attractiveness of a state is measured by multiplying his expected return for that state by his marginal utility. This product gives the rate at which expected utility increases with each initial dollar spent on the state. Equation (8) indicates that the optimal payoff is chosen so that the extra expected utility received from the last dollar spent on each state is the same for each state. Intuitively, this must be the case as otherwise there would be an incentive to reallocate initial wealth from states where the rate of increase in expected utility is low to states where it is high.

2.2.2. Optimal payoffs and personalized prices. Another interpretation of the optimal payoff arises from multiplying (8) by $\frac{B_0 q(S)}{\lambda_i}$ and integrating over S :

$$\frac{1}{\lambda_i} \int_0^\infty p_i(S) U_i'[\phi_i(S)] dS = B_0 \int_0^\infty q(S) dS = B_0.$$

Substituting into (8) gives:

$$\pi_i(S) \equiv \frac{p_i(S) U_i'[\phi_i(S)]}{\int_0^\infty p_i(S) U_i'[\phi_i(S)] dS} = q(S). \quad (9)$$

The LHS can be interpreted as a probability density since it is positive and integrates to one. Thus (9) states that the optimal payoff is chosen so that each investor equates his personalized risk-adjusted probability density to the risk-neutral density. More generally, multiplying (8) by $\frac{B_0 q(S) f(S)}{\lambda_i}$ and integrating over S gives:

$$B_0 \int_0^\infty \pi_i(S) f(S) dS = B_0 \int_0^\infty f(S) q(S) dS \equiv V_0[f],$$

after substituting in (9). For payoffs such as $f(S) = S$ or $f(S) = (S - K)^+$, the RHS is the observable market price of the stock or option. Thus, each individual chooses his optimal payoff so that the personalized value of each asset equates with the market value.

2.2.3. General properties of optimal payoffs. The optimal payoff can be determined by solving (8) for $\phi_i(S)$:

$$\phi_i(S) = (U_i')^{-1} \left(\lambda_i B_0 \frac{q(S)}{p_i(S)} \right). \quad (10)$$

We next consider the general shape of this payoff under various conditions.

Equality of beliefs and prices. If an investor assigns the same probability to each state as does the risk-neutral density (i.e. $p_i(S) = q(S)$), then the optimal payoff is independent of the stock price, and so the optimal investment is to plunge all wealth into the riskless asset, i.e. $\phi_i(S) = W_0^i/B_0$. When $p = q$, the investor believes that the expected return on the stock is the riskless rate. Since the investor is risk-averse, he has no incentive in this case to hold stock. Furthermore, when $p = q$, the higher moment structures (e.g. variance, skewness and kurtosis) are identical, so there is no reason to buy other risky assets such as straddles, risk reversals or strangles.

Beliefs and prices differ in mean. If personal probabilities differ from the risk-neutral density (i.e. $p(S) \neq q(S)$), then we cannot have $p < q$ for all S or $p > q$ for all S since the densities both integrate to 1. It follows that when $p \neq q$, there are states in which $p > q$ and other states in which $p < q$. The payoff in any state for which $p > q$ is strictly greater than the payoff for any state in which $p < q$, since risk aversion forces the inverse of the marginal utility function $(U_i')^{-1}(\cdot)$ to be a decreasing function of its argument.

If an investor believes that the expected return on the stock is greater than the risk-free rate, then $\int S p_i(S) dS \geq \int S q(S) dS$, and so we would expect $p_i(S)$ to exceed $q(S)$ for most of the higher states. Since p cannot dominate q for all states, we would also expect p to be less than q for most of the lower states. This suggests an optimal payoff which is usually increasing with respect to S when the investor's expected excess return is positive. Conversely, if this expected excess return is negative, then p exceeds q for most of the lower states, and the optimal payoff is usually decreasing.

Beliefs and prices differ in volatility. Turning to the second moment, if the investor believes that volatility is higher than indicated by q , then the investor has $p > q$ for very high and very low states, with $p < q$ for intermediate states. Thus from (10), the corresponding payoff is U-shaped for most of its domain, which would require long positions in options. Conversely, if the investor believes volatility is lower than is implied by the market prices of options, then one would expect the optimal payoff to be an inverted U for most of the domain, although risk aversion would suggest that infinitely negative payoffs would generally be avoided.

Risk aversion and derivatives. Focusing on the effect of preferences on the optimal payoff, note that the greater the risk aversion, the more marginal utility falls with each dollar of terminal wealth, and the smaller is the required response of the optimal payoff to a deviation of the personal density from the risk-neutral density. Let $T_i[\phi_i(S)] \equiv -\frac{U_i''[\phi_i(S)]}{U_i'[\phi_i(S)]}$ denote the investor's risk tolerance and let $D_i(S) \equiv \ln \left(\frac{p_i(S)}{q(S)} \right)$ measure the deviation of the investor's beliefs from the market. Taking the logarithmic derivative of both sides of the first order condition (8) with respect to the stock price yields the decomposition in Leland [35] of an investor's optimal *exposure* into the product of his preferences and beliefs:

$$\phi_i'(S) = T_i[\phi_i(S)] D_i'(S). \quad (11)$$

Thus, when one terminal price is compared to an adjacent one, the investor increases his payoff at the higher price if his personal density grows faster than the risk-neutral density, and decreases it otherwise. However, the lower the risk tolerance, the smaller the required response of payoffs to deviations in growth rates of personal probabilities over risk-neutral probabilities.

Equation (11) is an ordinary differential equation (ODE) governing the optimal payoff $\phi_i(S)$. Under certain regularity

conditions on the risk tolerance, this ODE may be solved subject to the budget constraint:

$$\int_0^{\infty} \phi_i(S) B_0 q(S) dS = W_0^i. \quad (12)$$

The solution to this problem is given by (10), where the parameter λ_i is obtained by substituting the optimal payoff (10) in (12):

$$\int_0^{\infty} (U_i')^{-1} \left(\lambda_i B_0 \frac{q(S)}{p_i(S)} \right) B_0 q(S) dS = W_0^i. \quad (13)$$

In order to develop fully explicit solutions for optimal payoffs, we next restrict preferences.

3. Restrictions on preferences

As mentioned in the introduction, the purpose of this paper is to describe the derivatives positions that are optimal for investors. If derivatives are not held in our economy, then the investor confines his holdings to the bond and the stock and the optimal derivatives position is zero. Clearly, a necessary condition for the optimality of a zero derivatives position is that two-fund monetary separation hold¹¹. In turn, Cass and Stiglitz [11] have shown that a necessary condition for two-fund monetary separation to hold is that investors have linear risk tolerance (LRT):

$$T_i(W) \equiv -\frac{U_i'(W)}{U_i''(W)} = \tau_i + \gamma_i W. \quad (14)$$

The parameter γ_i is frequently called cautiousness in the literature¹² as it describes how risk tolerance changes with wealth. In this paper, we only consider utility functions with increasing risk tolerance, i.e. $\gamma_i \geq 0$. To avoid negative risk tolerance, the utility function is defined only for wealth levels $W_i \geq -\frac{\tau_i}{\gamma_i}$. For positive cautiousness, this lower bound is finite and as terminal wealth approaches it, the tolerance for risk approaches zero. Thus, LRT investors with positive cautiousness invest so as to create a floor of $-\frac{\tau_i}{\gamma_i}$ on final wealth. In order that this floor be attainable, we require $W_0^i \geq -B_0 \tau_i / \gamma_i$.

Solving the differential equation (14) for marginal¹³ utility $U_i'(W)$ gives:

$$U_i'(W) = \begin{cases} \rho_i (\tau_i + \gamma_i W)^{-1/\gamma_i} & \text{if } \gamma_i > 0; \\ \rho_i \exp\left(-\frac{W}{\tau_i}\right) & \text{if } \gamma_i = 0, \end{cases} \quad (15)$$

¹¹ Working in a multiasset setting with arbitrary concave utility functions, Ross [44] describes the distributional restrictions which are necessary and sufficient to generate two-fund monetary separation.

¹² See Huang and Litzenberger [30] p 134 for example.

¹³ Integrating (15) once implies that LRT utility functions are positive linear transformations of:

$$U_i(W) = \begin{cases} \frac{1}{\gamma_i - 1} (\tau_i + \gamma_i W)^{1-1/\gamma_i}, & \text{if } \gamma_i \neq 1, 0; \\ \ln(\tau_i + W), & \text{if } \gamma_i = 1; \\ -\tau_i \exp\left(-\frac{W}{\tau_i}\right), & \text{if } \gamma_i = 0, \end{cases}$$

for $W \geq -\frac{\tau_i}{\gamma_i}$.

where the arbitrary positive constant ρ_i can be interpreted as the individual's rate of time preference. Substituting the inverse of this function in (13), solving for λ , and substituting in (10) implies¹⁴ that LRT investors prefer a payoff of the form:

$$\phi_i(S) = \begin{cases} -\frac{\tau_i}{\gamma_i} + \frac{W_0^i + B_0 \tau_i / \gamma_i}{V_0 [e^{\gamma_i D_i}]} e^{\gamma_i D_i(S)}, & \text{if } \gamma_i > 0; \\ \frac{W_0^i - \tau_i V_0 [D_i]}{B_0} + \tau_i D_i(S), & \text{if } \gamma_i = 0, \end{cases} \quad (16)$$

where $V_0 [e^{\gamma_i D_i}]$ is the initial value of the payoff $e^{\gamma_i D_i(S)} = (p_i(S)/q(S))^{\gamma_i}$, and $V_0 [D_i]$ is the value of the payoff $D_i(S) \equiv \ln(p_i(S)/q(S))$:

$$V_0 [e^{\gamma_i D_i}] = B_0 \int_0^{\infty} [p_i(S)/q(S)]^{\gamma_i} q(S) dS$$

$$V_0 [D_i] = B_0 \int_0^{\infty} \ln[p_i(S)/q(S)] q(S) dS.$$

The optimal payoffs in (16) separate into the riskless asset and a single customized derivative. Under homogeneous beliefs and identical cautiousness, the derivative fund is the same across investors, so that two-fund monetary separation obtains. Cass and Stiglitz [11] have shown that these restrictive conditions are also necessary. In what follows, we will allow for heterogeneous beliefs and cautiousness parameters in order to explore how differences in these attributes affect the optimal customized payoff. As indicated in (16), the nature of this payoff depends on whether the cautiousness parameter γ_i is positive or zero. We discuss each case in turn.

