GEORGE MASON UNIVERSITY
SCHOOL OF LAW

TAX SHELTERS, DUTCH BOOKS, AND THE FUNDAMENTAL THEOREM OF ASSET PRICING

Terrence R. Chorvat, George Mason University School of Law

07-21

Virginia Tax Review, Forthcoming

GEORGE MASON UNIVERSITY LAW AND ECONOMICS RESEARCH PAPER SERIES

This paper can be downloaded without charge from the Social Science Research Network at http://ssrn.com/abstract_id=989825
Tax Shelters, Dutch Books, and the Fundamental Theorem of Asset Pricing

Terrence R. Chorvat*

Abstract

Because the tax law makes many distinctions not based on fundamental economic differences, taxpayers can exploit these inconsistencies to create opportunities for tax arbitrage. This article argues that any interpretative system which requires consistency in the application of rules where the system itself is based on fundamental inconsistencies will always allow for arbitrage. This result applies equally to purposive interpretations of the tax law as well as to more formal or literal systems of interpretation. That is, the problem of tax arbitrage is concomitant with the existence of non-economic principles of tax law embedded in a system that interprets these rules in a consistent manner. While this idea has been applied to questions of whether a tax system can impede the ability of the economy to reach equilibrium, it has not been applied to the legal analysis of tax shelters themselves. The result of this line of reasoning is that tax shelters are the inevitable result of any tax system with inherent inconsistencies which attempts to bring an artificial consistency to these non-economic based distinctions. The article goes on to argue that one of the methods by which tax shelters are addressed by the courts is the use of inconsistent methods of applying the law to these transactions.

The current U.S. income tax has many examples of inconsistent treatment of essentially identical transactions. For example, under the realization doctrine one can exchange a portfolio for one with essentially the same risk and expected return but which is composed of different securities and recognize a loss for tax purposes. If instead one continues to hold the same portfolio, the loss is not recognized\(^1\) and the present value of the tax benefit of the losses is reduced.\(^2\) It is the thesis of this article that inconsistencies

---

* Associate Professor, George Mason University School of Law; M.S. candidate, Department of Mathematics, The University of Chicago. The author would like to thank Niels Nygaard and the participants of the Symposium for their comments on this article.

1 Cottage Savings Association v. Commissioner, 499 U.S. 554 (1991). This case involved an exchange of portfolios of mortgages. The portfolios of mortgages exchanged were designed so that the risks and returns were as close as possible to each other while the mortgages themselves involved different debtors. The court held that, because the portfolios involved different mortgages, the exchange constituted a realization event.

2 For a discussion of the present value of taxes see Myron Scholes et al., Taxes and Business Strategy 58-74 (3rd ed., 2004). For a discussion with specific reference to the realization doctrine see
such as this allow tax shelters to exist. If instead all income was taxed according to economic income and loss, it would essentially be impossible to create tax shelters.

One can classify tax shelters into two types: transactions which are arguably consistent with current tax law and those which clearly are not. This article will focus on those tax shelters which can be argued to be consistent with the rules. For those transactions that plainly violate the rules, the important questions revolve around the best method to encourage taxpayers to follow the tax law. This may require greater enforcement activities, or it may require a different approach to enforcement.\(^3\) In any case, the questions relevant to these tax shelters are not addressed in this paper.

While there are a variety of definitions of tax shelters, the one used in this article is a transaction or series of transactions whose purpose is to reduce the taxes paid on the gain from other transactions. That is, these transactions shelter from tax income that has been generated from other sources. Such shelters present a problem for the tax system because of the economic waste that results from the costs associated with creating and entering into such transactions, any potential loss from the distortion of economic incentives, and of course the concomitant decrease in government revenue.

The tax law has adopted a number of approaches to such "shelters", not all of which are consistent (e.g., the economic substance doctrine).\(^4\) This article argues that it is the inconsistency of these doctrines which prevents such shelters from gaining greater prominence. That is, the fundamental inconsistency in the tax rules has led to a matching parallel inconsistency in the legal doctrine which addresses tax shelters. The natural


\(^4\) Discussed in part III, *infra*. 

---

conclusion of this line of argument is that there is likely no potential comprehensive consistent strategy to deal with tax shelters which can be developed from the rules themselves. While it may be the case that some alterations to the rules may be helpful to reduce the number of tax shelters by making it more difficult to enter into some transactions, I would claim that any attempt to address the problem through a method of interpretation of the rules in a consistent manner is highly unlikely to succeed. However, by the adoption of a strategy which allows for the appropriate form and measure of inconsistency, the system can and has been able to continue to function.

