This paper will appear in the January, 1991, issue of *Mathematical Finance*.

CASH STREAM VALUATION IN THE FACE OF TRANSACTION COSTS AND TAXES

Jaime Cuevas Dermody
Finance Dept. (m/c 168), University of Illinois at Chicago
Chicago, IL 60607

and

R. Tyrrell Rockafellar
Applied Mathematics Dept. FS-20, University of Washington
Seattle, Washington 98195

6 January 1990
(revised 21 July 1990)

Abstract. The usual notion of every future cash stream having a net present value determined from a single term structure breaks down when transaction costs are taken into account, especially the sizable costs associated with short-borrowing. The difficulties are compounded by taxes, which can lead to paradoxes of disequilibrium if elementary NPV is assumed to be a rational basis for decision making. This paper systematically develops a theory of valuation which overcomes these shortcomings by accepting the multiplicity of no-arbitrage term structures that may be present for each tax class of investors and using the entire set of them to impute both a "long price" and a "short price" for every cash stream, regardless of the sign of the future payments. The valuation operators giving these prices are nonlinear but readily calculated from linear programming formulas.

1. INTRODUCTION

A fundamental problem in finance is that of determining for any investor the value of a given future cash stream. Such a cash stream, if essentially riskless in nominal dollars, can be viewed as a vector $w = (w_1, \ldots, w_m)$, where w_i is the amount of money to be received at time i, for instance a bond payment date. Each w_i can be positive, zero, or negative. An investor faced with the prospect of acquiring w at a price c must decide whether w is overvalued or undervalued at that price, or neither. The cash stream w could be derived from some special asset or project opportunity, or on the other hand, it could correspond to a traded security or portfolio of such securities. Either way, the issue is deeply tied to price equilibrium and the theory of value, as well as practical questions of investment opportunity and arbitrage.

A standard procedure has been to compare c with the so-called *net present value*, or NPV, of w, as given by a linear formula

$$dw = d_1 w_1 + \cdots + d_m w_m \tag{1.1}$$

relative to a single vector $d = (d_1, \ldots, d_m)$, called a *term structure.* Here d_i is a factor for discounting income at time i to the present. (See Ruback (1986), for instance.) An appropriate vector d has commonly been thought to exist for any investor or class of investors on the basis of utility theory and economics, and to be derivable from no-arbitrage models of financial markets. A key issue, however, is whether d is uniquely determined in a given context, for if not, there can be a whole range of values appropriately assigned to w by (1.1) and the concept of NPV ceases to have the same significance.

No-arbitrage models impute value to a cash stream w from the current prices for traded securities and other instruments that transfer money between present and future. In the theory of such models laid down by Ross (1976)(1978), the existence of a unique term structure d is established when the market is complete and transaction costs are ignored. But transaction costs, which include the costs of short-selling a security as well as those associated with the bid-ask spread, are far from negligible in practice. More serious still, the suppression of transaction costs entails the assumption that *every* traded security (whatever its type or maturity) can be shorted by every investor (regardless of holdings) at the same price at which it can be purchased.[1] Utility models likewise fail to provide a unique term structure d when transaction costs are present, so the difficulty cannot be thought of as arising because arbitrage models are rougher than utility models.

[1] These assumptions are implicit in the proofs of Ross (1978) and the view that the valuation operator obtained there is enforced by arbitrage being available otherwise through actual market transactions.

Transaction costs in a no-arbitrage model imply the existence of not just one equilibrium term structure d but an infinite set of them for each tax class of investors, as shown by Dermody and Prisman (1988). This indeterminacy has nothing to do with error; it is inherent in the market. On the other hand, the nonuniqueness cannot be dealt with by assuming that the market happens at any moment to be reacting to just one of the identified term structures. Besides the lack of any theoretical principle behind such an assumption, there are strong mathematical reasons for believing that the nonuniqueness indicates underlying nonlinearities in the behavior of value that cannot be captured by a single term structure.

Serious questions must then be answered, since it is essential to financial theory to have a solid mathematical foundation for the notion of what a future cash stream should be worth to an investor, taking reasonable account of the circumstances that affect financial decisions. Our aim here is to study, in the face of transaction costs and taxes, the case of *nonspeculative* investment and arbitrage involving *essentially riskless* cash streams (where the future payments are not regarded by the investor as entailing significant risk). Mathematically much of the work could be extended to a state space model for risky cash streams, but this is eschewed in order to keep the main ideas in better perspective.

We take the multiplicity of term structure as the starting point, and going farther than Dermody and Prisman (1988), use it to develop a consistent and rigorous approach to the valuation of general cash streams in the context of readily available financial data. We introduce, for each cash stream w and tax class $k = 1, \ldots, K$ of investors, an *imputed long price* $V^k(w)$ and an *imputed short price* $v^k(w)$. The first is the lowest amount of dollars through which an investor in class k can acquire w from a combination of transactions in the present market, i.e., take a "long position" in w. The second is the highest amount of dollars an investor can extract from the market in return for paying w in the future, i.e., taking a "short position" in w. In contrast to the NPV expression in (1.1), the formulas for $V^k(w)$ and $v^k(w)$ are *nonlinear* in the components w_i of w. We demonstrate that they correspond dually to linear programming computations over a term structure packet D^k, consisting of all the no-arbitrage term structures d for investor class k relative to current market prices. This set, which turns out to be a convex polyhedron in m-dimensional space, furnishes a new geometric basis for understanding various market phenomena, the geometry being quite different from that in Dermody and Prisman (1988,1990), where the graphs of transaction cost functions dominate.

The proper treatment of taxes has long been a serious challenge in the theory of finance. To quote the words of Dybvig and Ross (1986): "In the study of investments, taxes are largely an embarassment to financial economists. We know that taxes are significant, but we do not know the equilibrium effect of taxes on asset pricing and the consequent effect on portfolio choice." Moreover, in trying to understand the effect of taxes on the value of a cash stream or security there is a conceptual hurdle. Term structures, as a means of measuring worth to an individual investor or class of investors, can legitimately be applied only to *after-tax* cash streams, but the same asset can provide different after-tax cash streams to investors in different classes.

It is generally impossible for an investor in one tax class to transfer an after-tax cash stream, as such, to an investor in another class. On the other hand, financial assets, which can be transfered, have no inherent utility in themselves, apart from the after-tax income they bring. There would be little sense, in general, in thinking of investors as having utility functions based on individual preferences among different "baskets" of securities that are all riskless and yield the same after-tax results. Therefore, the buyers and sellers in a financial market are not necessarily trading in a common set of goods in the sense imagined for other economic markets. This means that a term structure appropriate for an investor in one tax class might on such grounds alone be inappropriate for an investor in another class. While $d = (d_1, ..., d_m)$ can be viewed as a sort of price vector, with d_i interpreted as the amount an investor would be willing to pay now for the rights to one tax-free dollar of income at time i, a theoretical economic mechanism for bringing such "prices" into line across different tax classes is lacking.

One mathematical counter to this situation has been to set up a different system of term structure equations for each tax class. Let us suppose that security j at price P_j provides for class k the after-tax cash stream $(a^k_{1j}, ..., a^k_{mj})$. Under the assumption that transaction costs play no role (and correspondingly that every security can be shorted at the same price P_j), the standard system of equations

$$(1.2) \qquad d^k_1 a^k_{1j} + \cdots + d^k_m a^k_{mj} = P_j \quad \text{for } j = 1, \ldots, n$$

has been taken as characterizing the term structure $d^k = (d^k_1, \ldots, d^k_m)$ appropriate for investors in class k. Trouble arises with real market data, however, in that the equations are typically inconsistent (with not enough unknowns) and have no solution d^k at all, much less a unique one for each class k. At best one can look then for "approximate" solutions of some sort.