3.1. Positive cautiousness

When $\gamma_i > 0$, (16) implies that the optimal payoff in each state is linear in a power of the expected return from that state $p_i(S)/q(S)$:

$$\phi_i(S) = -\frac{\tau_i}{\gamma_i} + \frac{W_0^i + B_0 \tau_i / \gamma_i}{V_0 [(p_i/q)^{\gamma_i}]} \left(\frac{p_i(S)}{q(S)} \right)^{\gamma_i}. \quad (17)$$

The position in the riskless fund is taken to ensure that the floor of $-\frac{\tau_i}{\gamma_i}$ is preserved. Since $\gamma_i > 0$, whether the investor is long or short in the fund depends on the sign of τ_i . When $\tau_i < 0$, the investor is long in the riskless fund and the remaining wealth is used to buy the risky fund with payoff $\left(\frac{p_i(S)}{q(S)}\right)^{\gamma_i}$. When $\tau_i = 0$, then from (14), the investor has proportional risk tolerance and invests all his assets in the risky fund. When $\tau_i > 0$, the investor is short in the riskless fund and invests the proceeds and his initial wealth in the risky fund. In this last case, the lower the cautiousness, the more the investor shorts in the riskless fund.

In any case, all of the excess of initial wealth over the present value of this floor is invested in a long position in the risky fund with customized payoff $\left(\frac{p_i(S)}{q(S)}\right)^{\gamma_i}$. This payoff can be synthesized with investor-specific positions in bonds, stocks and options. Since the risky fund has a non-negative

¹⁴ The solutions can also be obtained by substituting (14) in the ODE (11) and solving this linear ODE subject to (12).

payoff, options must be used if the implicit position in either the bond or the stock is negative. Thus, if the investor believes the expected return on the stock is very positive, then the investor might borrow to obtain leverage, but in this case the investor will also buy puts to guarantee a non-negative payoff. Similarly, if the investor believes the expected return on the stock is below the risk-free rate, then the investor would short in the stock, but in this case the investor would also buy calls to prevent the payoff from going negative at high prices. Turning to the second moment, if the investor believes the volatility of the stock to be substantially lower than is implied by the market prices of options, then one would expect the risky fund's payoff to be a bell-shaped curve, since the payoffs cannot go negative.

In summary, an LRT investor with positive cautiousness first invests in the riskless fund to establish the floor, and then invests the remainder in a limited liability risky asset. The next subsection shows that an LRT investor with zero cautiousness behaves quite differently.

3.2. Zero cautiousness

Consider rewriting the optimal payoff (17) with positive cautiousness as:

$$\phi_i(S) = \frac{W_0^i}{V_0 [(p_i/q)^\gamma]} \left(\frac{p_i(S)}{q(S)} \right)^\gamma + \frac{\tau_i}{\gamma_i} \left[\frac{B_0}{V_0 [(p_i/q)^\gamma]} \left(\frac{p_i(S)}{q(S)} \right)^\gamma - 1 \right].$$

If we require $\tau_i > 0$ and let γ_i approach zero, then L'Hôpital's rule implies that the payoff approaches¹⁵ the zero-cautiousness payoff:

$$\phi_i(S) = \frac{W_0^i - \tau_i V_0 [D_i]}{B_0} + \tau_i D_i(S), \quad (18)$$

where recall $V_0 [D_i]$ is the initial cost of the customized payoff $D_i(S) \equiv \ln \left(\frac{p_i(S)}{q(S)} \right)$. Given the availability of this payoff, the zero cautiousness investor buys τ_i units of this risky fund. Substituting $\gamma_i = 0$ in (14) implies that the number of units that the investor buys is fixed at his risk tolerance, in contrast to the case with positive cautiousness. The zero-cautiousness investor then invests all remaining wealth in the riskless fund. In further contrast to the case with positive cautiousness, the risky fund can have an arbitrarily large negative payoff. Since risk tolerance is independent of final wealth, low wealth realizations do not induce the zero-cautiousness investor to place a floor on final wealth.

4. Optimal investment in a partial equilibrium

Thus far, the optimal payoff has been expressed in terms of the investor's personal probability density p_i , the investor's utility function U_i and in terms of the risk-neutral density q . In general, the latter is determined by equilibrating investor demand for assets with the fixed supply. In the absence of an options market, the investor would need to know the tastes and

beliefs of all other participants in order to determine his optimal position. Thus, a fundamental role of the options markets in our setting is to summarize the tastes and beliefs of all other participants, so that it is not necessary for a given investor to solve a general equilibrium problem in order to determine his optimal position. In principle, each investor can back out the risk-neutral density from option prices and combine this information with his personal information to solve for his optimal position.

The next section demonstrates a general equilibrium setting in which all asset prices are determined endogenously. In this section, we will take bond, stock and option prices as given, and demonstrate an investor's optimal position in our partial equilibrium framework. For tractable results in closed form, we assume log-normality for both the personal densities and the risk-neutral density. The existence of general equilibrium models with this property are discussed in Rubinstein [45], Breeden and Litzenberger [6], Brennan [7], Stapleton and Subrahmanyam [48], Bick [4, 5] and He and Leland [29]. A sufficient condition for log-normality in both the personal densities and the risk-neutral density is constant proportional risk aversion (CPRA) on the part of all investors. To our knowledge, necessary conditions for log-normality in both densities have not been established. Thus, we will first derive results for arbitrary preferences and then specialize them to linear risk tolerance, which includes CPRA as a special case.

The partial equilibrium Black–Scholes model also has the property that personal densities and risk-neutral densities are both log-normal. In this section, one can think of the Black–Scholes model as holding, except that the investor under consideration is unable to trade continuously. This investor is assumed to invest so as to maximize the expected utility of wealth at the next trading opportunity. The Rubinstein and Black–Scholes models assume that investors have the same volatility, which implies that the risk-neutral and personal volatilities are equal¹⁶.

It is important from both an empirical and a theoretical perspective to consider the case when personal and risk-neutral volatilities differ. On the empirical side, there is considerable evidence that implied volatilities are substantially higher than historical volatilities. This phenomenon is most likely due to risk aversion, which causes extreme states to be priced higher than their relative frequency of occurrence. Since one would expect rational investors to account for this phenomenon in formulating their personal volatilities, one would expect that personal volatilities are closer to historical volatilities than risk-neutral volatilities. A second reason for considering unequal personal and risk-neutral volatilities is purely theoretical. The equality of the two volatilities in the Black–Scholes model is a consequence of Girsanov's theorem, which only applies to diffusion processes. For continuous time jump processes or for all discrete time processes (for example the binomial process), the two volatilities generally differ, except in the diffusion limit.

¹⁶ In the Black–Scholes model, if an investor disagreed with the risk-neutral volatility, then he would take an arbitrarily large position, stressing the price-taking assumption.

¹⁵ Thus, the payoffs in (16) are continuous in γ_i for all $\gamma_i \geq 0$.

A second justification for studying the log-normal distribution is that as a two-parameter distribution, it allows us to address the effect of differences of opinion on the mean and variance of return. As in the last two sections, we first develop results for arbitrary preferences and then specialize to LRT preferences.

4.1. Log-normal beliefs and arbitrary preferences

This subsection specializes the results on optimal positioning to the case where preferences are arbitrary, but where the personal and risk-neutral distributions are both log-normal. Letting $x \equiv \ln(S/S_0)$ denote the log price relative, the log-normal density is given by:

$$\ell(S; \mu, v) = \frac{1}{\sqrt{2\pi}vS} \exp \left\{ -\frac{1}{2} \left[\frac{x - (\mu - v^2/2)}{v} \right]^2 \right\}, \quad (19)$$

where μ and v are the mean and volatility of return¹⁷ over our single period. We now suppose that the investor's personal density is $\ell(S; \mu_i, v_i)$, while the risk-neutral density is $\ell(S; r, \sigma)$, where r is the assumed constant riskless rate and σ is the implied volatility.

Under log-normal densities, the deviation of the investor's probability density from that of the market is quadratic in $x \equiv \ln(S/S_0)$:

$$D_i(S) \equiv \ln \left(\frac{p_i(S)}{q(S)} \right) = A_i + B_i x - C_i \frac{x^2}{2}, \quad (20)$$

where:

$$\begin{aligned} A_i &= \ln \left(\frac{\sigma}{v_i} \right) + \frac{(r - \sigma^2/2)^2}{2\sigma^2} - \frac{(\mu_i - v_i^2/2)^2}{2v_i^2} \\ B_i &= \frac{\mu_i}{v_i^2} - \frac{r}{\sigma^2} \\ C_i &= \frac{1}{v_i^2} - \frac{1}{\sigma^2}. \end{aligned} \quad (21)$$

To focus on the effects of an investor differing from the market only in the mean parameter, we temporarily assume that the personal and risk-neutral volatilities are the same. In a diffusion setting, this is a consequence of Girsanov's theorem and this case was also considered in the context of a static model under homogeneous beliefs by Brennan and Solanki [9]. If $v_i = \sigma$, then $C_i = 0$, so (20) simplifies to:

$$D_i(S) = A_i + S_i x, \quad (22)$$

where $A_i \equiv \frac{1}{2} \left(\mu_i - r - \frac{\mu_i^2 - r^2}{\sigma^2} \right)$ and $S_i \equiv \frac{\mu_i - r}{\sigma^2}$ is defined as the investor's Sharpe ratio¹⁸.

¹⁷ Note that μ is the log of the expected value of the periodically compounded gross return, $\mu = \ln E \frac{S}{S_0}$, while v is the standard deviation of the continuously compounded return, $v = \text{Std} \ln(S/S_0)$.

¹⁸ The Sharpe ratio is usually defined as the expected excess return relative to the standard deviation of return. We define it relative to the variance for ease of reference.