One of the assumptions of this article is that the government will continue to utilize a tax which is more or less consistent with an income tax as a major source of revenue. That is, we will continue to have a tax system where gains are taxed and losses are permitted to offset gains on other transactions. While there may be limits on the utilization of losses, this article assumes that these limits will clearly be the exception to the rule.5

The article will proceed in three parts. The first part discusses the connection between the notion of "Dutch books" and the theorem commonly referred to as the Fundamental Theorem of Asset Pricing. This part discusses the manner in which the inconsistency of preferences or of prices results in the potential for arbitrage. The second part of the article connects these ideas to tax shelters by explaining that tax shelters are a form of arbitrage. This part goes on to demonstrate that a tax system can easily allow for tax arbitrage. This is based on the proposition that, for any income tax system in which items that are economically identical are taxed at different rates and not all persons are

subject to the same tax rules, there will always be the potential for arbitrage of the tax system. While this proposition has been applied to the question of whether the tax system affects the financial market's ability to reach equilibrium, it has as yet not been applied to the analysis of the optimal legal approach to tax shelters. Finally, Part III discusses why we do not observe arbitrage schemes commonly operating in the world and argues that the tax system already utilizes the same mechanisms which individuals use to avoid being arbitraged in order to prevent the spread of tax shelters. In particular, there are two main methods. The first is the operation of market pricing mechanisms. The second is the adoption of matching inconsistent preferences to place limitations on arbitrage. The article concludes by demonstrating the manner in which the tax system addresses tax shelters essentially mirrors the methods by which individuals avoid being arbitraged.

I. DUTCH BOOKS AND THE FUNDAMENTAL THEOREM OF ASSET PRICES

A. Dutch Books and Arbitrage

One important requirement of the standard analysis of general equilibrium is that there are no arbitrage opportunities.\(^6\) That is, for a general equilibrium to exist under standard economic analysis, it must not be possible for markets to be arbitraged. Because we do not observe significant levels of arbitrage, it is generally assumed that markets operate in a fashion that prevents arbitrage opportunities from arising.\(^7\) In particular, it is assumed that the preferences of individuals are such that arbitrage is generally not

---

\(^6\) ANDREU MAS-COLELL ET AL. 702 MICROECONOMIC THEORY (1995) (The standard proofs of the existence of general equilibrium under uncertainty assume it must be the case that there are no arbitrage opportunities).

\(^7\) Id, at 699-708 (describing the conditions necessary for the existence of a so-called Radner equilibrium, or an equilibrium in asset markets under uncertainty).
possible. A cornerstone of this analysis is the so-called Dutch book argument. The argument runs that, where an individual holds inconsistent preferences, there will always be a possible Dutch book or series of transactions which will generate a profit from the inconsistent preferences of the individual, without actually improving the welfare of the individual. The basic notion of a Dutch book can illustrated using a simple example. Imagine that there are three possible goods: A, B, and C. If there is an individual (Individual 1) whose preferences can be represented as $A \succ B \succ C \succ A$, with some minimal increment in between each step (say $1$), then (without loss of generality) if Individual 1 possesses A and some additional assets, another individual (Individual 2) could arrange a series of transactions to derive a profit from Individual 1 without transferring anything of value. The series of transactions would begin with Individual 2 transferring good C to Individual 1 in exchange for good A and $1$, and then transferring good B to Individual 1 in exchange for C for and $1$. Finally, Individual 2 would then transfer good A back to Individual 1 in exchange for good B and $1$. In this case, Individual 1 will be back to owning A, but will have $3$ less wealth.

This is a simple example of a Dutch book. The standard conclusion from this line of analysis is that individuals must not have inconsistent preferences, or else we would commonly see the existence of such arrangements.

B. Inconsistency and Arbitrage

---

8 HAL VARIAN, MICROECONOMIC ANALYSIS 393 (1992)
10 The notation $A \succ B$ means that good A is strictly preferred to good B.
11 This example assumes that the individual's preferences are constant throughout the series of exchanges.
12 MAS-COLELL ET AL., supra note 6, at 181.
The standard definition of arbitrage is the interposing of transactions which generate a current positive cash flow but which have a value of zero at all times in the future. That is, in a simple two period model, an arbitrage generates a positive cash flow in period one and no net positive or negative flow in the second period. Effectively, a Dutch book strategy is a type of arbitrage. That is, it is series of transactions by which one can make a profit without transferring any real value or taking on any risk.

Under what is referred to as the fundamental theorem of asset pricing, where prices are "consistent" in the manner described below, there is no possibility of arbitrage of the asset market and, where prices are "inconsistent," it will necessarily follow that there will be a possibility for arbitrage. This is essentially another way of stating the Dutch book theorem. Before discussing the theorem, I will first explain some of the terms used. A probability measure is a method of assigning probabilities to a set of potential states of nature such that the sum of all of the probabilities sums to one and none of the probabilities is below zero. In addition, probabilities must follow some simple rules such as additivity and other properties of a measure space. Finally, a risk neutral valuation is one which