An example furnished by Schaefer (1982a) demonstrates, however, that price equilibrium can actually be incompatible with (1.2). If the equations are written more carefully as

$$(1.3) \qquad d^k_1 a^k_{1j}(P_j) + \cdots + d^k_m a^k_{mj}(P_j) = P_j \quad \text{for } j = 1, \ldots, n$$

in order to reflect the fact that the after-tax cash stream furnished to an investor in class k by security j depends to an extent on the price P_j at which the security is acquired, there may well be no set of prices P_j such that a solution $d^k = (d^k_1, \ldots, d^k_m)$ exists for every $k = 1, \ldots, K$. Schaefer has interpreted this apparent paradox as indicating the need to incorporate shorting costs, or restrictions on shorting, in financial models.[2]

[2] To disallow shorting is to take the cost of short-borrowing so large as to equal the entire bid price of the security.

Others have sought alternative ways around the paradox. Dammon and Green (1987) have investigated tax properties that might allow for equilibrium despite an absence of transaction costs. Dybvig and Ross (1986) have referred to "the fallacy of assuming that agents have linear schedules (with linear marginal tax rates) in a model allowing short sales," thereby likewise taking the position that nonlinear taxes, rather than theoretical attention to transaction costs, should be the answer.

Litzenberger and Rolfo (1984) have proposed the introduction of a separate term structure for each of three categories of income from a security (namely ordinary income, capital gains, and nontaxable return of purchase price) as a means of avoiding the specific pitfalls in Schaefer's two-period example. The additional unknowns would provide more degrees of freedom that might be helpful in theory. But there is no insurance against the existence of equally devastating examples involving more than two time periods, and for these the proposed trick would not suffice. A more fundamental objection to the approach of Litzenberger and Rolfo is that it corresponds mathematically to the supposition that an investor cannot pool "long" income from different categories to cover "short" obligations. Dermody (1988) has demonstrated that if pooling were indeed allowed (as of course it is in actual markets), the proposed separate term structures for different categories would have to be proportional to each other. In that case the number of degrees of freedom in the equations would revert to what it was before, and no advantage would have been gained by the tactic of viewing bond payments as three different categories of income.

The possibility of separate term structures d^k has been eyed by some researchers in finance as an undesirable complication. A heuristic device for avoiding it has been to introduce a "representative" investor. In following such an approach, as pioneered by McCulloch (1975) and Carelton and Cooper (1976), and improved by Jordan (1984), one bypasses the tax classes that may actually be present and works with equations of the form

$$d_1 a_{1j}(r_1, \ldots, r_L) + \cdots + d_m a_{mj}(r_1, \ldots, r_L) = P_j \text{ for } j = 1, ..., n \tag{1.4}$$

where $(r_1, \ldots, r_L)$ is a parameter vector of tax rates on L classes of income, e.g., untaxed return of purchase price, capital gain, and ordinary income from coupons. The expression $a_{ij}(r_1, \ldots, r_L)$ refers to the after-tax payment by security j at time i as a function of these rates. Each tax class k corresponds in this conception to a specific rate vector $(r_1^k, \ldots, r_L^k)$ yielding $a_{ij}^k = a_{ij}(r_1^k, \ldots, r_L^k)$, but instead of solving the systems (1.2), one tries to solve the system (1.4) for a single combination of $(r_1, \ldots, r_L)$ and $(d_1, \ldots, d_m)$. In practice this means using a form of regression to determine vectors $(\bar{r}_1, \ldots, \bar{r}_L)$ and $(\bar{d}_1, \ldots, \bar{d}_m)$ that give as good a fit to the equations as possible, according to some criterion. The tax rates $(\bar{r}_1, \ldots, \bar{r}_L)$ are defined then as characterizing the so-called "representative" or "marginal" investor for purposes of bond pricing. The term structure $(\bar{d}_1, \ldots, \bar{d}_m)$ is proposed as a kind of aggregate for the market as a whole, which could be justified econometrically through statistical evidence that it reflects the way the market behaves.

The theory presented here looks elsewhere than equation-based models. It relies instead on systems of linear inequalities, which incorporate shorting costs and other transaction costs in specifying the packet D^k of no-arbitrage term structures associated with a tax class k. The linear programming formulas we express through D^k provide a practical methodology for answering questions about the real cost of capital for various groups of investors, as well as a framework for analyzing clientele effects, where different tax classes of investors exhibit different patterns of holdings.

A precedent for the utilization of linear programming and inequalities in the theory of term structure with multiple tax classes can be found in Hodges and Schaefer (1977). Those authors disallowed all shorting, although recognizing the importance of the fundamental issues it raises. They were not concerned with general valuation operators like our V^k and v^k and did not introduce a term structure packet D^k, but rather they identified a particular term structure relative to the optimal replication of a particular cash stream for a given investor. This term structure (actually there is no reason to believe it would be unique) was viewed as a kind of *term structure estimate* relative to the special interests of the investor in question.

In other work, Schaefer (1981)(1982b) applied linear programming ideas in schemes for estimating term structure which did not use least-squares regression. This was extended by Ronn (1987), once again with the idea of finding a single, "best" term structure d^k to use in the NPV formula (1.1). Implicit in Ronn's estimation problem is a term structure set $\hat{D}^k$ like our set D^k, the chief difference being that bid prices are used in place of short prices. In effect, Ronn minimizes a piecewise linear penalty function for membership in $\hat{D}^k$: when $\hat{D}^k$ is nonempty, the minimum is 0 and *every* term structure d in $\hat{D}^k$ is a "best estimate," whereas if $\hat{D}^k$ is empty, the availability of riskless arbitrage is reported and the market is seen as in disequilibrium. Thus, Ronn's paper implicitly concerns an infinite set of term structures, as does this paper, but where Ronn (in the case of a market in equilibrium and not presenting opportunities for arbitrage) stops with calculating an arbitrary element d of his set,[3] the corresponding linear valuation then being offered as the appropriate one for application, we utilize all the elements d of our set D^k to obtain nonlinear valuation operators V^k and v^k.

As to the analysis of clientele effects, the framework we adopt in this paper[4] has some ties to the article of Hodges and Schaefer (1977) but more broadly can be compared with Dybvig and Ross (1986). The efforts of those authors go in a direction quite different from ours, however. In any

---

[3] The empirical results presented by Ronn (1987) presumably refer to equilibrium cases. Although mathematically the term structure estimate d could therefore be any element of his set $\hat{D}^k$, his computer program comes up with a particular one.

[4] This framework for clientele effects has been sketched in part by Dermody and Prisman (1988) as cited from our working papers, Dermody and Rockafellar (1987a), (1987b).

case, their model is not "implementable," because it requires knowledge of every agent's complete utility function. In contrast, our model is based on ordinary market data.

We begin in Section 2 by explaining the model, and continue in Section 3 with a study of the corresponding valuation operators V^k and v^k. Nonlinearities in these operators are investigated in Section 4 in connection with a condition of "complete no-arbitrage." Clientele effects are discussed in Section 5.

2. THE MARKET MODEL AND TAX STRUCTURE

The times $i = 1, \ldots, m$ will refer to a fixed horizon of m time periods or payment dates. The securities $j = 1, \ldots, n$ will refer to essentially riskless securities, such as various marketable, non-callable national-government securities, but perhaps also forward contracts on those securities and even certain arrangements with banks, as will be discussed in due course. Nothing is truly riskless, obviously. We use the words "essentially riskless" to indicate transfers of money between future time periods and the present that would commonly be regarded as not involving substantive risk. To some degree it is up to the modeler to make this more precise in any application by deciding which securities to include.

Each security that is included must play out its life within the chosen time horizon. A security that still had payments coming after time $i = m$ would have an uncertain residual value which would have to be considered, and this would undermine its claim to being essentially riskless. While the problems of speculative investment, centered on the unknown future motion of security prices, are extremely important in finance, we do not treat them here, preferring first to lay a better foundation for the nonspeculative case. Thus, we concentrate on questions of arbitrage and investment that can be addressed in terms of *buy-and-hold positions where the securities are held until maturity.*

Security j provides its owner with the before-tax cash stream $(a_{1j}, \ldots, a_{mj})$. The current price at which it can be purchased is P_j, the so-called ask price. There is also a bid price, slightly lower, at which security j can be sold, and the difference between these two prices—the bid-ask spread—reflects a type of transaction cost. This is not, however, the transaction cost that most attracts our attention.

Transaction costs in the form of a bid-ask spread exist in most economic markets. Financial markets, however, are distinguished by the fact that price equilibrium is maintained not only by ordinary buying and selling but also by arbitrageurs taking long and short positions.