Recall the multiplicative decomposition (11) of the investor's exposure $\phi'_i(S)$ into his risk tolerance $T_i[\phi_i(S)]$ and the deviation of his beliefs from the market:

$$\phi'_i(S) = T_i[\phi_i(S)] D'_i(S). \quad (23)$$

Differentiating (22) and substituting into (23) gives an exposure of:

$$\phi'_i(S) = \frac{T_i[\phi_i(S)]}{S} S_i.$$

Since risk tolerance is positive, the sign of the investor's exposure matches the sign of the Sharpe ratio S_i . Thus, under log-normal beliefs and risk-neutral densities with equal volatilities, the investor's optimal payoff is monotonically increasing if his mean μ_i is higher than the market's, is flat if means match, and is decreasing otherwise. Thus, an investor with an expected return above the risk-free rate should have an increasing payoff, even if he believes this premium is insufficient to cover the risk borne.

4.2. Log-normal beliefs and linear risk tolerance

In order to completely determine the optimal payoff, this subsection assumes both log-normal beliefs and LRT preferences. Substituting (20) in (16) and simplifying implies that the optimal payoff is:

$$\phi_i(S) = \begin{cases} -\frac{\tau_i}{\gamma_i} + \frac{W_0^i + B_0 \tau_i / \gamma_i}{V_0 [e^{\gamma_i (B_i x - C_i x^2 / 2)}]} e^{\gamma_i (B_i x - C_i x^2 / 2)}, & \text{if } \gamma_i > 0; \\ \frac{W_0^i - \tau_i (B_i V_0 [x] - C_i V_0 [x^2 / 2])}{B_0} + \tau_i \left(B_i x - C_i \frac{x^2}{2} \right), & \text{if } \gamma_i = 0. \end{cases} \quad (24)$$

4.2.1. Positive cautiousness and homogeneous volatility

To analyse (24), we temporarily assume that cautiousness is positive and that the risk-neutral and personal volatilities are equal. Setting $C_i = 0$ and $x = \ln(S/S_0)$ in the $\gamma_i > 0$ case of (24) gives:

$$\phi_i(S) = -\frac{\tau_i}{\gamma_i} + \frac{W_0^i + B_0 \tau_i / \gamma_i}{V_0 [S^{\gamma_i S_i}]} S^{\gamma_i S_i}, \quad (25)$$

where for the log-normal risk-neutral density, $V_0 [S^{\gamma_i S_i}] = S_0^{\gamma_i S_i} B_0 e^{\gamma_i S_i (r - \sigma^2/2) + \gamma_i^2 S_i^2 \sigma^2 / 2}$. Thus, for $\gamma_i > 0$, the optimal payoff¹⁹ is the sum of the riskless fund and a risky fund whose payoff takes the simple form of the stock price raised to a power. The power is the product of the investor's cautiousness γ_i and his Sharpe ratio $S_i = (\mu_i - r)/\sigma^2$. As shown in the last subsection, if $\mu_i > r$, then the optimal payoff is increasing in the stock price, while if $\mu_i < r$, it is decreasing. If $\mu_i > r + \gamma \sigma^2$, the optimal payoff is convex in the stock price, while if μ_i is below this level, it is concave. For the optimal position to not require derivatives, it must be the case

¹⁹ For proportional risk tolerance, the optimality of this payoff in the continuous-time context is discussed in Cox and Huang [15].

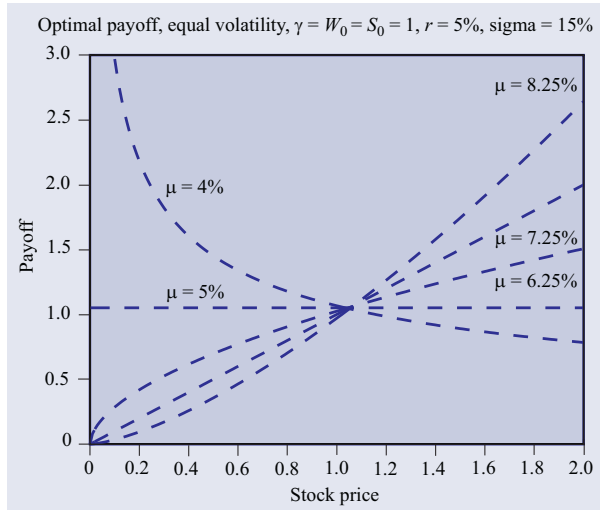


Figure 1. Optimal payoffs when volatilities agree.

that $\mu_i = r + \gamma\sigma^2$, i.e. that the investor's independently determined mean return happens to agree with the cost of financing the stock and bearing the risk. Although this can occur for some individuals, most investors will optimally hold derivative positions.

Figure 1 shows the effect of varying the expected return on the optimal payoff for an investor with proportional risk tolerance (i.e. $\tau_i = 0$, $\gamma_i > 0$).

When $\mu_i > r + \gamma_i\sigma^2$, the investor takes a leveraged position in stock. He also buys puts to protect the downside and buys calls to convexify the upside. As expected, return is lowered, the investor borrows less, buys less stock and reduces his option purchases. When $\mu_i = r + \gamma_i\sigma^2$, the investor holds only stock. As expected, return is reduced below this point, the investor starts to be long in the riskless asset but continues to maintain a long position in stock as well. The investor now judiciously sells puts and calls. When $\mu_i = r$, the investor holds only the riskless asset. Finally, when $\mu_i < r$, the investor shorts in the stock and buys puts to convexify the effect of stock price declines, while buying calls to protect against stock price rises.

4.2.2. Zero cautiousness and homogeneous volatility.

Setting $C_i = 0$ and $x = \ln(S/S_0)$ in the $\gamma_i = 0$ case of (24) gives the optimal payoff for an investor with constant risk tolerance τ_i :

$$\phi_i(S) = \frac{W_0^i - \tau_i S_i V_0[\ln(S/S_0)]}{B_0} + \tau_i S_i \ln(S/S_0), \quad (26)$$

where for log-normal risk-neutral density $V_0[\ln(S/S_0)] = B_0(r - \sigma^2/2)$. Thus, for constant risk tolerance, the investor either has no exposure to the risky asset (if $S_i = 0$) or else uses derivatives in conjunction with the underlying stock. If $\mu_i > r$, then the optimal payoff is increasing and concave. It is also unbounded, tending to $-\infty$ as S tends to 0, and tending to ∞ as S tends to ∞ . The position is a leveraged position in stock combined with aggressive selling of out-of-the-money

options. If $\mu_i < r$, then the optimal position is decreasing, convex, and unbounded. This position is created by shorting the stock with the proceeds used to buy bonds and out-of-the-money options. The puts are held to steepen the payoff as the stock falls and the calls are held to dampen the decline as the stock rises.

In summary, for non-negative cautiousness, stocks and derivatives are not held if the investor is in complete agreement with the risk-neutral density or else is infinitely risk averse. In addition, derivatives are not held if the stock's independently determined expected return lines up with the cost of financing the stock purchase and bearing the risk. In all other cases, options are used as part of the optimal payoff even though the investor's personal volatility agrees with that implied. The positions taken in options will be even larger in absolute terms when the investor's personal volatility differs from implied volatility, as shown next.

4.2.3. Positive cautiousness and unequal volatility.

Recall the optimal payoff (24) for an investor with positive cautiousness γ_i :

$$\phi_i(S) = -\frac{\tau_i}{\gamma_i} + \frac{W_0^i + B_0\tau_i/\gamma_i}{V_0[e^{\gamma_i(B_i x - C_i x^2/2)}]} e^{\gamma_i(B_i x - C_i x^2/2)}, \quad (27)$$

where when the risk-neutral density is log-normal:

$$V_0[e^{\gamma_i(B_i x - C_i x^2/2)}] = \frac{v}{\sigma} \exp\left\{-\frac{1}{2}\left[\frac{(r - \sigma^2/2)^2}{\sigma^2} - \left(\gamma_i B_i + \frac{(r - \sigma^2/2)}{\sigma^2}\right)^2 v^2\right]\right\},$$

with $\frac{1}{v^2} \equiv \gamma_i \frac{1}{v_i^2} + (1 - \gamma_i) \frac{1}{\sigma^2}$.

When an investor's personal volatility v_i exceeds the risk-neutral volatility σ , then the investor's precision is below that of the market, and thus the coefficient C_i defined in (21) will be negative. It follows from differentiating (27) twice that the optimal payoff will tend to be convex. Figure 2 shows the effect of lowering the personal volatility on the optimal payoff of an investor with proportional risk tolerance ($\tau_i = 0$, $\gamma_i > 0$) and zero Sharpe ratio.

When $v_i > \sigma$, the optimal payoff is U-shaped and amounts to overlaying straddles or strangles onto a riskless bond position. As volatility is lowered, the investor cuts back his option purchases and invests more in the riskless bond, until when $v_i = \sigma$, the investor holds only the riskless asset. As volatility is lowered further, the payoff is similar in form to a log-normal density function. Thus the investor now sells near-the-money options to be consistent with the volatility view, but buys further out-of-the-money options to enforce the floor.