13 This is equivalent to having a current value of zero, but having a positive probability of a positive value in the future. TOMAS BJORK, ARBITRAGE THEORY IN CONTINUOUS TIME 7 (2nd ed., 2004)
14 This could also be reversed, (no cash flow in the current period, but positive cash flow in the future which exceeds the time value of money) because this can be turned into an arbitrage by going short a risk-free asset in period one (and receiving a payment in period one) an amount exactly equal to the amount of the guaranteed payment in period two.
15 One can think of taking on risk as a form of transferring value, because in general individuals seek to avoid risk.
16 BJORK, supra note 13, at 422 defines probability measure. For additivity and other properties of a measure space, see id. at 401.
17 (i.e., the probabilities are such that if \( \theta_i \) and \( \theta_j \) are mutually exclusive states of the world then \( P(\theta_i \text{ or } \theta_j) = P(\theta_i) + P(\theta_j). \))
assigns a price to an asset equal to the expected value of the payoffs of that asset in each of the possible states of the world which have non-zero probability.\(^\text{18}\)

The theorem states that, if that there exists a risk-free asset (we can denote the corresponding risk free interest rate by \( r \)), then the market is arbitrage free if and only if there exists a probability measure \( Q \) such that \( S_0 = \frac{1}{1 + R} E^Q[S_1] \), where \( S_0 \) is the price of the asset at time zero, and \( E^Q[S_1] \) is the expected price of the asset given the probability measure \( Q \).\(^\text{19}\)

One consequence of this theorem is that, if one can find a way to allocate probabilities in a manner which follows the standard rules described above, then a set of prices derived from a risk-neutral valuation based on these probabilities is not capable of being arbitraged and, conversely, if the market is not capable of being arbitraged, there is at least one set of probabilities (although there may be possibly infinitely many if the market is incomplete) that one could use that would give the observed prices via a risk-neutral valuation. For example, if we have a pricing system which has a positive risk-free rate of return (which again we denominate \( r \)) then, if the price of an asset today is $100 while the value of the asset tomorrow will be either $90 or $80, the system will permit arbitrage because there is no probability measure consistent with this system.\(^\text{20}\) If the price in the first period was instead between \( \frac{80 + \varepsilon}{1 + r} \) and \( \frac{90 - \varepsilon}{1 + r} \), then the price system would not be capable of being arbitraged because one could assign non-zero

\(^{18}\) BJORK, supra note 13, at 8.

\(^{19}\) The exposition of the theorem is adapted from BJORK, supra note 13, at 29.

\(^{20}\) It is not possible to have a solution to the system \( 100 = 90a + 80b \) where \( a/(1+r) + b/(1+r) = 1, a > 0, b > 0 \), and \( r > 0 \).
probability to each state which could result in such a price consistent with a risk-neutral valuation.

Another way of thinking of this is that, if it is not possible to derive a consistent set of probabilities which generate the observed prices as the expected value, then it is possible to arbitrage the price system. This is similar to the Dutch book idea that if preferences are not consistent, it will be possible to arbitrage these preferences.²¹

To further illustrate how arbitrage can result from inconsistent prices, it is helpful to introduce the concept of replication. An asset can be replicated if its payoffs are a linear combination of the payoffs of other assets in the market. To illustrate how this could occur, imagine that there are two possible states of nature, both of which have strictly positive probability. Imagine further that there is an asset (Asset 1) which has a payoff of 1 in the first state of the world and zero in the second. This can be represented by the payoff vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We will denominate the price of asset one as $P_1$. There is also a second asset (Asset 2) which has a payoff vector of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and a price of $P_2$.²² Let us assume that we have a third asset (Asset 3), which has a payoff vector of $\begin{bmatrix} 5 \\ 3 \end{bmatrix}$ with a price of $P_3$. The price of Asset 3 will need be equal to the same as five times $P_1$ plus three times $P_2$ in order to avoid arbitrage.²³ If the relative price level differs from this, then there

---

²¹ If there are only a finite number of possible choices, this is the same as saying that there is no method by which one could assign a utility value to the goods in the market. Mas-Colell et al., supra note 6, at 8-9.

²² Asset 1 and 2 are a type of asset commonly referred to as Arrow-Debreu securities, which payoff a value of 1 in one state of nature and zero in all others. One can think of the value of these securities as the risk-neutral probability assigned to the state of the world where each of these assets pays off.

²³ It will always be the case that if there are more securities than states of nature, then at least one asset will be replicated by a linear combination of some of the other assets.
will be the possibility of arbitrage. One can show this directly \(^\text{24}\) or as a consequence of the theorem, because only prices consistent with risk-neutral valuation are arbitrage-free.