A "long" position in security j corresponds to the purchase of one or more units of that security in the usual sense. For each unit purchased, the investor pays P_j now and receives the cash stream $(a_{1j}, \ldots, a_{mj})$ later. A "short" position, on the other hand, corresponds to a special transaction in which the investor assumes the obligation of paying out (for each unit of security j that is shorted) the cash stream $(a_{1j}, \ldots, a_{mj})$ in return for receiving a cash amount p_j in the present. The amount p_j is distinctly less than the bid price for security j;

$$0 \le p_j < P_j \quad \text{for } j = 1, \ldots, n. \tag{2.1}$$

We find it convenient to call P_j the *long price* and p_j the *short price* for security j.

If security j cannot be shorted at all (as for instance with certain long-term bonds in today's financial markets), then $p_j = 0$ in our model. This effectively precludes a short position because an investor would get no reward for taking on the future payment obligations connected with the shorting. As mentioned earlier, one can think of the shorting cost in such a case as amounting to the entire bid price. The model can in this manner cover even a market where all shorting of bonds is forbidden, which falls quite outside the scope of the kind of term structure theory that supports linear valuation in the pricing of assets, as in Ross (1978).

Shorting is a form of borrowing, and shorting costs can be seen as additional costs that must be paid for such borrowing. The additional cost, beyond that already built into the cash stream associated with a given security, works out to a few percent of face value per year that the bond is shorted (if the shorting is possible at all). This is far too large a cost not to have significant consequences in an extremely liquid market like the one for U.S. Treasury obligations.

Although we have used the language of bonds in speaking about long and short positions, the reader should bear in mind that the securities in our model do not all have to be bonds. In particular, some could be bank deposits or fully collateralized loans that can be made to a bank (at a normalized amount), in which case the corresponding short positions consist of loans that can be obtained from a bank. Then the difference between P_j and p_j is more properly the current *borrowing-lending* spread, but for simplicity we still speak anyway of P_j and p_j as long and short prices. The inclusion of such securities in the model could be a way of compensating for an inability to short long-term bonds. Loans from a bank provide alternative means for transfering future cash payments into the present.

The general market transaction we study here for an investor is represented by two vectors $X = (X_1, \ldots, X_n)$ and $x = (x_1, \ldots, x_n)$, where $X_j \ge 0$ is the long position taken in security j (in number of units), and $x_j \ge 0$ is the short position. As already mentioned, the positions are to be regarded as buy-and-hold positions until maturity. There is no mathematical reason to insist that only one of the quantities X_j and x_j can be positive for any j. We keep the notation cleaner by relying on the economies of cost to discourage this. The vector pair (X, x) is called a *trade*. We view a trade (X, x) as a supplement to any previously acquired portfolio of the investor in question, but for present purposes we disregard the ways an investor might be able to take advantage of such existing holdings.

Let A denote the $m \times n$ matrix whose entries are the income amounts a_{ij}. The n columns of A correspond to the various securities j, while the m rows correspond to the various times i at which one or more of the securities makes a payment. In undertaking a trade (X, x), the investor pays the amount $P_1X_1 + \cdots + P_nX_n - p_1x_1 - \cdots - p_nx_n = PX - px$ in the present in return for receiving $a_{i1}X_1 + \cdots + a_{in}X_n - a_{i1}x_1 - \cdots - a_{in}x_n$ at time i for $i = 1, \ldots, m$, or in other words the future cash stream $A(X - x)$. This is an essentially riskless *before-tax* cash stream.

Investors are divided into separate classes indexed by $k = 1, \ldots, K$, according to the way their taxes are calculated. One of these can be a "tax-free" class, but we do not single it out with special notation. More than one government can be involved in the collection of the taxes.

For investors in class k, the amount of tax due at time i on each unit of security j held "long" is denoted by $t^k_{ij}(P_j)$, and the after-tax income at that time is then $a_{ij} - t^k_{ij}(Pj)$. For each unit of security j held "short," the tax amount $s^k_{ij}(p_j)$ appears correspondingly as a subsidy or credit against the obligation to pay a_{ij} at time i, so that the net after-tax payment to be made is $a_{ij} - s^k_{ij}(p_j)$. The taxes do depend in general on the prices P_j and p_j, because these prices set the "basis" for calculations of interest and capital gains. It is useful to us that this dependence be indicated explicitly, since later there will be important consequences for the question of when an investor will wish to liquidate certain current holdings. The tax functions $t^k_{ij}(\cdot)$ and $s^k_{ij}(\cdot)$ could well be the same in many cases, but we allow for asymmetry to cover examples like municipal bonds in the U.S., which are tax-free in a long position but lose this feature in a short position.

We speak of $t^k_{ij}(P_j)$ as a tax amount due at time i, but the modeler can interpret this in different ways in order to gain flexibility. Most work on taxes has adopted a pay-as-you-go pattern, where taxes are subtracted from taxable income as soon as it is received, but it is perfectly possible to figure taxes as being paid on the dates when they actually are due. In this case all the tax collection dates relevant for the specified time horizon must obviously be included among the times $i = 1, \ldots, m$. Similar treatment is available for the subsidy amounts $s^k_{ij}(p_j)$.

Let $T^k(P)$ be the $m \times n$ matrix whose entries are $t^k_{ij}(P_j)$, and $S^k(p)$ the one whose entries are $s^k_{ij}(p_j)$. The *after-tax* cash stream received by an investor in class k from a trade (X, x) is then $A(X - x) - T^k(P)X + S^k(p)x$. Here we are obviously treating taxes as simply additive. This is justified on the grounds that our model is largely concerned with investors whose tax *rates* will not be affected by the trades in question.

We shall assume that taxes and other payments can always be handed over in advance, if desired (without any corresponding discount of the amount paid). This minor assumption is not typically made explicit in the financial literature. Some modelers include instead a provision for depositing, at a positive rate of interest, any cash received earlier than needed. Such deposit provisions are admissible in our model as special "securities" to the extent that the interest rates can be locked in on the basis of arrangements made in *current markets.*

The advance payment assumption is incorporated in our model by a vector $h = (h_1, \ldots, h_m)$ called a *holdover schedule.* The quantity $h_i \geq 0$ represents a credit or reserve of money held over from time $i-1$ against payments that might be necessary at time i. Introducing the $m \times m$ matrix

$$J := \begin{bmatrix} 1 & -1 & & & & & \\ & 1 & -1 & & & & \\ & & 1 & \cdot & & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & -1 & \\ & & & & & 1 & -1 \\ & & & & & & 1 \end{bmatrix}$$

(with zeroes in the blank positions), we get the vector

$$Jh = (h_1 - h_2, h_2 - h_3, \ldots, h_{m-1} - h_m, h_m)$$

as giving the amounts reserved for such payments. Relative to a trade (X, x) and a holdover schedule h, therefore, the investor obtains the future after-tax cash stream

$$\big[A - T_k(P)\big]X - \big[A - S_k(p)\big]x + Jh$$

at a current cost of

$$PX - px + h_1 = PX - px + e_1 h, \text{ where } e_1 = (1, 0, \ldots, 0).$$

For purposes of matrix notation, it is evident that we are treating X, x and h always as column vectors. This is also the case with cash streams w. On the other hand, the vectors P and p (as well as d, eventually) are row vectors.

3. ARBITRAGE AND IMPUTED PRICES

A basic assumption about the model needs now to be stated.

DEFINITION. *The weak no-arbitrage condition WNA^k is satisfied for investor class k, relative to the current price vectors P and p, when it is impossible for any investor in this class to find a trade $(X,x) \geq (0,0)$ and holdover schedule $h \geq 0$ such that*

$$\big[A - T_k(P)\big]X - \big[A - S_k(p)\big]x + Jh \geq 0 \quad \text{and} \quad PX - px + e_1 h < 0. \tag{3.1}$$

If WNA^k holds for every $k = 1, \ldots, K$, we say the no-arbitrage condition WNA is satisfied.