4.2.4. Zero cautiousness and unequal volatility.

Recall the optimal payoff (24) for an investor with constant risk tolerance τ_i :

$$\begin{aligned} \phi_i(S) &= \frac{W_0^i - \tau_i(B_i V_0[x] - C_i V_0[x^2/2])}{B_0} + \tau_i \left(B_i x - C_i \frac{x^2}{2}\right), \\ & \quad (28) \end{aligned}$$

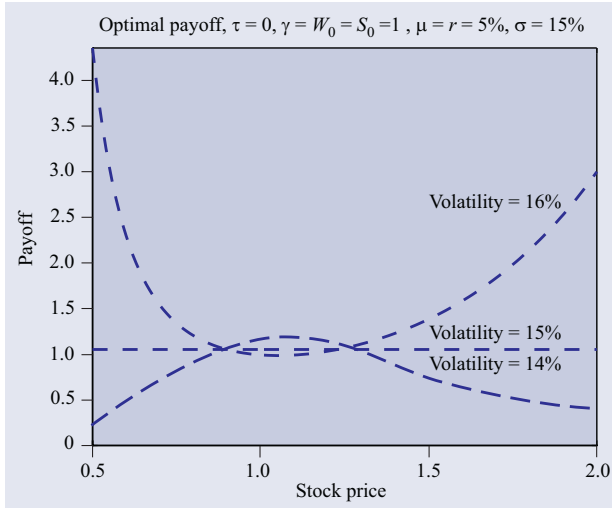


Figure 2. Optimal payoff with zero Sharpe ratio.

where for log-normal risk-neutral density, $V_0[x] = B_0(r - \sigma^2/2)$ and $V_0[x^2/2] = \frac{B_0}{2} [\sigma^2 + (r - \sigma^2/2)^2]$. Given the existence of a contract paying the log price relative $x = \ln(S/S_0)$ and another paying its square, the zero cautiousness investor confines his investments to these two contracts and the riskless asset. In this economy, the log contract is used to speculate on expected return²⁰. The investor is long in the log contract if $\mu_i > r \frac{v_i^2}{\sigma^2}$, is out of the contract if $\mu_i = r \frac{v_i^2}{\sigma^2}$, and is short otherwise. The greater the investor's expected return μ_i , the larger B_i is, and thus the larger the position in the log contract. In contrast, the log squared contract is a vehicle for speculating on volatility. The investor is long in this contract if $v_i > \sigma$, is out of the contract if $v_i = \sigma$, and is short otherwise. The larger the investor's volatility v_i , the smaller the coefficient C_i indicating the difference in precisions, and thus the larger the position in the log squared contract. Furthermore, the larger v_i is, the lower B_i is, and thus the lower the position in the concave log contract. Thus, the larger the personal volatility, the more convex the payoff.

Determining the conditions under which the investor is uniformly long or uniformly short options is complicated by the concavity of the log contract. However, if $\mu_i = r \frac{v_i^2}{\sigma^2}$, then $B_i = 0$ and the curvature (gamma) of the total position has the same sign as the excess of personal volatility over implied. Consequently, in contrast to the positive-cautiousness case, the zero-cautiousness investor who thinks volatility is lower than implied sells options at all strikes.

5. Optimal positioning in general equilibrium

Our results on optimal positioning thus far are mainly affected by the attributes of the individual investor under consideration. The attributes of other investors were summarized by the

²⁰ Neuberger [39] and Dupire [19] stress the role of the log contract in allowing investors to speculate on variance. However, their analysis requires that the investor be able to trade continuously.

risk-neutral density, which aggregates their preferences and beliefs. As a practical matter, the empirical estimation of this risk-neutral density from option prices is probably a robust procedure. However, our results thus far are theoretically incomplete, as it is not clear whether, in equilibrium, someone is available to take the other side of the investor's optimal derivatives position. To address this issue, we consider the solution of a general equilibrium model.

Consider an economy in which multiple investors simultaneously optimize their holdings. We no longer take option prices as given, and so the form of the risk-neutral density must be solved for endogeneously. We require that the risk-neutral density must price the bond:

$$B_0 \int_0^\infty 1q(S)dS = B_0, \tag{29}$$

or equivalently, that the risk-neutral density $q(\cdot)$ integrates to one. We further assume that bonds and options are in zero net supply and thus in the aggregate, it is just the stock that is held:

$$\sum_{i=1}^n \phi_i(S) = S, \tag{30}$$

which implies that the sum of the exposures is unity:

$$\sum_{i=1}^n \phi'_i(S) = 1. \tag{31}$$

Finally, the above equations imply that the the risk-neutral expected return on the stock is the riskless rate. To see this, recall that each investor is endowed with β_i shares, where $\sum_{i=1}^n \beta_i = 1$. Since $W_0^i \equiv \beta_i S_0$, initial wealths sum to the initial stock price:

$$\sum_{i=1}^n W_0^i = S_0.$$

Substituting in the budget constraint (12) and interchanging summation and integration implies:

$$\int_0^\infty B_0 \sum_{i=1}^n \phi_i(S)q(S)dS = S_0.$$

Finally, substituting in (30) gives the desired result:

$$B_0 \int_0^\infty Sq(S)dS = S_0. \tag{32}$$

5.1. The risk-neutral density in equilibrium

In any equilibrium model, only relative prices are determined. Thus, we will take S_0 as given, and solve for the risk-neutral density $q(S)$, the bond price B_0 and then determine optimal payoffs $\phi(S)$ in terms of S_0 . To obtain an expression for the risk-neutral density in general equilibrium, recall the

multiplicative decomposition (23) of exposures into beliefs and preferences:

$$\phi'_i(S) = T_i[\phi_i(S)]D'(S) = T_i[\phi_i(S)]\frac{d}{dS} \ln[p_i(S)/q(S)]. \quad (33)$$

Summing over i implies:

$$\sum_{i=1}^n \phi'_i(S) = \sum_{i=1}^n T_i[\phi_i(S)] \left[\frac{d}{dS} \ln p_i(S) - \frac{d}{dS} \ln q(S) \right] = 1, \quad (34)$$

from (31). Solving for $q(S)$ gives the desired expression:

$$q(S) = q(0) \exp \left\{ - \int_0^S \frac{1}{T(Z)} dZ \right\} \times \exp \left\{ \int_0^S \sum_{i=1}^n \frac{T_i[\phi_i(Z)] p'_i(Z)}{T(Z) p_i(Z)} dZ \right\}, \quad (35)$$

where $T(S) \equiv \sum_{i=1}^n T_i[\phi_i(S)]$ is the total risk tolerance in state S . Since the optimal payoff ϕ_i depends on q , this is not an explicit expression for the risk-neutral density. Furthermore, under heterogeneous beliefs, (35) indicates that the equilibrium risk-neutral density is the product of a factor reflecting total risk tolerance, i.e. $\exp \left\{ - \int_0^S \frac{1}{T(Z)} dZ \right\}$, and a factor reflecting the personal beliefs, which we term the market view. The greater the risk tolerance of a given investor, the more his probability density gets reflected in the market view.

Under homogeneous beliefs, (35) simplifies into:

$$q(S) = q(0) \exp \left\{ - \int_0^S \frac{1}{T(Z)} dZ \right\} p(S). \quad (36)$$

The first factor is a positive declining function of S which changes the mean in the market view to the riskless rate, and may add negative skewness. For example, if $p(S)$ is a normal density and aggregate risk tolerance is constant, then $q(S)$ is also normal but with a shifted mean. However, if $p(S)$ is log-normal and risk tolerance is constant, then $q(S)$ is not in the log-normal family and is skewed to the left with the density having a fatter left tail. For linear aggregate risk tolerance, similar results hold. If aggregate risk tolerance is infinite or equivalently, if there exist individuals with zero risk aversion, then $q(S) = p(S)$. It follows that the disparity between the risk-neutral density and the density describing homogeneous beliefs is a consequence of universal risk aversion and the requirement that the risky stock be held in equilibrium.

5.2. The optimal payoffs in general equilibrium

To obtain the optimal payoffs in our general equilibrium, we substitute (36) into (33) to express the optimal exposure in terms of preferences and beliefs:

$$\phi'_i(S) = \frac{T_i[\phi_i(S)]}{T(S)} + T_i[\phi_i(S)] \times \left[\frac{d \ln p_i(S)}{dS} - \sum_{i=1}^n \frac{T_i[\phi_i(S)]}{T(S)} \frac{d \ln p_i(S)}{dS} \right]. \quad (37)$$

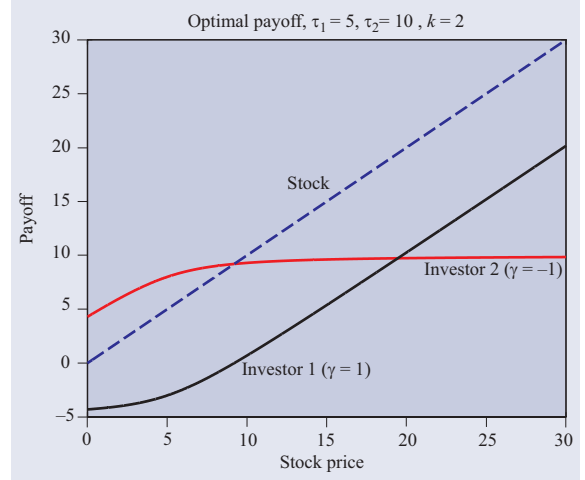


Figure 3. Optimal payoff under homogeneous beliefs and opposite cautiousness.

The first term reflects the investor's risk tolerance relative to the population total. The second term is a composite of the investor's risk tolerance and the extent to which the investor's beliefs differ from a risk tolerance weighted average of the beliefs of other investors in the economy.

If investors have homogeneous beliefs (i.e. $p_i(S) = p(S) \forall i$), then the second term vanishes:

$$\phi'_i(S) = \frac{T_i[\phi_i(S)]}{T(S)}. \quad (38)$$

Since the right side is positive, homogeneous beliefs imply that all investors must have an increasing payoff. The greater the investor's risk tolerance relative to the total, the greater the exposure of the investor's position. The next two subsections show that LRT investors with homogeneous beliefs and identical cautiousness will not hold derivatives. Appendix 3 shows that derivatives are held in an economy with two LRT investors with homogeneous beliefs, but *opposite* cautiousness. In particular, if $T_1[W_1] = \tau_1 + \gamma W_1$ and $T_2[W_2] = \tau_2 - \gamma W_2$, then:

$$\phi_1(S) = \frac{S}{2} - \frac{\tau}{2\gamma} + \sqrt{\left(\frac{S}{2} + \frac{\tau_2 - \tau_1}{2\gamma} \right)^2 + k^2}$$

$$\phi_2(S) = \frac{S}{2} + \frac{\tau}{2\gamma} - \sqrt{\left(\frac{S}{2} + \frac{\tau_2 - \tau_1}{2\gamma} \right)^2 + k^2}, \quad (39)$$

where $\tau \equiv \tau_1 + \tau_2$ and k is an arbitrary constant.

Thus, in this simple economy, a three-fund separation occurs in which each investor holds equal positions in the stock and offsetting positions in the bond and the derivative²¹. The optimal derivative security is the square root of the sum of a positive constant and a squared linear position in the stock. Figure 3 plots the optimal payoffs.