As long as all individuals have the same valuation for all asset payoffs, it will be possible to derive an equilibrium set of prices. For any potential payoffs of the assets, there exist risk neutral prices for an equilibrium valuation.\(^\text{25}\) If not all persons have the same relative valuation of payoffs, it may be possible that there is no set of equilibrium prices consistent with all valuations of all persons. To illustrate this in the context of the example above, we first note that Individual 1 has the payoff vector for Asset 3 of \(\begin{bmatrix} 5 \\ 3 \end{bmatrix}\). If we assume that there is another person in the market whose payoff vector for asset three is \(\begin{bmatrix} 5 \\ 4 \end{bmatrix}\), there would be no set of prices under which an equilibrium between the two persons could be reached. That is, there would no set of prices which could simultaneously satisfy the two equations, (unless either \(P_1 = 0\) or \(P_2 = 0\) which, under our assumptions, is not possible)\(^\text{26}\):

\[
5P_1 + 3P_2 = P_3 \\
5P_1 + 4P_2 = P_3
\]

One can see that in order to ensure that equilibrium is possible, the prices of assets must be consistent with the payoffs of the assets and the payoffs must be consistent

\(^{24}\) For example if the price of asset one is .5 and the price of asset two is .5, then if the price of asset three is 4.7, then an individual could short asset three and receive 4.7 in the current period, and then have a payoff of either -4 or -3 in the next period. However, the individual could buy five of asset one and three of asset two in the first period, which would cover his position in period two. This hedge would cost him 4. He would then have generated .7 of net current cash flow and no possibility of a negative cash flow in the next period.

\(^{25}\) That is for any matrix of payoffs, one can multiply this by a vector of probabilities which are consistent with the rules of a probability measure to derive a risk-neutral set of prices.

\(^{26}\) Because we have assumed that the probability of each of the states is strictly greater than zero, then if the prices of either asset one or asset are zero, then are simpler arbitrage strategies.
across individuals. If both conditions are not met, it may be possible to derive arbitrage profits from the system.

II. TAX SHELTERS AND TAX ARBITRAGE

As defined above, tax shelters exist if a person that has already earned taxable income arranges a series of transactions which generate a tax benefit without generating any significant economic losses or incurring any significant economic risk. Because the profit generated from the transactions is either entirely or almost entirely generated at the expense of the tax system, this form of arbitrage is often referred to as tax arbitrage.

There are two separate theorems which tie the idea of an arbitrage-free market to the situation where taxes are imposed on the payoffs. Together, these theorems demonstrate that where different tax rates are applied to economically identical transactions and not all persons are subject to the same rates, there will always exist the potential for tax arbitrage.

To discuss the first theorem, it is helpful to clarify the meaning of a few additional terms. For purposes of this discussion, the matrix $X$ is the matrix of payoffs of the different securities in the different states of nature. If there are $m$ states of nature and $n$ different securities, the payoff matrix is an $m \times n$ matrix, where each column represents the payoff of a particular asset and each row represents the payoffs of all assets in particular state of nature. The matrix $\bar{X}$ is the matrix corresponding to $X$ which gives the level of taxable income that corresponds to each of the payoffs. The parameter $t$ is the tax level, which is assumed to be constant (of course, we could vary this rate by varying $\bar{X}$ appropriately). The vector $\hat{w}^i$ is a vector of portfolio weights for the $i$th investor,
where each entry in the vector, $\hat{w}_j^i$, represents the portfolio weight of the $j$th asset in the portfolio of the $i$th investor. We can see that $X\hat{w}^i$ is the vector of the pre-tax payoffs of the portfolio of the $i$th investor in the various states of nature. Therefore, $(X - \bar{X}_i)\hat{w}^i$ is the vector of the after-tax payoffs of the portfolio. Finally, a semipositive payoff vector is one that is never negative in any state of the world but is positive in some states.

The theorem states that there exist "no-tax-arbitrage" prices for an economy with constant marginal tax rates $T_i(\cdot) = t_i, i = 1, ..., l$, if and only if there is no set of asset holdings $\hat{w}^i \in R^n, i = 1, ..., l$, solving the following system of inequalities:

$$(X - \bar{X}_i)\hat{w}^i \geq 0, i = 1, ..., l$$

$$\sum_{i=1}^{l} \hat{w}^i = 0,$$

where, for at least one $i$, the above inequalities are semipositive.28

Rephrased, this theorem states that tax system is arbitrage-free if it is not possible for all of the positions held by all investors to net to zero, but yet the net total after-tax value of these assets is positive. For example, if it is possible for one individual’s position to exactly balance another individual's position,29 yet the net after-tax value of the two portfolios is positive, then tax arbitrage is possible. Indeed, only the possible source of this positive value would be the tax system. One could view this theorem as

27 See Appendix I.
28 Dammon & Green, supra note 5, at 1155.
29 For example, if one individual is long 100 shares of a stock and the other individual is short 100 shares of the same stock, their positions net to no net ownership of stock.
simply restating the "no arbitrage" condition. However, when read together with the next theorem, we can see that the result is more than a simple restatement of arbitrage.

The second theorem states that if an asset's payoff can be replicated by a portfolio of other assets, and if the replicating portfolio has different tax results than the replicated asset, it may be possible that one can find a portfolio \( \hat{\omega} \) that simultaneously satisfies both conditions of the theorem. That is, because of the different tax treatment, one could acquire a portfolio with positive tax characteristics and sell a portfolio with identical payoffs but with less beneficial tax attributes and derive a net benefit from the tax system.\(^{30}\) However, in order for this to occur, the prices of the asset must be such that they do not entirely reflect the benefit derived by the taxpayer. Such will be the case if not all persons in the market are subject to the same tax rules. In this case, the after-tax payoffs of individuals will differ and, as discussed above, arbitrage will then be possible.