The WNA^k condition means that no investor in class k is able to gain positive "walk-away" money (the money the investor receives in the present from a short position minus the cost of a long position, all of which is net of transaction costs) from an essentially riskless transaction that would leave no net positive payments to be made at any time in the future. Linear programming techniques can be used to test whether WNA is satisfied relative to current prices, and if not, to find specific tax classes and trades for which essentially riskless arbitrage is available.

Fixing on any one of the classes k, we now consider an essentially riskless *after-tax* cash stream $w = (w_1, \ldots, w_m)$ (with payments w_i that could be positive, zero or negative) and address the question of what it might be worth to a generic investor in this class.

DEFINITION. *The imputed long price, or long value, for the after-tax cash stream w relative to tax class k is the amount*

$$\begin{aligned} V^k(w) = \; & \text{minimum of } PX - px + e_1 h \text{ subject to} \\ & \big[A - T^k(P)\big]X - \big[A - S^k(p)\big]x + Jh \geq w \text{ with } X \geq 0,\ x \geq 0,\ h \geq 0. \end{aligned} \tag{3.2}$$

This is the lowest amount of present cash with which every investor in class k can procure, through some trade (X,x) and holdover schedule h, a future after-tax cash stream at least as good as w. The imputed short price, or short value, for w relative to tax class k, on the other hand, is the amount

$$\begin{aligned} v^k(w) = \; & \text{maximum of } -PX + px - e_1 h \text{ subject to} \\ & -\big[A - T^k(P)\big]X + \big[A - S^k(p)\big]x - Jh \leq w \text{ with } X \geq 0,\ x \geq 0,\ h \geq 0. \end{aligned} \tag{3.3}$$

This is the highest amount of present cash that every investor in class k can obtain by taking on the obligation of paying a future cash stream for which the after-tax burdens are no worse than w.

In effect, $V^k(w)$ is the price for a long position in w while $v^k(w)$ is the price for a short position—as seen in after-tax terms by a generic investor in class k. No investor in class k should pay more than the amount $V^k(w)$ for w, but on the other hand every such investor should be willing to pay at least $v^k(w)$ for w, because this net amount can immediately be obtained by a trade that uses w to subsidize all its net future payments.

THEOREM 3.1. *If condition WNA^k is satisfied, the imputed long and short prices $V^k(w)$ and $v^k(w)$ satisfy*

$$-\infty < v^k(w) \le V^k(w) < \infty, \qquad v^k(w) = -V^k(-w). \tag{3.4}$$

If the condition WNA^k is not satisfied, however, then $V^k(w) = -\infty$ and $v^k(w) = \infty$.

PROOF. Formula (3.3) can equally be written as

$$\begin{aligned} -v^k(w) = &\text{ minimum of } PX - px + e_1 h \text{ subject to} \\ &\big[A - T^k(P)\big]X - \big[A - S^k(p)\big]x + Jh \ge -w \text{ with } X \ge 0,\ x \ge 0,\ h \ge 0. \end{aligned} \tag{3.5}$$

This minimization also expresses $V^k(-w)$ by the definition in (3.2) and is the basis for asserting (3.4). We must demonstrate, however, that V^k is a well defined function.

Fix any w. There does exist at least one choice of X, x, and h satisfying the constraints in (3.2): if one were to take $X = 0$, $x = 0$, and $h = r(m, \dots, 3, 2, 1)$ with $r > 0$ large enough that $r \ge w_i$ for all i, it certainly would be true that $\big[A - T^k(P)\big]X - \big[A - S^k(p)\big]x + Jh \ge w$. Similarly the constraints in (3.5) can be satisfied, and these are identical to the constraints in (3.3). Thus $V^k(w) < \infty$ and $v^k(w) > -\infty$.

To prove that $v^k(w) \le V^k(w)$, we consider numbers $a' > V^k(w)$ and $a'' < v^k(w)$, and verify that the relation $a'' \ge a'$ is impossible. Because $a' > V^k(w)$, we can find nonnegative vectors X', x', and h' satisfying

$$\big[A - T^k(P)\big]X' - \big[A - S^k(p)\big]x' + Jh' \ge w \quad \text{and} \quad PX' - px' + h'_1 < a'. \tag{3.6}$$

Because $-a'' > -v^k(w)$ and thus $-a'' > V^k(-w)$, we can likewise find nonnegative vectors X'', x'', and h'' satisfying

$$\big[A - T^k(P)\big]X'' - \big[A - S^k(p)\big]x'' + Jh'' \ge -w \quad \text{and} \quad PX'' - px'' + h''_1 < -a''. \tag{3.7}$$

Then the trade $(X, x) = (X' + X'', x' + x'')$ and holdover schedule $h = h' + h''$ have

$$\big[A - T^k(P)\big]X - \big[A - S^k(p)\big]x + +Jh \ge 0 \quad \text{and} \quad PX - px + h_1 < a' - a'',$$

as seen from the addition of (3.7) to (3.6). Unless $a'' < a'$, the WNA^k condition, which excludes (3.1), would be violated. Thus $v^k(w) \le V^k(w)$ as claimed.

In particular, $V^k(w)$ and $v^k(w)$ must be finite. Because these finite values are calculated by linear programming, each must be attained. This is well known in linear programming theory. For the converse part, we note that when WNA^k is not satisfied, unlimited arbitrage is available to every investor, and this results in infinite values for $V^k(w)$ and $v^k(w)$. □

With these considerations it is apparent that although reliance on NPV with respect to a single term structure is untenable in the face of transaction costs, a definite *range* of values can be placed on firm ground for each tax class of investors, under the assumption of WNA. As far as a generic investor in class k is concerned, if an asset producing the after-tax cash stream w is offered at price c it will be considered *overvalued* if $c > V^k(w)$ but *undervalued* if $c < v^k(w)$. Prices in the interval $\left[v^k(w), V^k(w)\right]$ will be regarded as not out of line.

According to definition, the long and short prices $V^k(w)$ and $v^k(w)$ can be calculated by solving certain linear programming problems in X, x and h. We now develop a dual linear programming formula for the long and short prices, which will show that these prices reflect precisely the multiplicity of no-arbitrage term structures that may exist when transaction costs are present and describe the range of NPV values relative to those term structures.

DEFINITION. *For investor class k, the current term structure packet for evaluating after-tax cash streams is the set D^k consisting of all the vectors $d = (d_1, \ldots, d_m)$ such that*

$$
\begin{aligned}
& d_1[a_{ij} - t^k_{1j}(P_j)] + \cdots + d_m[a_{ij} - t^k_{mj}(P_j)] \le P_j \ \text{ for } j = 1, \ldots, n, \\
& d_1[a_{ij} - s^k_{1j}(p_j)] + \cdots + d_m[a_{ij} - s^k_{mj}(p_j)] \ge p_j \ \text{ for } j = 1, \ldots, n, \\
& 1 \ge d_1 \ge d_2 \ge \cdots \ge d_m \ge 0.
\end{aligned}
\tag{3.8}
$$

THEOREM 3.2. *The term structure packet D^k is nonempty if, and only if, the weak no-arbitrage condition, WNA^k, holds. The long and short values $V^k(w)$ and $v^k(w)$ for after-tax cash streams w can then be calculated by*

$$
V^k(w) = \max_{d \in D^k} dw \quad \text{and} \quad v^k(w) = \min_{d \in D^k} dw. \tag{3.9}
$$

On the other hand, D^k is completely determined by knowledge of V^k or v^k:

$$
D^k = \{d \mid dw \le V^k(w) \ \text{ for all } w\} = \{d \mid dw \ge v^k(w) \ \text{ for all } w\}. \tag{3.10}
$$

PROOF. The formulas in (3.9) follow by linear programming duality from the definitions (3.2) and (3.3) of $V^k(w)$ and $v^k(w)$. The formula for $V^k(w)$, for instance, is known to be valid if and only if $V^k(w)$ is finite, and by Theorem 3.1 that is true if and only if WNA^k holds. This formula asserts in the language of convex analysis that the function V^k on R^m is the *support function* of the set D^k; see Rockafellar (1970), §13. Inasmuch as D^k is a closed, convex set by its definition, we may conclude from that general theory that D^k can be recovered from V^k in the manner indicated in (3.10). The case of $v^k(w)$ can be argued similarly, or one can simply invoke the relationship between $v^k(w)$ and $V^k(w)$ in Theorem 3.1. □

The vectors $d \in D^k$ are appropriately called the *no-arbitrage term structures* for investor class k, by virtue of the (equivalent) formulas in (3.10). The first of these formulas says, for instance, that d belongs to D^k if and only if no combination of a trade (X, x) and a hold-over schedule h can secure any cash stream w at a cost lower than the NPV of w relative to d.