²¹ Appendix 3 also shows that if $k = 0$, then the investors no longer hold derivatives.

Although our results pertain to only two investors, it is worth quoting from Dumas [18]²²:

The two-investor equilibrium is as basic to financial economics as is the two-body problem in mechanics.

In order to obtain explicit solutions for the optimal payoff in an n investor economy, we next restrict preferences. In particular, the next subsection assumes generalized logarithmic utility²³, $U_i(W) = \ln(\tau_i + W)$, while the following subsection considers negative exponential utility $U_i(W) = -\tau_i \exp\left(-\frac{W}{\tau_i}\right)$. In both cases, we examine the cross section of investor beliefs and preferences in order to explain the positions held.

5.3. Generalized logarithmic utility

Setting $\gamma_i = 1$ in (17), the optimal payoff becomes:

$$\phi_i(S) = -\tau_i + \frac{R_0^i p_i(S)}{B_0 q(S)}, \quad (40)$$

since $V_0[p_i/q] = B_0 \int_0^\infty \frac{p_i(S)}{q(S)} q(S) dS = B_0$ and since $R_0^i \equiv W_0^i + B_0 \tau_i \geq 0$ is defined as the risk capital of investor i . Note that the greater τ_i is, the greater the risk tolerance $T_i[W_i] = \tau_i + W_i$ and the greater the risk capital R_0^i . Summing (40) over investors and invoking the market clearing condition (30) gives:

$$S = -\tau + \sum_{i=1}^n \frac{R_0^i p_i(S)}{B_0 q(S)}, \quad (41)$$

where $\tau \equiv \sum_{i=1}^n \tau_i$. Solving (41) for the risk-neutral density gives:

$$q(S) = \sum_{i=1}^n \frac{R_0^i p_i(S)}{B_0 S + \tau}. \quad (42)$$

In order that $q(S)$ be non-negative for all non-negative S , we require $\tau \geq 0$, i.e. that the aggregate floor $-\tau$ cannot be positive. Thus, positive floors on the part of some must be compensated for by negative floors on the part of others. From (42), the greater the risk capital R_0^i of an investor, the more impact his beliefs have on the risk-neutral density.

Substituting (42) in (40) gives the optimal payoff:

$$\phi_i(S) = -\tau_i + \frac{R_0^i p_i(S) \tau}{\sum_{i=1}^n R_0^i p_i(S)} + \frac{R_0^i p_i(S) S}{\sum_{i=1}^n R_0^i p_i(S)}. \quad (43)$$

Thus, each investor first establishes a floor at $-\tau_i$ and then invests all remaining wealth in two derivative funds. The holdings in the riskless fund and the first customized derivative sum to zero across investors, while the holdings in the second customized derivative sum to the stock price. The higher τ_i is, the higher the investor's risk tolerance and risk capital, and

²² In a highly original paper, Dumas [18] numerically solves for an equilibrium without derivatives in an intertemporal setting with two investors with different utility functions.

²³ See Rubinstein [46] for further motivation for the use of this preference structure.

the larger his position in each customized derivative. Note that each customized derivative is a limited liability claim. The first customized derivative is also bounded above by the absolute value of the aggregate floor τ , while the second customized derivative is bounded above by the stock price. If the two customized payoffs are made available to investors, then no one holds the stock directly, although holdings must sum to the stock.

5.3.1. Generalized logarithmic utility and homogeneous beliefs. Under generalized log utility, derivatives are not held if beliefs are homogeneous. In this case, payoffs simplify to:

$$\phi_i(S) = -\tau_i + \frac{R_0^i \tau}{\sum_{i=1}^n R_0^i} + \frac{R_0^i S}{\sum_{i=1}^n R_0^i}. \quad (44)$$

Thus, each investor holds the riskless asset and a long position in the stock. The higher is τ_i or W_0^i , the higher the investor's risk tolerance and risk capital, and the larger the position in stock. Thus, for generalized log utility investors²⁴ with homogeneous beliefs, differences in risk tolerance do not induce demand for derivatives, but instead only affect the division between the riskless asset and the stock.

5.3.2. Generalized logarithmic utility and derivative fund theorems. Under certain conditions, the customized derivatives optimal for a given investor can be decomposed into a linear combination of payoffs of universal interest. Under heterogeneous beliefs, an m -fund separation arises if each investor's density can be represented as:

$$p_i(S) = p(S) \sum_{k=1}^m c_{ik} f_k(S), \quad i = 1, \dots, n, \quad (45)$$

where $p(S)$ is the unknown true density and $\{f_k(S), k = 1, \dots, m\}$ is a collection of basis functions. In words, each investor's density differs from the true density by a multiplicative error, which can be represented by a finite number of basis functions. When (45) holds, we have:

$$R_0^i p_i(S) = p(S) \sum_{k=1}^m R_0^i c_{ik} f_k(S),$$

$$\sum_{i=1}^n R_0^i p_i(S) = p(S) \sum_{k=1}^m \theta_k f_k(S),$$

where $\theta_k \equiv \sum_{i=1}^n R_0^i c_{ik}$. Substituting into (43) gives an optimal payoff of:

$$\begin{aligned} \phi_i(S) = & -\tau_i + \sum_{k=1}^m R_0^i c_{ik} \frac{f_k(S) \tau}{\sum_{k=1}^m \theta_k f_k(S)} \\ & + \sum_{k=1}^m R_0^i c_{ik} \frac{f_k(S) S}{\sum_{k=1}^m \theta_k f_k(S)}. \end{aligned} \quad (46)$$

Thus, each investor's holdings separate into $2m + 1$ funds. The first fund is the riskless fund, which is used to establish

²⁴ It can be shown that this result generalizes to the case when investors have identical cautiousness and beliefs.

the floor of $-\tau_i$. Each investor holds $R_0^i c_{ik}$ units of each of the $2m$ derivative funds, where the funds have a payoff of $\left\{ \frac{(\tau_i)^{1-l} S^l f_k(S)}{\sum_{k=1}^m \theta_k f_k(S)}, k = 1, \dots, m, l = 0, 1 \right\}$. No one holds the stock individually, although the collective holdings sum to the stock.

5.3.3. Log-normal beliefs and zero aggregate floor Under further restrictions on preferences and beliefs, we can obtain explicit formulae for the risk-neutral density, for the bond price, and for the optimal payoffs in equilibrium. When each generalized log utility investor has log-normal beliefs, then the density is:

$$\ell(S; \mu_i, v_i) = \frac{1}{\sqrt{2\pi} v_i S} \exp \left\{ -\frac{1}{2} \left[\frac{\ln(S/S_0) - (\mu_i - v_i^2/2)}{v_i} \right]^2 \right\} \quad (47)$$

Substituting in (42) and setting the aggregate floor of $-\tau$ to zero gives a risk-neutral density of:

$$q(S) = \sum_{i=1}^n \frac{R_0^i \ell(S; \mu_i, v_i)}{B_0 S}. \quad (48)$$

Multiplying (48) by $S B_0$ and integrating over S implies that capital at risk aggregates to the initial stock price:

$$\sum_{i=1}^n R_0^i = S_0, \quad (49)$$

from²⁵ (32). Completing the square in the log-normal in (48) implies that the risk-neutral density can also be written as:

$$q(S) = \sum_{i=1}^n \frac{R_0^i}{B_0} \frac{e^{-\mu_i + v_i^2}}{S_0} \ell(S; \mu_i - v_i^2, v_i). \quad (50)$$

Integrating over S and invoking (29) gives the bond pricing equation:

$$B_0 = \sum_{i=1}^n \frac{R_0^i}{S_0} e^{-\mu_i + v_i^2}. \quad (51)$$

Thus, from (49), the bond price is a risk-capital weighted average of each investor's expectation of $\frac{S_0}{S}$. Recalling that the risk-capital $R_0^i = W_0^i + B_0 \tau_i$ depends on the bond price, substitution gives an explicit expression:

$$B_0 = \frac{\sum_{i=1}^n \beta_i e^{-\mu_i + v_i^2}}{1 - \sum_{i=1}^n \frac{\tau_i}{S_0} e^{-\mu_i + v_i^2}}. \quad (52)$$

This expression simplifies if we further assume that $\tau_i = 0 \forall i$:

$$B_0 = \sum_{i=1}^n \beta_i e^{-\mu_i + v_i^2}. \quad (53)$$

Thus, each investor's expectation of $\frac{S_0}{S}$ is now weighted by their initial stock endowment.

²⁵ From (42), (49) holds for any density when $\tau = 0$.

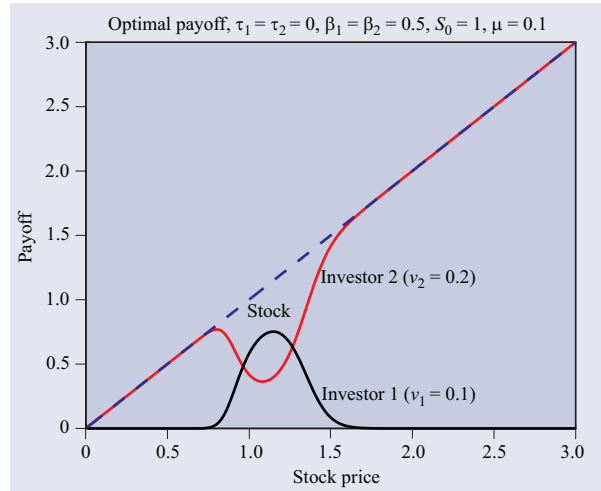


Figure 4. Optimal payoffs when means agree.

In general, (50) indicates that the risk-neutral density is not log-normal, even though each investor has log-normal beliefs. However, under homogeneous beliefs, the risk-neutral density simplifies to:

$$q(S) = \ell(S; \mu - v^2, v). \quad (54)$$

In this case, the risk-neutral density is log-normal with mean $\mu - v^2$ and volatility v . The requirement (32) that the stock's risk-neutral expected return be the risk-free rate implies that $\mu - v^2 = r$, so that the Black–Scholes formula holds for options, as shown in Rubinstein [45]. However, no one holds any options in this economy, as shown in (44) or subsection 4.2.1.