To illustrate the application of the results of these theorems, let us use a modified version of the earlier example in which there are three securities and two states of the world.\(^{31}\) We can represent the pre-tax payoff of all three securities as a matrix

\[
\begin{pmatrix}
1 & 0 & 5 \\
0 & 1 & 3
\end{pmatrix}
\]  

\(^{32}\) If we assume that the market behaves such that both states are equally likely, then \( P_1 \) will be \$.50 and \( P_2 \) will be \$.50 as before.\(^{33}\) As discussed earlier, \( P_3 \) will then have to be \$4. To include the effects of the tax system we introduce the matrix \( X \).

To do this, let us have a very simple tax system in which the gains to Asset 1 are taxed,


\(^{31}\) Such a system will by definition require that one of the securities is a linear combination of the other securities.

\(^{32}\) Here, this means that the payoff of asset three in state 1 is five times that of its price.

\(^{33}\) That is, a risk-neutral person would value a payoff of \$1 with a 50% chance of occurring as \$.50.
but losses are not and gains and losses on Assets 1 and 2 are not taxed at all. The entries
in this matrix $\overline{X}$ are the amounts of the payoffs of the assets that are subject to tax. That
is, the matrix of the contribution of the payoffs to taxable income would be
$$\begin{bmatrix}
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}. \quad 34$$

Note here that the payoff of Asset 3 is treated less favorably for tax purposes than the
payoff of five times the payoff of Asset 1 plus three times the payoff of Asset 2. Let us
further imagine that, if the tax rate is 50%, then the after tax payoff matrix $[X - t\overline{X}]$ is
then
$$\begin{bmatrix}
1 & 0 & 5 \\
0 & 0 & 1 \\
0 & 1 & 3
\end{bmatrix} - \begin{bmatrix}
-0.5 & 0 & 1 \\
0 & 0 & 0 \\
0 & 1 & 3
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 4.5 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}. \quad \text{In this case, the equilibrium set of prices must be such that } P_3 \text{ would need to be 4.5 times } P_1 \text{ plus 3 times } P_2. \text{ If all persons in the market are taxed in the same manner, then arbitrage will still not be possible, because all persons would face the same after-tax returns. However, if there is a person in the market who is not taxed and whose payoff matrix is still } \begin{bmatrix}
1 & 0 & 5 \\
0 & 0 & 1 \\
0 & 1 & 3
\end{bmatrix}, \text{ there is no set of prices which is consistent with the valuations of both persons. That is, there will not be a set of prices (where } P_1 \neq 0, P_2 \neq 0) \quad 35 \text{ which satisfies both of the equations:}

$$
\begin{align*}
5P_1 + 3P_2 &= P_3 \\
4.5P_1 + 3P_2 &= P_3
\end{align*}
$$

If the prices that prevail in the market are inconsistent with a particular person's
payoff matrix, then that person could arrange the portfolio weights such that both the net
amount invested in a portfolio is zero and the after tax level of the returns are greater than

---

$34$ The upper right hand column entry is 1 because the payoff in that case is 5 and the price of the asset was 4.
$35$ This is not acceptable as a price for either asset 1 or asset 2 if we assume that the probability of the both states of nature is strictly greater than zero.
zero. One can also note that if $\bar{X} = aX$ (that is, it is simply a proportional representation of $X$), and if prices in the market are consistent with the pre-tax prices, then they will be consistent with after-tax prices, because the two equations will be multiples of each other.

Based on these theorems, one can conclude that, if the tax system has inconsistent rules concerning the inclusion of income, there will always be the possibility of tax arbitrage. Even though the theorem requires that there exist not merely inconsistency across asset payoffs, but also inconsistency across persons, this requirement is almost trivial because there are numerous tax-exempt entities and markets generally have large numbers of foreign participants not subject to domestic tax rules.

Although these theorems are stated with respect to linear tax rates, it turns out that the results apply to non-linear schedules as well. More specifically, it can be demonstrated that arbitrage can also exist in the context of progressive rates. Depending on the structure of the tax system, progressive tax rates may limit the scope of arbitrage because, as taxable income falls as a result of arbitrage, the rates of all persons become more closely aligned.\textsuperscript{36}

Until now we have assumed that one can always replicate an asset by a portfolio of other assets. This is somewhat of a strong assumption. If it is not true that all asset payoffs can be replicated by combinations of other assets, the market is not complete. Indeed, real asset markets are not complete.\textsuperscript{37} However, what is required for the existence of tax arbitrage is not completeness of the market, but rather the existence of replicating portfolio for some assets which are not taxed at the same level as the asset

\textsuperscript{36} Dammon & Green, note 5 supra at 1156-58, see also Schaefer, note 30 supra at 776-77.