The current term structure packet D^k is a convex polyhedron in R^m, because it is defined by a finite system of weak linear inequalities in the variables $d_1, \ldots, d_m$ and thus can be seen as the intersection of a finite collection of closed half-spaces. The formulas in (3.9) show that the imputed long and short prices $V^k(w)$ and $v^k(w)$ can be obtained by linear programming over this polyhedron. If D^k happened to reduce to just a *single* vector d^k, one would have $V^k(w) = d^k w = v^k(w)$ for all w, i.e., the linear form of valuation that has commonly been used until now. But transaction costs effectively prevent this from being the case, as will be seen in the next section.

The idea of nonlinear valuations based on maximizing or minimizing over a convex set of vectors is well understood in the mathematics of optimization. Although we are encountering it here in the context of a basic financial model, it could naturally be expected anywhere that the dual variables associated with the constraints in a problem of optimization might not be unique. Utility models in finance, and more generally in economics, where the dual variables form "price" vectors instead of term structures, fall into this pattern, for instance. Such dual vectors can usually be interpreted as the subgradients of some convex function at a point representing current resources or equilibrium. The right and left derivatives of this function are then determined by formulas exactly like those in (3.9) relative to the set of all such vectors; see §23 of Rockafellar (1970).

4. NONLINEARITIES IN THE VALUATIONS

The properties of $V^k(w)$ and $v^k(w)$ as functions of the cash stream w will now be explored. It will be demonstrated in particular that, under a no-arbitrage condition slightly stronger than WNA^k, both valuation operators are definitely nonlinear relative to trade-offs between different cash stream, although they remain linear to scale. We begin with some properties of $V^k(w)$ and $v^k(w)$ that follow without additional assumptions from the results already obtained.

THEOREM 4.1. *For each tax class k, the upper valuation operator V^k is convex while the lower valuation operator v^k is concave. Both are piecewise linear in general, but linear to scale: one has*

$$\begin{aligned} &V^k(w + w') \leq V^k(w) + V^k(w') \text{ and } v^k(w + w') \geq v^k(w) + v^k(w') \text{ for all } w, w', \\ &V^k(\lambda w) = \lambda V^k(w) \text{ and } v^k(\lambda w) = \lambda v^k(w) \text{ for all } \lambda \geq 0, \text{ with } V^k(0) = v^k(0) = 0. \end{aligned} \tag{4.1}$$

PROOF. The term structure packet D^k, being a nonempty, bounded, convex polyhedron, is the convex hull of its vertex points, of which there are finitely many. Let the vertex elements be denoted by $d^l = (d_1^l, \ldots, d_m^l)$ for $l = 1, \ldots, r_k$. In the linear programming problem presented by the first formula in (3.6), the maximum over D^k relative to any given w will be attained at some vertex point of D^k. It follows that for every w one has $V^k(w) = \max\{d^l w \mid l = 1, \ldots, r_k\}$. Thus, V^k is the maximum of a collection of finitely many linear functions of w. This implies V^k is convex and piecewise linear. In a similar vein, $v^k(w) = \min\{d^l w \mid k = 1, \ldots, r_k\}$, so v^k is concave and piecewise linear. The relations in (4.1) are obvious from these formulations. □

It should be emphasized, to avoid misunderstanding, that although the linearity to scale in (4.1) holds mathematically for arbitrarily large $\lambda > 0$, we are not claiming our model can be applied indiscriminately for such large values. The model is still to be thought of as having only *local* validity around the current price equilibrium. If the trades (X, x) contemplated in the valuation formulas were to get very large, the current prices P_j and p_j could not be trusted, or perhaps the trade could not even be completed.

Another general property worthy of note is the following.

THEOREM 4.2. *The valuation operators V^k and v^k are cumulatively monotone, in the sense that for any pair of cash streams $w = (w_1, \ldots, w_m)$ and $w' = (w'_1, \ldots, w'_m)$ and prices c and c' such that*

$$-c \leq -c' \text{ and } -c + w_1 + \cdots + w_i \leq -c' + w'_1 + \cdots + w'_i \text{ for } i = 1, \ldots, m, \tag{4.2}$$

one has

$$V^k(w) - c \leq V^k(w') - c' \text{ and } v^k(w) - c \leq v^k(w') - c'. \tag{4.3}$$

Furthermore, if strict inequalities hold in (4.2), they hold also in (4.3).

PROOF. For any term structure d in D^k and any cash stream w, we can write

$$\begin{aligned} -c + dw = {} & (1 - d_1)(-c) + (d_1 - d_2)(-c + w_1) + (d_2 - d_3)(-c + w_1 + w_2) + \cdots \\ & + (d_{m-1} - d_m)(-c + w_1 + \cdots + w_{m-1}) + d_m(-c + w_1 + \cdots + w_m). \end{aligned} \tag{4.4}$$

Membership in D^k entails $1 \geq d_1 \geq d_2 \geq \cdots \geq d_m \geq 0$ in particular, so the difference coefficients in (4.4) are all nonnegative. It follows that if w and w' satisfy the inequalities in (4.2) relative to the prices c and c', we have

$$-c' + dw' \geq -c + dw \text{ for all } d \in D^k. \tag{4.5}$$

Taking the maximum on both sides of (4.5) yields $V^k(w) - c \leq V^k(w') - c'$ by the first formula in (3.9). Taking the minimum yields $v^k(w) - c \leq v^k(w') - c'$ by the second formula in (3.9).

If strict inequality holds for every i in (4.2), then it also holds in (4.5) because the nonnegative difference quotients in (4.4) add up to 1 and therefore cannot all vanish. In this case in taking the maximum and minimum we get strict inequality in (4.3). This is true since both are certain to be attained. □

The interpretation of condition (4.2) is that the cash stream w' at price c' is at least as good as w at price c, because it costs no more now, i.e., $c' \leq c$, and yet it delivers at least as much net income as w up to any future time and might even deliver it earlier. Theorem 4.2 says that this kind of superiority is always reflected in our long and short valuations.

To probe further into the properties of D^k, V^k, and v^k, we introduce a sharper condition.

DEFINITION. *The complete no-arbitrage condition CNA^k is satisfied for investors in tax class k when there is no combination of a trade $(X,x) \geq (0,0)$ and a holdover schedule $h \geq 0$, with $(X,x) \neq (0,0)$, such that*

$$\big[A - T^k(P)\big]X - \big[A - S^k(p)\big]x + Jh \geq 0 \quad \text{and} \quad PX - px + e_1 h \leq 0. \tag{4.6}$$

The first inequality states that the investor's after-tax cash flow (incoming) from (X,x) and h is nonnegative in each future period, while the second asserts that the present net cost of putting this together, and thereby moving to a position different from the investor's previous one (since $(X,x) \neq (0,0)$) is nonpositive, i.e., no outlay of cash is required. In forbidding such a trade, CNA^k is more restrictive than WNA^k, which corresponds to having the second inequality in (4.6) be strict. It is more restrictive also than the following.

DEFINITION. *The strong no-arbitrage condition SNA^k is satisfied for class k when there is no combination of a trade $(X,x) \geq (0,0)$ and a holdover schedule $h \geq 0$ satisfying (4.6), except perhaps ones such that*

$$\big[A - T^k(P)\big]X - \big[A - S^k(p)\big]x + Jh = 0 \quad \text{and} \quad px - PX - e_1 h = 0. \tag{4.7}$$

Strong no-arbitrage conditions have been developed and explained in many places, e.g. in Garman (1978) and Dermody and Prisman (1988). The version stated here is slightly different in that it incorporates holdovers h as well as taxes.