Returning to the case of heterogeneous log-normal beliefs, substituting (47) in (43) gives the optimal payoff as:

$$\phi_i(S) = -\tau_i + \frac{R_0^i \ell(S; \mu_i, v_i) S}{\sum_{i=1}^n R_0^i \ell(S; \mu_i, v_i)}. \quad (55)$$

since the aggregate floor is zero. Figure 4 shows the optimal payoffs for a two-investor economy when the investors have the same initial wealths, the same floor of zero and agree on the mean. Investor 1 believes volatility is 10%, while investor 2 thinks it is 20%. The optimal payoff for investor 1 resembles a bell-shaped curve, consistent with his low volatility view and his floor of zero. The optimal payoff of investor 2 accommodates the payoff of investor 1 and the requirement that payoffs sum to the stock price.

Assuming that investors all agree on volatility, $v_i = v \forall i$, then the optimal payoff simplifies to:

$$\phi_i(S) = -\tau_i + \frac{\hat{R}_0^i (S/S_0)^{p_i} S}{\sum_{i=1}^n \hat{R}_0^i (S/S_0)^{p_i}}, \quad (56)$$

where $\hat{R}_0^i \equiv R_0^i e^{p_i(\mu_i - v)/2}$ and $p_i \equiv \frac{\mu_i}{v^2}$. Thus, the optimal customized payoff when means differ is a power of the stock price divided by a sum of powers. Figure 5 shows the optimal payoffs for a two-investor economy when the investors have the same initial wealths, the same floor of zero and agree

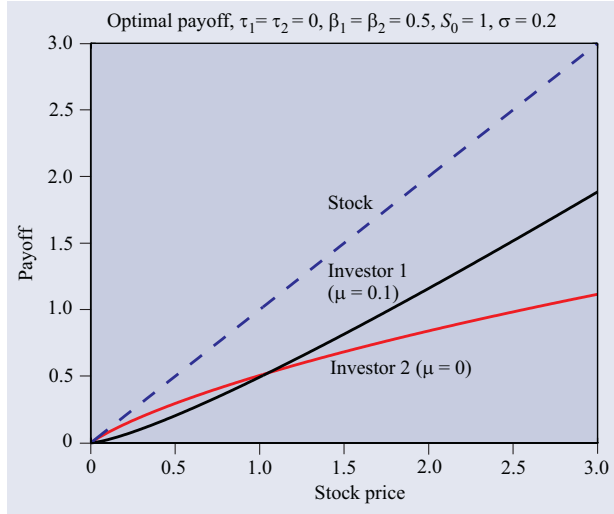


Figure 5. Optimal payoffs when volatilities agree.

on the volatility. Investor 1 believes the expected return is 10%, while investor 2 thinks it is 0%. Both payoffs would be synthesized using long positions in the stock. However, the more bullish investor 1 borrows at the risk-free rate and buys options while the less bullish investor 2 lends at the risk-free rate and sells options. Thus, in contrast to the case with homogeneous beliefs, options are used in the optimal portfolio, even if investors agree on volatility.

To obtain a separation result for the n investors case with equal volatility, let us further suppose that the n investors select their means from among $m < n$ possible values μ_1, \dots, μ_m . In particular, if investor i believes that the mean is μ_j , where μ_j is one the m possible values, then his optimal payoff is:

$$\phi_i(S) = -\tau_i + \frac{\widehat{R}_0^j(S/S_0)^{p_j} S}{\sum_{k=1}^m \widehat{R}_0^k(S/S_0)^{p_k}}.$$

In aggregate, the n investors hold m risky funds, although each investor only has non-zero holdings in one risky fund.

To summarize, investors with unitary cautiousness do not hold derivatives if beliefs are homogeneous, while derivatives are held when beliefs differ. The next subsection shows that both conclusions also hold when cautiousness is zero.

5.4. Negative exponential utility

Consider an n person economy in which all investors have constant risk tolerances, i.e. $T_i[\phi(S)] = \tau_i \forall i$.

From (37), each investor has an optimal exposure of the form:

$$\phi'_i(S) = \frac{\tau_i}{\tau} + \tau_i \left[\frac{d \ln p_i(S)}{dS} - \sum_{i=1}^n \frac{\tau_i}{\tau} \frac{d \ln p_i(S)}{dS} \right], \quad (57)$$

where now $\tau \equiv \sum_{i=1}^n \tau_i$ is the total risk tolerance. Integration gives the optimal payoff in terms of bonds, stocks and

derivatives:

$$\phi_i(S) = \kappa_i + \frac{\tau_i}{\tau} S + \tau_i d_i(S), \quad (58)$$

where $d_i(S) \equiv \ln p_i(S) - \sum_{i=1}^n \frac{\tau_i}{\tau} \ln p_i(S)$. The constant of integration κ_i is determined by substituting (58) in the budget constraint (12).

$$\kappa_i = \frac{W_0^i - \frac{\tau_i}{\tau} S_0 - \tau_i V_0[d_i]}{B_0}.$$

In this economy, each investor's stock and derivatives position does not depend on his initial wealth. Thus, the bond position is used to finance the positions in stocks and derivatives. The magnitude of this position in stock and derivatives depends on their risk tolerance. The greater the risk tolerance, the greater the exposure to stocks and derivatives. Each investor's stock position does not depend on his beliefs. In contrast, each investor's derivatives position depends mainly on the extent to which his beliefs differ from those in the market. Thus, the open interest in derivatives markets is primarily a reflection of the heterogeneity of beliefs. If investors have homogeneous beliefs but differing risk aversion, then they do not hold derivatives. Differences in risk aversion under homogeneous beliefs affect only the division between the riskless asset and the stock.

To obtain separation results under constant risk tolerance and heterogeneous beliefs, note from (47) that the log of the log-normal density is a linear combination of the log-price relative and its square. Suppose more generally that the log of each personal density can be written as a linear combination of basis functions:

$$\ln p_i(S) = \sum_{k=1}^m c_{ik} f_k(S). \quad (59)$$

Then (58) implies that the optimal payoff separates into $m+2$ funds:

$$\phi_i(S) = \kappa_i + \frac{\tau_i}{\tau} S + \tau_i \sum_{k=1}^m \left[\left(c_{ik} - \sum_{i=1}^n \frac{\tau_i}{\tau} c_{ik} \right) f_k(S) \right]. \quad (60)$$

Furthermore, the m derivative funds are the m basis functions which make up the log of the density. The optimal holding in the k th fund is $\tau_i \left(c_{ik} - \sum_{i=1}^n \frac{\tau_i}{\tau} c_{ik} \right)$. Thus, if investors agree on the coefficient of $\ln p$ on the j th basis function, i.e. $c_{ij} = c_j$, then that fund is not held by anyone.

Under constant risk tolerance, the risk-neutral density given in (35) simplifies to:

$$q(S) = \kappa \exp(-S/\tau) \prod_{i=1}^n [p_i(S)]^{\frac{\tau_i}{\tau}}, \quad (61)$$

where κ is a normalizing constant given by the requirement that q integrates to 1. Thus, the market view is a risk-tolerance weighted geometric average of the individual densities.

Given a specification of probability beliefs and an array of risk tolerances, it is straightforward to use (61) to value an option or any other derivative.

Note that under homogeneous beliefs, (61) simplifies to:

$$q(S) = \kappa \exp(-S/\tau) p(S). \quad (62)$$

Thus, if $p(S)$ is normal, then $q(S)$ is also normal with the same variance and with mean equal to the forward price as shown in Brennan [7]. The next subsection assumes that p is log-normal, and shows that q ends up in a different class from p . We also allow for heterogeneity in means and volatilities.

5.4.1. Zero cautiousness and log-normal beliefs. Recall that when all investors have log-normal beliefs, the log of each density is quadratic in $x \equiv \ln(S/S_0)$:

$$\ln l(S; \mu_i, v_i) = -\ln(\sqrt{2\pi} v_i S_0) - x - \frac{1}{2} \left[\frac{x - (\mu_i - v_i^2/2)}{v_i} \right]^2. \quad (63)$$

To obtain the optimal payoff under log-normal beliefs and constant risk tolerance of τ_i , substitute (63) in (58):

$$\phi_i(S) = \frac{W_0^i - m_i V_0[x] + p_i V_0[x^2/2] - S_0 \tau_i / \tau}{B_0} + \frac{\tau_i}{\tau} S + m_i x - p_i \frac{x^2}{2},$$

where

$$m_i \equiv \tau_i \left(\frac{\mu_i}{v_i^2} - \sum_{i=1}^n \frac{\tau_i \mu_i}{\tau v_i^2} \right)$$

and

$$p_i \equiv \tau_i \left(\frac{1}{v_i^2} - \sum_{i=1}^n \frac{\tau_i}{\tau v_i^2} \right).$$

Under constant risk tolerance and log-normal beliefs, the optimal payoff for each investor involves just two derivatives, one paying the log of the stock price and the other paying its square. As discussed in subsection 4.2.4, the log contract is used to speculate on expected return, while the squared-log contract is used to speculate on variance. If investors agree on the ratio of the mean to the variance, then from the definition of m_i , they do not hold the log contract. Similarly, if investors agree on volatility (i.e. $v_i = v \forall i$), then they do not hold the log-squared contract. If, in addition, investors agree on the mean (i.e. $\mu_i = \mu \forall i$), then no derivatives are held, consistent with (38).