\textsuperscript{37} Schaefer, supra note 30, at 774 points out that in general equity markets do not allow for the replication of all assets and so are not complete. Schaefer goes on to argue that bond markets are closer to being complete.
If the degree to which an asset can be closely approximated by a portfolio exceeds the inconsistency, arbitrage will still be possible. Therefore, the degree of inconsistency allowable in the tax rules is limited by completeness of the market. The more complete the market is, the greater are the limits on the potential inconsistency of the tax system.

III. Tax Arbitrage and Inconsistency of the Tax Rules

The preceding sections show that if there are inconsistencies in the tax rules, then the system can admit arbitrage. There are many ways in which the current tax system treats similar transactions in a distinctly different manner. One of the most commonly arising inconsistencies is that created by the realization doctrine. Under this doctrine, income is not taxed until it has been "realized", which generally requires some kind of sale or other completed transaction. Because the income from an asset which has been sold is subject to a higher rate of tax than one which has not been sold, the gains from these assets are subject to different rates of tax, even though the assets may have exactly the same risk and returns. Under the standard view, an increase in value of $10 which is unrealized has the same economic value as a $10 realized gain. Therefore, unrealized gain is treated more favorably than economically identical realized income. Furthermore, there are also inconsistencies that result from the taxation of capital gains as compared to

---

38 A market can be incomplete as still have some assets that can be replicated by other assets. Dammon and Green, note 5 supra at 1154
39 Dammon & Green, note 5 supra at 1155-56
40 Dammon & Green, note 5 supra at 1155-56.
41 See BORIS I. BITTKER & LAWRENCE LOKKEN, FEDERAL TAXATION OF INCOME, ESTATES AND GIFTS ¶ 5.2, AT 5-17 (3D ED. 1999)
42 Constantinides, note 2 supra at 611.
The division of income into ordinary income and capital gain does not have any significant economic meaning.\textsuperscript{44} Dividends which are received on shares of stock have traditionally been treated as ordinary income,\textsuperscript{45} while the taxation of retained earnings that increase the value of a share of stock are taxed at capital gains rates on the sale of the shares. Therefore, the same earnings are taxed at different rates depending on whether they are distributed to the shareholders or retained within the corporation.\textsuperscript{46} Another important non-economic distinction is the difference between treatment of the income of a corporation which is attributable to claims of debt holders and that which is attributable to claims of equity holders.\textsuperscript{47} As was shown by Modigliani and Miller,\textsuperscript{48} these claims are, at least in theory, economically indistinguishable, yet the tax treatment of these two different types of income is distinctly different. Interest payments are deductible, whereas dividend payments to shareholders are not.\textsuperscript{49} Consequently, earnings attributable to equity capital are subject to two layers of tax, while earnings attributable to debt capital are only subject to tax in the hands of the creditor.

Analogous problems arise with regard to individual preferences. The economics literature discusses many instances in which individuals appear to exhibit inconsistent preferences.\textsuperscript{50} Perhaps the most important examples in this context are dynamically

\textsuperscript{43} IRC § 1222 defines capital gains as gains from the sale or exchange of capital assets. IRC § 1221 defines capital assets as all assets other than inventory accounts receivable and certain other specified assets.


\textsuperscript{45} Currently, under IRC § 1 (h) interest and dividends are tax at the same rate as capital gains.


\textsuperscript{49} Chorvat, note 47, supra at 248-49.

\textsuperscript{50} Daniel Kahneman and Amos Tversky, Prospect Theory: A Theory of Decision Under Risk, 47 ECONOMETRICA 263 (1979).
inconsistent marginal rates of intertemporal substitution.\textsuperscript{51} This phenomenon is often referred to as hyperbolic discounting to contrast it with standard exponential discounting. It has been found that in many contexts individuals seem to exhibit a higher rate of discount between periods of time close to the current period as compared to periods of time in the distant future. For example, the discount rate between receiving an amount of money today versus receiving it tomorrow appear to be larger than the discount rate between receiving the same amount of money a year from now versus a year and one day from now.\textsuperscript{52} One relatively simple method of modeling such preferences is represented by the $\beta - \delta$ model of David Laibson.\textsuperscript{53} Like many standard models, this model assumes that time periods can be represented in a discrete fashion and that time preferences are separable. It represents the utility at the beginning of a stream of consumption as being broken up into the value of the consumption at each of the individual time periods, discounted back to the present time. This can be represented as $U(C(t)) = U(c_1) + \beta[\delta U(c_2) + \delta^2 (c_3) + \cdots]$. Under this model, one can easily show that the rate of time discount between the first period and the second period is $\frac{U'(c_1)}{U'(c_2)} = \beta \delta$, while the discount rate between the second and third period is $\frac{U'(c_2)}{U'(c_3)} = \delta$.\textsuperscript{54} At time period 2, the utility from future consumption can be represented as $U(C(t)) = U(c_2) + \beta[\delta U(c_3) + \delta^2 (c_4) + \cdots]$, and now the rate of discount between the

\textsuperscript{52} That this does not follow standard exponential discounting is demonstrated in Appendix II.
\textsuperscript{54} This is demonstrated in Appendix III.
second and the third period is \( \frac{U'(c_2)}{U'(c_3)} = \beta \delta \). That is, the rate of discount between period one and period two changes as we move into period one from period zero from \( \delta \) to \( \beta \delta \). Because the rate of discount between these two periods changes, this is known as dynamically inconsistency in the rates of time preference.