In words, WNA^k excludes the possibility of an immediate "free lunch" for an investor in class k, whereas SNA^k excludes the possibility of guaranteeing a "free lunch" to be received at any time, either present or future. Under SNA^k the investor cannot make a trade that brings in positive cash at some time without incurring an outflow of cash at some other time. The CNA^k condition goes further in excluding even a "free trip," where the investor is able to move to a new position without ever paying any money currently or in the future. CNA^k can be seen, like SNA^k and WNA^k, as expressing a level of market equilibrium. It appears justified as an assumption for models with transaction costs, such as in this paper. If CNA^k did not hold, every investor in class k could undertake certain trades without it ever costing anything. This would be a more subtle form of arbitrage in which the gain is not money itself but a new position that otherwise could not be achieved for free.

THEOREM 4.3. *If CNA^k holds, the term structure packet D^k has nonempty interior (and therefore exhibits infinite multiplicity). This interior consists then of the vectors d satisfying*

$$\begin{aligned} &d_1[a_{ij} - t^k_{1j}(P_j)] + \cdots + d_m[a_{ij} - t^k_{mj}(P_j)] < P_j \text{ for } j = 1,\ldots,n, \\ &d_1[a_{ij} - s^k_{1j}(p_j)] + \cdots + d_m[a_{ij} - s^k_{mj}(p_j)] > p_j \text{ for } j = 1,\ldots,n, \\ &1 > d_1 > d_2 > \cdots > d_m > 0. \end{aligned} \tag{4.8}$$

PROOF. Suppose that CNA^k holds. Let D_0^k denote the set of all vectors d satisfying (4.8). We argue first that if D_0^k is nonempty, then D^k has nonempty interior, in which event D_0^k is this interior. Certainly D_0^k is an open set included in the convex polyhedron D^k, so all points of D_0^k belong to the interior of D^k. The assertion that D_0^k, when nonempty, is actually the whole interior of D^k can be justified as follows.

If d is an arbitrary point in the interior of D^k, we can express d relative to any d^0 in D_0^k by $d = (1-\lambda)d^0 + \lambda d^1$, where d^1 is some other point of D^k and $0 < \lambda < 1$. Using the fact that d^0 satisfies the strict inequalities in (4.8), while d^1 satisfies the corresponding weak inequalities which define D^k in (3.8), we can obtain from this representation of d the conclusion that d satisfies the strict inequalities and hence belongs to D_0^k. Thus, the interior of D^k cannot be larger than D_0^k and therefore must equal it.

To finish the proof, it will be enough now to demonstrate that if D_0^k is empty, then CNA^k fails to hold. The nonexistence of a vector d satisfying all the inequalities in (4.8) means the following, when stated in terms of auxiliary vectors $Z = (Z_1, \ldots, Z_m)$ and $z = (z_1, \ldots, z_m) \in R^m$: the set

(4.9) $$B = \{(Z, z, d) \mid Z_j < P_j \text{ and } z_j > p_j \text{ for } j = 1, \ldots, n;\ 1 > d_1 > \cdots > d_m > 0\},$$

which is nonempty by (2.1), has no point in common with the set

(4.10) $$I = \{(Z, z, d) \mid Z = d[A - T^k(P)] \text{ and } z = d[A - S^k(p)]\,\},$$

where in the latter d is an unrestricted vector in R^m. The set B is open and convex, while I is a linear subspace of $R^n \times R^n \times R^m$ (hence convex as well). Therefore, their intersection is empty if and only if they be separated (as explained in Rockafellar (1970, Theorem 11.1)): there exist vectors $X \in R^n$, $x \in R^n$ and $w \in R^m$, not all 0, along with a scalar value c, such that

(4.11) $$\begin{aligned} ZX - zx + dw &\le c \text{ for all } (Z, z, d) \in B,\\ ZX - zx + dw &\ge c \text{ for all } (Z, z, d) \in I. \end{aligned}$$

The second condition in (4.11) can be seen from the definition of I in (4.10) as saying that $d([A - T^k(P)]X - [A - S^k(p)]x + w) \ge c$ for all $d \in R^m$. This is true if, and only if,

(4.12) $$w = -[A - T^k(P)]X + [A - S^k(p)]x \quad \text{and} \quad c \le 0.$$

The fact that X, x and w cannot all be 0 implies then that $(X, x) \ne (0, 0)$. The first condition in (4.11) is equivalent on the other hand, by the definition of B in (4.9), to

(4.13) $$\max_{Z \le P} ZX - \min_{z \ge p} zx + \max_{d \in M} dw \le c, \text{ where } M = \{d \mid 1 \ge d_1 \ge \cdots \ge d_m \ge 0\}.$$

The first maximum in (4.13) is PX if $X \ge 0$, but otherwise it is ∞; similarly, the minimum in (4.13) is px if $x \ge 0$, but otherwise it is $-\infty$. The formula

$$\max_{d \in M} dw = \max_{dJ \ge e_1} dw = \min_{h \ge 0,\ Jh = w} e_1 h$$

holds by linear programming duality, where the maximum and minimum are necessarily attained. Thus the first condition in (4.11), through its expression in (4.13), is equivalent to having

$$X \geq 0,\ x \geq 0,\ \text{ and for some } h \geq 0 \text{ also } PX - px + e_1 h \leq c.$$

By combining this with (4.12), we obtain that (X, x) is a nonzero trade violating the CNA^k condition. □

Note that CNA^k does not follow simply from the inequalities $p_j < P_j$. These could hold and yet, mathematically, D^k could have less than the full dimension m (and in some cases even be just a single point). An immediate consequence of Theorem 4.3 is a strong assertion about V^k and v^k.

THEOREM 4.4. *If CNA^k holds, neither of the valuation operators V^k or v^k can actually be linear, not even relative to a neighborhood of $w = 0$. Furthermore,*

$$v^k(w) < V^k(w) \text{ for all } w \neq 0. \tag{4.14}$$

PROOF. The formulas in (3.9) must always give different answers for $w \neq 0$ when D^k has a nonempty interior. Therefore, by Theorems 3.1 and 4.3, we do have (4.14) under CNA^k. If either V^k or v^k were linear relative to some neighborhood of $w = 0$, it would follow from the second relation in (3.4) that $v^k(w) = V^k(w)$ on the neighborhood in question, in conflict with (4.14). Such linearity is therefore impossible. □

We have seen that WNA^k is equivalent to D^k being nonempty, whereas CNA^k corresponds to D^k being "fat" in the sense of possessing a nonempty interior. Our final result shows where SNA^k fits into this picture.

THEOREM 4.5. *Condition SNA^k holds if, and only if, the term structure packet D^k contains some $d > 0$.*

PROOF. In terms of the valuation theory in Section 3, SNA^k is equivalent to the assertion that the only cash stream $w \geq 0$ with $V^k(w) \leq 0$ is $w = 0$. Using the cumulative monotonicity of V in Theorem 4.2, we can reduce this to the assertion that $V^k(w) > 0$ for all w of the form $(0, \cdots, 0, \lambda)$ with $\lambda > 0$, since $V^k(0) = 0$ by Theorem 4.1. This translates by the positive homogeneity (linearity to scale) in Theorem 4.1 into the simple property $V^k(e_m) > 0$, where $e_m = (0, \ldots, 0, 1)$. The formula for $V^k(e_m)$ in (3.9) identifies this with the existence of a vector $d = (d_1, \cdots, d_m)$ in D^k having $d_m > 0$. The chain of equivalences ties SNA^k to the existence of a vector $d > 0$ in D^k. □

5. PRICE EQUILIBRIUM AND CLIENTELE EFFECTS

The fact that each tax class k of investors has its own valuation operators V^k and v^k does not pose any difficulty in itself for price equilibrium in financial markets. These valuation operators apply to after-tax cash streams, and a traded security can yield a different after-tax cash stream for each class. By Theorem 3.1, the WNA condition is equivalent to the finiteness of the valuations and nothing more. To say it another way, the different ranges of value $\big[v^k(w), V^k(w)\big]$ assigned to an after-tax cash stream w by the various tax classes of investors *definitely do not* lead to the existence of arbitrage, at least not of the kind that an investor can realize individually through a trade (X, x).