In order to determine each investor's position in the riskless fund, the two derivative funds must be priced. Substituting (63) in (61) implies that the equilibrium risk-neutral density is:

$$q(S) = \kappa' \exp(-S/\tau) \ell(S; \hat{\mu}, \hat{v}) \quad (64)$$

where the aggregate precision $\frac{1}{\hat{v}^2} = \sum_{i=1}^n \frac{\tau_i}{\tau} \frac{1}{v_i^2}$ is a risk-tolerance weighted average of the individual precisions, and $\hat{\mu} = \frac{\sum_{i=1}^n (\tau_i/v_i^2) \mu_i}{\sum_{i=1}^n (\tau_i/v_i^2)}$ is a weighted average of the individual means, where the weights are given by the ratio of the risk tolerance to the risk. The constant κ' is determined by requiring that q integrate to 1:

$$\kappa' = 1 / \left[\int_0^\infty \exp(-S/\tau) \ell(S; \hat{\mu}, \hat{v}) dS \right]. \quad (65)$$

Unfortunately, the denominator is the Laplace transform of a log-normal density, which must be determined numerically. Once κ' is known, the values of the two derivatives funds are also obtained by quadrature:

$$V_0[x^j] = B_0 \int_0^\infty [\ln(S/S_0)]^j q(S) dS, \quad j = 1, 2.$$

To obtain the bond price, multiply (64) by $B_0 S$ and integrate over S :

$$B_0 \int_0^\infty S \kappa' \exp(-S/\tau) \ell(S; \hat{\mu}, \hat{v}) dS = S_0,$$

from (32). Substituting $S \ell(S; \hat{\mu}, \hat{v}) = S_0 e^{\hat{\mu}} \ell(\hat{\mu} + \hat{v}^2, \hat{v})$ gives the bond pricing equation:

$$B_0 = 1 / \left[\int_0^\infty \kappa' \exp(\hat{\mu} - S/\tau) \ell(S; \hat{\mu} + \hat{v}^2, \hat{v}) dS \right].$$

From (64), we note that $q(S)$ is not a log-normal density even though each investor believes that the stock price is log-normally distributed. However, the market view is log-normal since it is a geometric average of the log-normal individual views. The negative exponential adds negative skewness to this log-normal density²⁶. As a result, a graph of Black–Scholes implied volatilities against strike prices will slope down, as is observed in equity index option markets.

6. Summary and future research

Our primary contribution is the delineation of the optimal payoffs which arise for investors in both a partial and a general equilibrium context. In each case, the optimal payoff is chosen so that the marginal utility of the initial investment in each state is equalized across states. This Marshallian principle leads to an optimal payoff given by the inverse of the marginal utility function evaluated at the state price per unit of probability. Optimal *positions* are hence seen as the outcome of combining investor *preferences*, *probabilities* and *prices* of state contingent dollars.

In our partial equilibrium setting, we observe that investors use derivatives even when their personal volatility agrees with implied volatility. They also use derivatives when beliefs suggest selling options, while risk aversion suggests buying them. In particular, when implied volatility exceeds personal volatility, the optimal position involves writing near-the-money options to capture the volatility view, coupled with buying out-of-the-money options to eliminate unlimited downside exposure.

In our general equilibrium setting, we show that for unitary or zero cautiousness, homogeneous beliefs induce investors to shun derivatives, even though they differ in risk aversion. However, under heterogeneous beliefs or other preference specifications, investors optimally hold derivatives individually, even though they are not held in aggregate. Under negative exponential utility and log-normal beliefs, a four-fund separation occurs in which in addition to the bond and

²⁶ However, if one investor is risk-neutral, say the n th investor, then aggregate risk tolerance is infinite, and $q(S) = \ell(S; \mu_n, v_n)$.

the stock, investors take positions in two other derivatives: one which pays the log of the price and the other which pays the square of the log. The log contract is primarily used to express views on the mean, whereas the squared-log contract is a vehicle for trading volatility. If investors use options to create the squared-log contract, then the discontinuity in slope at the current stock price induces relatively large positions in at-the-money options. In a multiperiod setup, movement of the stock price would induce a large trading volume in such options, a phenomenon which is universally observed in listed options markets.

It would also be useful to investigate more fully the relationship between the implications of heterogeneous beliefs and the consequences of background risk as studied in Franke *et al* [22]. For example, it would be interesting to investigate whether higher background risk corresponds to a larger effective volatility view held by an investor engaged in a buy and hold strategy. Other interesting directions for future research would be to extend these results to a multiperiod or intertemporal setting. In a continuous time economy with continuous trading opportunities, jumps of random size would induce the demand for options demonstrated here. The optimal control problem induced by this extension is solved in Carr and Madan [10]. To our knowledge, the general equilibrium formulation with jumps in continuous time remains an open problem. Hence, an investigation of the properties of such an equilibrium is an interesting topic for future research.

Acknowledgments

We thank Michael Brennan, Mark Cassano, George Constantinides, Phil Dybvig, Rahim Esmailzadeh, Bob Jennings, Vikram Pandit and Lisa Polsky for their comments. We are also grateful to participants in finance workshops at the University of Chicago, Columbia University, Cornell University, ESSEC, the Fields Institute, Indiana University, New York University, Purdue University, Poincare Institute, University of Texas at Austin and Washington University at St Louis, in the Queen's University 1997 Derivatives Conference, in the Global Derivatives 97 Conference and in Risk Conferences on Advanced Mathematics for Derivatives, and on Asset and Liability Management. Any errors are our own.

Appendix 1. Proof of equation (1)

The fundamental theorem of calculus implies that for any fixed F :

$$\begin{aligned} f(S) &= f(F) + 1_{S>F} \int_F^S f'(u) du - 1_{S<F} \int_S^F f'(u) du \\ &= f(F) + 1_{S>F} \int_F^S \left[f'(F) + \int_F^u f''(v) dv \right] du \\ &\quad - 1_{S<F} \int_S^F \left[f'(F) - \int_u^F f''(v) dv \right] du. \end{aligned}$$

Noting that $f'(F)$ does not depend on u and applying Fubini's theorem:

$$\begin{aligned} f(S) &= f(F) + f'(F)(S - F) + 1_{S>F} \int_F^S \int_v^S f''(v) du dv \\ &\quad + 1_{S<F} \int_S^F \int_S^v f''(v) du dv. \end{aligned}$$

Performing the integral over u yields:

$$\begin{aligned} f(S) &= f(F) + f'(F)(S - F) + 1_{S>F} \int_F^S f''(v)(S - v) dv \\ &\quad + 1_{S<F} \int_S^F f''(v)(v - S) dv \\ &= f(F) + f'(F)(S - F) + \int_F^\infty f''(v)(S - v)^+ dv \\ &\quad + \int_0^F f''(v)(v - S)^+ dv. \end{aligned} \tag{66}$$

Setting $F = S_0$, the initial stock price, gives equation (1). Note that if $F = 0$, the replication involves only bonds, stocks, and calls:

$$f(S) = f(0) + f'(0)S + \int_0^\infty f''(v)(S - v)^+ dv,$$

provided the terms on the right-hand side are all finite. Similarly, for claims with $\lim_{F \uparrow \infty} f(F)$ and $\lim_{F \uparrow \infty} f'(F)F$ both finite, we may also replicate using only bonds, stocks and puts:

$$f(S) = \lim_{F \uparrow \infty} f(F) + \lim_{F \uparrow \infty} f'(F)(S - F) + \int_0^\infty f''(v)(v - S)^+ dv.$$

Appendix 2. Proof of equations (3) and (4)

Given the existence of options of all strikes, absence of arbitrage and (66) imply:

$$\begin{aligned} V_0[f] &= [f(F) - f'(F)]B_0 + f'(F)S_0 \\ &\quad + \int_F^\infty f''(v)(S - v)^+ dv + \int_0^F f''(v)(v - S)^+ dv. \end{aligned} \tag{67}$$

Integrating (2) by parts gives:

$$\begin{aligned} V_0[f] &= [f(F) - f'(F)]B_0 + f'(F)S_0 f'(K)P(K) \Big|_0^F \\ &\quad - \int_0^F f'(K)P'(K) dK \\ &\quad + f'(K)C(K) \Big|_F^\infty - \int_F^\infty f'(K)C'(K) dK. \end{aligned}$$

Since $P(0) = 0$ and $C(\infty) = 0$ and $C(F) - P(F) = S_0 - FB_0$, the second, third and fifth terms cancel. Integrating by parts once again yields:

$$V_0[f] = B_0 f(F) - f(K)P'(K) \Big|_0^F + \int_0^F f(K)P''(K)dK \\ - f(K)C'(K) \Big|_F^\infty + \int_F^\infty f(K)C''(K)dK.$$

Noting that $C'(\infty) = P'(0) = 0$ and $P'(F) - C'(F) = B_0$ by differentiating put call parity, we observe that:

$$V_0[f] = B_0 \int_0^\infty f(K)q(K)dK,$$

where $q(K)$ is proportional to the second derivative with respect to strike of the option pricing function:

$$q(K) \equiv \begin{cases} \frac{1}{B_0} \frac{\partial^2 P(K)}{\partial K^2} & \text{for } K \leq F; \\ \frac{1}{B_0} \frac{\partial^2 C(K)}{\partial K^2} & \text{for } K > F. \end{cases}$$

Setting $F = S_0$ gives the desired result.

Appendix 3. Proof of equation (39)

Recall that under homogeneous beliefs, the optimal exposure simplifies to:

$$\phi'_i(S) = \frac{T_i[\phi_i(S)]}{T(S)},$$

where $T(S) \equiv \sum_{i=1}^n T_i[\phi_i(S)]$. Suppose we have $n = 2$ investors with linear risk tolerance:

$$\phi'_1(S) = \frac{\tau_1 + \gamma_1 \phi_1(S)}{\tau + \gamma_1 \phi_1(S) + \gamma_2 \phi_2(S)} \quad (68)$$

$$\phi'_2(S) = \frac{\tau_2 + \gamma_2 \phi_2(S)}{\tau + \gamma_1 \phi_1(S) + \gamma_2 \phi_2(S)}, \quad (69)$$

where $\tau \equiv \tau_1 + \tau_2$. This is a coupled system of nonlinear ODEs. Fortunately, it can be solved if we assume opposite cautiousness, i.e. $\gamma_1 = -\gamma_2$. Without loss of generality, let $\gamma_1 = -\gamma_2 = \gamma \geq 0$. Dividing (68) by (69) implies:

$$\frac{\phi'_1(S)}{\phi'_2(S)} = \frac{\tau_1 + \gamma \phi_1(S)}{\tau_2 - \gamma \phi_2(S)}.$$

Rearranging gives $\gamma[\phi_1(S)\phi'_2(S) + \phi_2(S)\phi'_1(S)] - \tau_2\phi'_1(S) + \tau_1\phi'_2(S) = 0$. Integrating both sides gives $\gamma\phi_1(S)\phi_2(S) - \tau_2\phi_1(S) + \tau_1\phi_2(S) = c$, where c is the constant of integration. Substituting $\phi_1(S) = S - \phi_2(S)$ gives a quadratic in ϕ_2 :

$$\gamma[S - \phi_2(S)]\phi_2(S) - \tau_2[S - \phi_2(S)] + \tau_1\phi_2(S) = c.$$