While it is sometimes said that inconsistent preferences such as this cannot be arbitrated,\(^{55}\) this is not true. To illustrate how an arbitrage could be arranged against someone with dynamically inconsistent preferences, imagine that there are three time periods, \( t_0, t_1 \) and \( t_2 \). At time \( t_0 \), let us say that the discount rate between the \( t_0 \) and \( t_1 \) is 20%, and the discount rate between \( t_1 \) and \( t_2 \) is 5%. If at \( t_1 \) the discount rate between \( t_1 \) and \( t_2 \) will change to 20%, then there is the potential for arbitrage. This could be accomplished by entering into a contract at time \( t_0 \) to borrow $100 at time \( t_1 \) for repayment at time \( t_2 \), with an addition of 5% interest, or payment of $105 at \( t_2 \). This contract will have a zero value at time \( t_0 \), because a person will be indifferent between entering into the contract and not entering into it. We know that at time \( t_1 \) the contract will have negative value to the lender of $15. Given a positive rate of interest, this payoff is inconsistent with zero value today. For the borrower, the contract will have a value of $15 at time \( t_1 \) (it can be sold for that). Therefore, at time \( t_0 \), an arbitrageur could enter into the loan agreement as the borrower, and then in addition borrow a sum of money today (that is, go short money) for repayment in period \( t_1 \), such that the amount of the repayment will be exactly $15.\(^{56}\) Then in period \( t_1 \), the arbitrageur could sell the contract


\(^{56}\) It does not matter whether the interest rate is 20% or 5% for this to work. This only controls the initial level of profit from the arbitrage, not the initial existence of arbitrage.
to lender and use the $15 received to pay off the loan. This satisfies the definition of arbitrage because it results in a positive cash flow in period \( t_0 \) with no net cost in any later time period. Therefore, dynamically inconsistent persons can be arbitraged. At a minimum, this implies that it cannot be the case that markets in equilibrium exhibit non-exponential discounting.

While this type of arbitrage is possible, so-called sophisticated hyperbolic discounters, who understand that they will have different preferences in the future, avoid this type of arbitrage.\(^{57}\) They may accomplish this by either refusing to enter into any such contracts in the first period, or alter the terms of the contract accordingly. By foreseeing their own inconsistencies, they prevent greater loss to utility by adopting matching inconsistencies in their preferences.

As discussed earlier there are many ways in which the tax system has inconsistent preferences. The conditions for potential tax arbitrage are met because different types of income are taxed at different rates, and there are substantial tax-exempt entities as well as significant numbers of non-U.S. participants even in U.S. markets, to say nothing of foreign markets. There are essentially three methods by which tax arbitrage is avoided. In the first method, market prices adjust to reduce and possibly eliminate tax arbitrage. In the earlier examples, if prices did not reflect the after-tax payoffs, it would be possible to derive arbitrage profits from the market. As market participants become aware of arbitrage strategies, in general prices will adjust to account for the tax benefits other participants can derive and the market will then act to eliminate such gains. Notice that while it may no longer be possible for any one person to derive an economic profit from

such transactions, because prices will have been altered by such transactions, the tax system may have continuing efficiency effects on the economy. By altering prices, this may alter the allocation of resources which would generally be thought to reduce the efficiency of the tax system.\textsuperscript{58}

The second method by which tax arbitrage is avoided is by the government placing restrictions on the use of losses from one transaction to offset gains from another. For example, the tax rules prevent a taxpayer from deducting interest from loans the proceeds of which were used to acquire assets whose income is tax-exempt.\textsuperscript{59} Similarly, losses on straddle positions are not allowed to be taken until the gains all of the positions in the straddle are realized.\textsuperscript{60} To the extent that such rules are introduced, the tax system moves away from being a pure income tax. While there are number of such restrictions, as long as the basic structure of the tax remains an income tax where losses offset gains, then tax arbitrage of the type discussed in the article will continue to be an issue.\textsuperscript{61}

It is the thesis of this part of the article that there is a third, and perhaps most controversial, method by which the tax system deals with tax shelters: the adoption counterbalancing "inconsistent" rules. I would argue that one example of a doctrine which illustrates this offsetting inconsistency is the economic substance doctrine. This doctrine holds that the desired tax treatment will apply to a transaction or series of transactions if they have "economic substance." The economic substance doctrine can result in different tax treatment of transactions that under the rules should receive the