Nonetheless there are certain relationships between the different valuation operators and the price vectors $P = (P_1, \ldots, P_m)$ and $p = (p_1, \ldots, p_m)$ that should be satisfied at equilibrium. Our goal in this section is to explore these relationships. We shall see that they provide a basis for clientele effects, where investors in different tax classes exhibit different preferences for long or short positions in certain traded securities. We assume throughout the discussion that WNA holds. Then by Theorem 3.2 the term structure packets D^k are all nonempty and can be used in expressing V^k and v^k. We can focus on the geometry of these packets D^k as a source of insights.

It will be helpful at this stage to have the notation $A_j = (a_{1j}, \ldots, a_{mj})$,

$$T_j^k(P_j) = \big(t_{1j}^k(P_j), \ldots, t_{mj}^k(P_j)\big) \quad \text{and} \quad S_j^k(p_j) = \big(s_{1j}^k(p_j), \ldots, s_{mj}^k(p_j)\big)$$

for the after-tax cash streams associated with security j in long and short positions. The inequalities (3.8) defining D^k can then be expressed in vector terms as $d\big[A_j - T_j^k(P_j)\big] \leq P_j$ and $d\big[A_j - S_j^k(p_j)\big] \geq p_j$, along with $dJ \geq e_1$.

The following interpretation can be made. To construct the polyhedron D^k for investors in class k, we should begin with the simple set

$$M = \{d \mid dJ \geq e_1\} = \{(d_1, \ldots, d_m) \mid 1 \geq d_1 \geq d_2 \geq \cdots \geq d_m \geq 0\}, \tag{5.1}$$

which is common to all classes and is a convex polyhedron with nonempty interior. Then for each security j we cut away from M all the term structures d (if any) in the open half-space $\{d \mid d\big[A_j - T_j^k(P_j)\big] > P_j\}$ and all those in the open half-space $\{d \mid d\big[A_j - S_j^k(p_j)\big] < p_j\}$. The remaining set is D^k.

The boundary hyperplane $\{d \mid d\big[A_j - T_j^k(P_j)\big] = P_j\}$ for a open half-space of the form $\{d \mid d\big[A_j - T_j^k(P_j)\big] > P_j\}$ might or might not touch the resulting D^k. If it does not, D^k lies in the complementary open half-space $\{d \mid d\big[A_j - T_j^k(P_j)\big] < P_j\}$. Then the inequality $d\big[A_j - T_j^k(P_j)\big] \leq P_j$ for this particular j plays no role ultimately in determining D^k. In this case we shall say that *security j is inactive on the long side* for D^k.

Similarly, the boundary hyperplane $\{d \mid d\big[A_j - S_j^k(p_j)\big] = p_j\}$ for one of the open half-spaces $\{d \mid d\big[A_j - S_j^k(p_j)\big] < p_j\}$ that we are excluding from D^k might or might not touch D^k. If it does not, we shall say that *security j is inactive on the short side* for D^k.

DEFINITION. *Security j will be called unattractive long for investors in class k when $V^k\big(A_j - T_j^k(P_j)\big) < P_j$ and unattractive short when $v^k\big(A_j - S_j^k(p_j)\big) > p_j$.*

In the first case of the definition, every investor in class k can obtain an after-tax cash stream at least as good as the one provided by a unit long position in security j by paying less than the current long price P_j. In the second case, every investor in class k can obtain more than the current short price p_j by agreeing to make future payments which, in after-tax terms, net out to a cash stream no worse than the one corresponding to a unit short position in security j. In the first case, investors in class k will not wish to take on a new long positions in security j, while in the second they will not want new short positions.

The next theorem connects these concepts to the geometric picture we have given of the construction of the term structure packet D^k.

THEOREM 5.1. *Security j is inactive on the long side for D^k if, and only if, it is unattractive long for investors in class k. Likewise, security j is inactive on the short side for D^k if, and only if, it is unattractive short for investors in class k.*

PROOF. By its definition, D^k is contained in the half-space $\{d \mid d\big[A_j - T_j^k(P_j)\big] \le P_j\}$. Therefore we always have

$$P_j \le \max_{d \in D^k} d\big[A_j - T_j^k(P_j)\big]. \tag{5.2}$$

The maximum in (5.2), which equals $V^k\big(A_j - T_j^k(P_j)\big)$ by Theorem 3.2, is attained by at least one d in D^k; this follows from the standard theory of linear programming. Only two cases are possible therefore. Either equality holds in (5.2) and at least one point $d \in D^k$ satisfies $d\big[A_j - T_j^k(P_j)\big] = P_j$, or instead one has $P_j > V^k\big(A_j - T_j^k(P_j)\big)$. In the first instance, security j is active on the long side for D^k. In the second instance, security j is unattractive long, by definition. The second half of the theorem is proved in parallel fashion using the v^k formula in Theorem 3.2. □

While a particular security may be unattractive long for certain classes and unattractive short for others, the unattractiveness could not be universal, or the prices would be "out of kilter." If security j were unattractive long for all investors, no one would rationally want to purchase any units of security j at price P_j. This would contradict the idea that P_j is the price at which trades involving long positions in security j are currently going on. In the same way, if security j were unattractive short for all investors, the short price p_j could not be representative of current trading. Some additional terminology will help pin this down and enable us to state a result about prices.

DEFINITION. *The long price P_j will be called unsupported in our model if, at this price, security j is unattractive long for every investor class $k = 1, \ldots, K$. Similarly, the short price p_j will be called unsupported if, at this price, security j is unattractive short for every investor class $k = 1, \ldots, K$.*

The reader should observe that this concept is not purely one of market clearance or price equilibrium. It refers specifically to the investor classes that have been included in our model. If the model failed to include some investor class whose transactions were the ones most responsible for the magnitude of P_j, it might happen that P_j is unsupported in our model in the sense of the definition, even though P_j does reflect current trading.

THEOREM 5.2. *If none of the current prices P_j or p_j is unsupported, then*

$$P_j = \max_{k=1,\dots,K} V^k\big(A_j - T_j^k(P_j)\big) \text{ and } p_j = \min_{k=1,\dots,K} v^k\big(A_j - S_j^k(p_j)\big) \text{ for } j = 1,\dots,n. \tag{5.3}$$

In this case every security j is active on the long side for at least one of the term structure packets D^k, and also active on the short side for at least one (possibly different) D^k.

PROOF. We know from (5.2) and (3.5) that $V^k\big(A_j - T_j^k(P_j)\big) \le P_j$ holds for every k. Strict inequality for a particular k corresponds to security j being unattractive long to investors in class k. Unless the long price P_j is unsupported, this cannot be true for every k. Our hypothesis therefore implies that $V^k\big(A_j - T_j^k(P_j)\big) = P_j$ for at least one k. The first equation in (5.3) is thus correct. The justification of the second equation in (5.3) follows the same lines in terms of the inequality $v^k\big(A_j - S_j^k(p_j)\big) \ge p_j$. □

One can think of the relations (5.3) as a system of nonlinear equations in the current market prices P_j and p_j. In this it is important to recognize, however, that the functions V^k and v^k on R^m depend in themselves on these prices. Really one could write $V^k(P_1,\dots,P_n;p_1,\dots,p_n;w)$ instead of just $V^k(w)$, and similarly for v^k. From this perspective, the equations would be written as

$$\max_{k=1,\dots,K} V^k\big(P_1,\dots,P_m;p_1,\dots,p_m;A_j - T_j^k(P_j)\big) - P_j = 0 \text{ and}$$
$$\min_{k=1,\dots,K} v^k\big(P_1,\dots,P_m;p_1,\dots,p_m;A_j - S_j^k(p_j)\big) - p_j = 0 \text{ for } j = 1,\dots,n.$$

Such relations are close in character to the main result of Dybvig and Ross (1986, Theorem 3), consisting actually of a system of highly nonlinear (and nondifferentiable) market price equations for currently traded assets, although not presented as such. Those authors worked in quite a different context of risky payments in a single time period, however. They excluded shorting and relied on cost-free buying and selling.[5] In deriving their equations through consideration of utility functions, they did not arrive at a pricing rule for general cash streams w.