Dividing by 2γ and rearranging gives:

$$\frac{1}{2}\phi_2^2(S) - \left[\frac{S}{2} + \frac{\tau}{2\gamma} \right] \phi_2(S) + \frac{\tau_2 S + c}{2\gamma} = 0,$$

with solution:

$$\phi_2(S) = \frac{S}{2} + \frac{\tau}{2\gamma} - \sqrt{\left(\frac{S}{2} + \frac{\tau}{2\gamma} \right)^2 - \frac{\tau_2 S + c}{\gamma}}. \quad (70)$$

Since $\phi_1(S) = S - \phi_2(S)$, we have:

$$\phi_1(S) = \frac{S}{2} - \frac{\tau}{2\gamma} + \sqrt{\left(\frac{S}{2} + \frac{\tau}{2\gamma} \right)^2 - \frac{\tau_2 S + c}{\gamma}}.$$

In order that both payoffs be real, we require:

$$\left(\frac{S}{2} + \frac{\tau}{2\gamma} \right)^2 - \frac{\tau_2 S + c}{\gamma} \geq 0.$$

Completing the square gives $\left(\frac{S}{2} - \frac{\tau_2 - \tau_1}{2\gamma} \right)^2 + \frac{\tau_1 \tau_2}{\gamma^2} - \frac{c}{\gamma} \geq 0$. Thus, a necessary condition for real payoffs is that $c \leq \frac{\tau_1 \tau_2}{\gamma}$. Choosing c so that this condition holds, define k^2 by:

$$c = \frac{\tau_1 \tau_2}{\gamma} - k^2 \gamma.$$

Then the optimal payoffs can be written as:

$$\phi_1(S) = \frac{S}{2} - \frac{\tau}{2\gamma} + \sqrt{\left(\frac{S}{2} - \frac{\tau_2 - \tau_1}{2\gamma} \right)^2 + k^2}$$

$$\phi_2(S) = \frac{S}{2} + \frac{\tau}{2\gamma} - \sqrt{\left(\frac{S}{2} - \frac{\tau_2 - \tau_1}{2\gamma} \right)^2 + k^2}.$$

Note that if we set $k = 0$, then the payoffs are linear:

$$\phi_1(S) = S - \frac{\tau_2}{\gamma} \quad \phi_2(S) = \frac{\tau_2}{\gamma}.$$

In any case, the positions sum to the stock as required. Furthermore, since $T_1[\phi_1(S)] = \tau_1 + \gamma\phi_1(S)$, we have:

$$T_1[\phi_1(S)] \\ = \frac{\gamma S - (\tau_2 - \tau_1)}{2} + \sqrt{\left(\frac{\gamma S - (\tau_2 - \tau_1)}{2} \right)^2 + \gamma^2 k^2},$$

which is non-negative. Similarly, since $T_2[\phi_2(S)] = \tau_2 - \gamma\phi_2(S)$, we have:

$$T_2[\phi_2(S)] \\ = -\frac{\gamma S - (\tau_2 - \tau_1)}{2} + \sqrt{\left(\frac{\gamma S - (\tau_2 - \tau_1)}{2} \right)^2 + \gamma^2 k^2},$$

which is also non-negative.

References

- [1] Arditti F and John K 1980 Spanning the state space with options *J. Financial Quantitative Anal.* **15** 1–9
- [2] Benninga S and Blume M 1985 On the optimality of portfolio insurance *J. Finance* **40** 1341–52
- [3] Benninga S and Mayshar J 2000 Heterogeneity and option pricing *Rev. Derivatives Res.* **4** 7–27
- [4] Bick A 1987 On the consistency of the Black–Scholes model with a general equilibrium framework *J. Financial Quantitative Anal.* **22** 259–76

- [5] Bick A 1990 On viable diffusion price processes of the market portfolio *J. Finance* **45** 672–89
- [6] Breeden D and Litzenberger R 1978 Prices of state contingent claims implicit in option prices *J. Business* **51** 621–51
- [7] Brennan M J 1979 The pricing of contingent claims in discrete time models *J. Finance* **34** 53–68
- [8] Brennan M J, Lagnado R and Schwartz E 1996 Strategic asset allocation *UCLA Working Paper*
- [9] Brennan M J and Solanki R 1981 Optimal portfolio insurance *J. Financial Quantitative Anal.* **16** 279–300
- [10] Carr P and Madan D B 2001 Optimal derivative investment in continuous time *Finance Stochastics* at press
- [11] Cass D and Stiglitz J 1970 The structure of investor preferences and asset returns and separability in portfolio allocation *J. Econ. Theory* **2** 122–60
- [12] Cassano M 2000 A two person general equilibrium analysis of option volume and the put call ratio *University of Calgary Working Paper*
- [13] Cox J C 1983 Optimum consumption and portfolio policies when asset prices follow a diffusion process *Graduate School of Business, Stanford University, Research Paper* 658
- [14] Cox J C and Huang C F 1989 Optimum consumption and portfolio policies when asset prices follow a diffusion process *J. Econ. Theory* **49** 33–83
- [15] Cox J C and Huang C F 1989 Option pricing theory and its applications *Theory of Valuation: Frontiers of Modern Financial Theory* ed S Bhattacharya and G Constantinides vol 1 pp 272–88
- [16] Draper D and Shimko D 1993 On the existence of ‘redundant securities’ *University of Southern California Working Paper*
- [17] Duffie D and Jackson M 1989 Optimal innovation of futures contracts *Rev. Financial Studies* **2** 275–96
- [18] Dumas B 1989 Two-person dynamic equilibrium in the capital market *Rev. Financial Studies* **2** 157–88
- [19] Dupire B 1993 Model art *Risk* **6** 118–20
- [20] Dybvig P H and Ingersoll J E 1982 Mean-variance theory in complete markets *J. Business* **55** 233–52
- [21] Evnine J and Henriksson R 1987 Asset allocation and options *J. Portfolio Management* **14** 56–61
- [22] Franke G, Stapleton R and Subrahmanyam M 1998 Who buys and who sells options: the role of options in an economy with background risk *J. Econ. Theory* **82** 89–109
- [23] Franke G, Stapleton R and Subrahmanyam M 2000 When are options overpriced? The Black Scholes model and alternative characterizations of the pricing kernel *Euro. Finance Rev.* **3** 79–102
- [24] Friesen P 1979 The Arrow Debreu model extended to financial markets *Econometrica* **47** 689–707
- [25] Gibson R 1996 *Asset Allocation: Balancing Financial Risk* 2nd edn (Chicago, IL: Irwin)
- [26] Green R C and Jarrow R A 1987 Spanning and completeness in markets with contingent claims *J. Econ. Theory* **41** 202–10
- [27] Hakansson N 1977 The superfund: efficient paths toward efficient capital markets in large and small countries *Financial Decision Making Under Uncertainty* ed H Levy and M Sarnat (New York: Academic)
- [28] Hakansson N 1978 Welfare aspects of options and supershares *J. Finance* **33** 759–76
- [29] He H and Leland H 1993 On equilibrium asset price processes *Rev. Financial Studies* **6** 593–617
- [30] Huang C F and Litzenberger R 1988 *Foundations for Financial Economics* (New York: North-Holland)
- [31] Jarrow R A and Madan D B 1997 Is mean-variance analysis vacuous: or was beta still born *Euro. Finance Rev.* **1** 15–30
- [32] Johnston L and McConnell J 1989 Requiem for a market: an analysis of the rise and fall of a financial futures contract *Rev. Financial Studies* **2** 1–24
- [33] Kreps D 1982 *Multiperiod Securities and the Efficient Allocation of Risk: A Comment on the Black–Scholes Option Pricing Model* (Chicago, IL: University of Chicago Press) chapter 6 pp 203–32
- [34] Leibowitz M, Bader L and Kogelman S 1996 *Return Targets and Shortfall Risks* (New York: Wiley)
- [35] Leland H 1980 Who should buy portfolio insurance *J. Finance* **35** 581–94
- [36] Markowitz H M 1952 Portfolio selection *J. Finance* **7** 77–91
- [37] Merton R C 1971 Optimum consumption and portfolio rules in a continuous time model *J. Econ. Theory* **3** 373–413
- [38] Nachman D 1988 Spanning and completeness with options *Rev. Financial Studies* **3** 311–28
- [39] Neuberger A 1990 Volatility trading *London Business School Working Paper*
- [40] Pliska S 1982 A discrete time stochastic decision model *Advances in Filtering and Optimal Stochastic Control* ed W H Fleming and L G Gorostiza (*Lecture Notes in Control and Information Sciences* **42**) (Berlin: Springer) pp 290–304
- [41] Pliska S 1986 A stochastic calculus model of continuous trading: optimal portfolios *Math. Operations Res.* **11** 371–82
- [42] Richards J I and Youn H K 1990 *Theory of Distributions: A Non-technical Introduction* (Cambridge: Cambridge University Press)
- [43] Ross S 1976 Options and efficiency *Q. J. Econ.* **90** 75–89
- [44] Ross S 1978 Mutual fund separation and financial theory—the separating distributions *J. Econ. Theory* **17** 254–86
- [45] Rubinstein M 1976 The valuation of uncertain income streams and the pricing of options *Bell J. Econ.* **7** 407–25
- [46] Rubinstein M 1976 The strong case for the generalized logarithmic utility model as the premier model of financial markets *J. Finance* **31** 551–71
- [47] Shimko D 1994 A tail of two distributions *Risk* **7** 123–30
- [48] Stapleton R and Subrahmanyam M 1984 The valuation of options when asset returns are generated by a binomial process *J. Finance* **39** 1525–39
- [49] Tilley J and Latainer G 1985 A synthetic option framework for asset allocation *Financial Analysts J.* **34** 25–32
- [50] Vince R 1995 *The New Money Management: A Framework for Asset Allocation* (New York: Wiley)