We can summarize the results so far as follows, under the assumption that none of the prices in our model is unsupported. Let

$$\begin{aligned} K_j^+ &= \text{ set of investor classes } k \text{ attaining the max in (5.3)}, \\ K_j^- &= \text{ set of investor classes } k \text{ attaining the min in (5.3)}. \end{aligned} \tag{5.4}$$

---

[5] Dybvig and Ross suggest that shorting could be covered by introducing “artificial securities” into their model, but if so, such securities would need to have negative prices—which they explicitly forbid.

For each security j, the investor classes $k \in K_j^+$ are the ones such that security j is active on the long side of D^k, while the investor classes $k \in K_j^-$ are the ones such that security j is active on the short side of D^k. The investor classes $k \notin K_j^+$ are the ones for which security j is unattractive long, while the investor classes $k \notin K_j^-$ are the ones for which security j is unattractive short.

In many discussions of clientele effects, it is taken for granted that an investor who is definitely adverse to taking a long position in security j must definitely be interested instead in a short position in that security. That fails to be true when transaction costs are brought into the picture. The gap between P_j and p_j allows for an "indifference band" between wanting to long or short security j.

These observations apply to *new* long and short positions, not to the question of whether the investors in tax class k will hang on to *previously acquired* positions in the light of current prices. That question is rather different and cannot be answered without going into some detail on the nature of an individual investor's holdings in each security. The holdings in security j may have been acquired at different times at different prices, and these earlier prices are important in determining the tax consequences. We should not think of liquidating an earlier position for present advantage unless, as part of the transaction, the remaining cash stream corresponding to that position is resecured in some essentially riskless manner.

Let us denote by P'_j the current bid price for security j, in comparison with its ask price P_j. This is the amount of present cash received from the market for liquidating a unit long position in security j. Let us similarly denote by p'_j the amount of present cash that currently must be paid for liquidating a unit short position in security j.

For any unit long position in j there is a current *basis* of value for tax purposes, which may be computed by tax rules from the price at which the particular position was obtained and the payments that have been made since. If this current basis value is Q_j, the future after-tax cash stream from the position is $A_j - T_j^k(Q_j)$, just as if the position were purchased in the present for the price Q_j instead of P_j. Similarly, any unit short position in j has a current basis value q_j such that the the future after-tax cash stream representing the payment obligations on the position is $A_j - S_j^k(q_j)$. But the basis values can vary among the long and short units of security j in an investor's holdings; there is not necessarily just one Q_j or q_j.

It will be supposed that for investors in class k the tax rate on capital gains is τ_k, where $0 \le \tau_k \le 1$. If such an investor sells a unit long position in security j for which the basis is Q_j, the difference $P'_j - Q_j$ is a capital gain (or, if negative, in effect a capital loss). For simplicity, we shall regard the tax on this capital gain as being deducted immediately from the proceeds, although a more complicated view could be taken in adding these future payment dates (four times per year in the United States for many investors) to the model. The net proceeds from liquidating the position are then

$$P'_j - \tau_k(P'_j - Q_j) = (1-\tau_k)P'_j + \tau_k Q_j.$$

By a parallel analysis, the net amount paid from liquidating a unit short position with basis q_j is

$$p'_j - \tau_k(p'_j - q_j) = (1-\tau_k)p'_j + \tau_k q_j.$$

THEOREM 5.3. *At current prices, no investor in class k should want to continue holding a previous long position in security j for which the current unit basis is Q_j, if*

$$V^k\big(A_j - T^k_j(Q_j)\big) < (1-\tau_k)P'_j + \tau_k Q_j. \tag{5.5}$$

Likewise, no investor in class k should want to continue holding a short position in security j with current unit basis q_j, if

$$v^k\big(A_j - S^k_j(q_j)\big) > (1-\tau_k)p'_j + \tau_k q_j. \tag{5.6}$$

PROOF. The after-tax cash stream $w = A_j - T^k_j(Q_j)$ provided by a unit long position in security j for which the current basis is Q_j can be replaced at the imputed long price $V^k(w)$. If this amount is less than the net amount the investor can receive in the present for liquidating the position, which is the quantity on the right side of (5.5), the investor should sell the unit, replace its cash stream, and pocket the price difference. In the case of a short position in j with basis q_j, the after-tax cash stream $w' = A_j - S^k_j(q_j)$ gives the remaining payment obligations, but in the current market an investor can obtain the imputed short price $v^k(w')$ for taking on these obligations. If this is greater than the net amount that must be paid to liquidate, the investor should get rid of the former short position, replacing it by a new equivalent, and take the difference as income. □

Note from Theorem 5.3 that despite the possible unattractiveness of a new long position in security j, an investor in tax class k could rationally hang on to an old long position because of the tax effects of the basis, and the same is true for a short position. Therefore, in analyzing the effects of taxes on financial markets, broad assertions about which classes of investors will *keep* which securities j may not be justifiable. On the class-wide level of generality and without delving into specific data on basis values, one must be content merely with knowing which securities are currently not unattractive, long or short.

REFERENCES

1. Carelton, W. and I. Cooper (1976). "Estimation and Uses of the Term Structure of Interest Rates." *Journal of Finance* **31**, pp. 1067–1083.
2. Dammon, R.M. and R.C. Green (1987). "Tax Arbitrage and the Existence of Equilibrium Prices for Financial Assets." *Journal of Finance* **42**, pp. 1143–1166.
3. Dermody, J.C. (1988). "Equilibrium in Government Bond Markets with Taxes: A Reexamination of Schaefer's Model Versus Litzenberger and Rolfo's Model." Office for the Study of Futures and Options Working Paper, University of Illinois at Chicago.
4. Dermody, J.C. and E.Z. Prisman (1988). "Term Structure Multiplicity and Clientele in Markets with Transaction Costs and Taxes." *Journal of Finance* **43**, pp. 893–911.
5. Dermody, J.C. and R.T. Rockafellar (1987a). "A New Approach to Term Structure for Markets with Shorting Costs." Working paper, Dept. of Finance, Arizona State University.
6. Dermody, J.C. and R.T. Rockafellar (1987b). "Tax Effects on Term Structure in Markets with Shorting Costs." Working paper, Dept. of Finance, Arizona State University.
7. Dybvig, P.H. and S.A. Ross (1986). "Tax Clienteles and Asset Pricing." *Journal of Finance* **41**, pp. 751–763.
8. Hodges, S.D. and S.M. Schaefer (1977). "A Model for Bond Portfolio Improvement." *Journal of Financial and Quantitative Analysis* **12**, pp. 243–260.
9. Jordan, J. (1984): "Tax Effects in Term Structure Estimation." *Journal of Finance* **39**, pp. 393–406.
10. Litzenberger, R.H. and J. Rolfo (1984). "An International Study of Tax Effects on Government Bonds." *Journal of Finance* **39**, pp. 1–21.
11. McCulloch, J.H. (1975). "The Tax Adjusted Yield Curve." *Journal of Finance* **30**, pp. 811–930.
12. Rockafellar, R.T. (1970): Convex Analysis. Princeton: Princeton University Press.
13. Ronn, E.H. (1987). "A New Linear Programming Approach to Bond Portfolio Management." *Journal of Financial and Quantitative Analysis* **22**, pp. 439–466.
14. Ross, S.A. (1976). "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory* **3**, pp. 343–362.
15. Ross, S.A. (1978). "A Simple Approach to the Valuation of Risky Streams." *Journal of Business* **51**, pp. 453–475.
16. Ruback, R.S. (1986). "Calculating the Market Value of Riskless Cashflows." *Journal of Financial Economics* **15**, pp. 322–339.
19. Schaefer, S.M. (1981). "Measuring a Tax-Specific Term structure of Interest Rates in the Market for British Government Securities." *The Economic Journal* **91**, pp. 415–438.
20. Schaefer, S.M. (1982a). "Taxes and Security Market Equilibrium." *Financial Economics: Essays in Honor of Paul Cootner*, ed. by W.F. Sharpe and C.M. Cootner. Englewood, NJ: Prentice-Hall, pp. 159–177.
21. Schaefer, S.M. (1982b). "Tax Induced Clientele Effects in the Market for British Government Securities." *Journal of Financial Economics* **10**, pp. 121–159.