\textsuperscript{58} The seminal work in the area is Frank P. Ramsey, \textit{A Contribution to the Theory of Taxation}, 37 \textsc{Econ. J.} 4, 58-59 (1927).
\textsuperscript{59} IRC § 265
\textsuperscript{60} IRC § 1092. A straddle is a group of transactions where an investor is long a position and short an essentially identical position but with a different expiration date. For example, an investor could be long April T-bills and short May T-bills.
\textsuperscript{61} Constantinides note 2 \textit{supra} at 635-36.
same treatment. In particular, if a transaction does not involve the transfer of a significant level of risk, yet the tax benefits are substantial, the doctrine can act to nullify the intended tax benefits. In other words:

The economic substance doctrine is not just a smell test, because it only applies to transactions which lack economic substance. The doctrine permits taxpayers to retain even the most egregious tax benefits if they arise from transactions with meaningful economic consequences.62

If significant tax benefits flow from a transaction in which there was no undertaking of economic risk, then the desired tax treatment will not be respected.63 The inconsistency here derives from the fact if there are two transactions that both meet the requirements the tax rules set forth to receive a particular tax treatment, their tax results will be different if one of them involves the undertaking of significant risk or has some significant non-tax reason for its structure and this is not the case for the other transaction.64 Therefore, even though the explicit rules have been followed, if a transaction essentially involves tax arbitrage, the desired tax benefits will be disallowed.

Based on these considerations, one can see the point of the economic substance doctrine as simply requiring than the transactions not arbitrage the tax rules. One way of reading the economic substance doctrine is simply as a reasonable backstop to arbitrage. One could note that this doctrine introduces an inconsistency designed to match other inconsistencies and so prevent the tax system from being arbitrated.

Classic arbitrage requires that the positions taken in the arbitrage be riskless. If instead the position requires some risk, then it is not classic arbitrage, but rather some

---

62 David Hariton, Sorting Out the Tangle of Economic Substance 52 TAX LAW. 235, 235 (1999)
63 See U.S. v. Phellis 257 U.S. 156, 165 (1921) (Holding that the substance of a transaction should govern its tax treatment, rather than its form).
64 Frank Lyon Co. v. U.S. 435 U.S. 561 (1978) (Allowing the taxpayer to retain the desired tax treatment because the form of the transactions was used to comply with banking regulations).
form of investment. However, preventing classic arbitrage, is likely not sufficient for preventing the problems associated with tax arbitrage. Because there is profit to be made from tax arbitrage, taxpayers might very well be willing to undergo some risk in order to obtain significant tax benefits. Furthermore, the benefits of a tax arbitrage might be out of proportion to the risk undertaken, which will general not the case in market arbitrage. Therefore, a simple rule about the complete lack of risk would not eliminate the ability of individuals to profit from the tax system. Therefore, the optimal doctrine would allow for the risk undertaken in a transaction to be balanced against the tax benefits received. Once one realizes that the undertaking of some risk should not prevent the categorization of a series of transactions as a tax shelter, then it must be the case that we engage in a complicated balancing between the tax benefits available and the amount of the risk undertaken, which must of necessity be based on the facts and circumstances of the situation. This is essentially the structure of the economic substance doctrine as we find it today.

If the problem of tax shelters is thought of as a problem of preventing tax arbitrage, then the solution is not to be found in the adoption of a purposive system of interpretation or any other doctrine of interpretation. It is not that the debate over the appropriate method of interpretation is not relevant to many issues, it is just that any such consistent doctrine of interpretation will necessarily allow for the creation of tax shelters if there is fundamental inconsistency in the tax law. Rather, the attack on these transactions must involve seeing them as a form of arbitrage and consequently adopting

---

65 Indeed, as long as risk aversion is not infinite, all taxpayers should be willing to undertake some risk to reduce their taxes.

66 Hariton, note 62, supra at 235-36.

67 For a discussion of the use of methods of interpretation to combat tax shelters, see generally Noël Cunningham and James Repetti, Textualism and Tax Shelters 24 VA. TAX REV. 1 (2004).
strategies that can be used to prevent arbitrage. Seen in this light, the key to attacking tax shelters is then the level of examining the risk undertaken in a transaction or series of transactions as compared to the tax benefits derived. Fortunately, this appears to be the focus of the economic substance doctrine.
APPENDIX I

To see this, we note that

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}$$

is the payoff matrix $X$, for $n$ securities in $m$ states of nature, then

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \begin{pmatrix} \omega_1^i \\ \vdots \\ \omega_n^i \end{pmatrix} = \begin{pmatrix} x_{11} \omega_1^i + \cdots + x_{1n} \omega_n^i \\ \vdots \\ x_{m1} \omega_1^i + \cdots + x_{mn} \omega_n^i \end{pmatrix}$$

the latter of which is the vector of the values of the portfolio in each of the $n$ states of nature.
APPENDIX II

To show that under exponential discounting the rate between today and tomorrow should be equal to the discount rate between one year from today and one year and one day from today, consider that $r$ is the annual discount rate. The discount form any day in the future is $e^{rt}$. For one day in the future the discount rate is then $e^{\frac{1}{365}}$, while the discount rate for a year from today would be $e^{r}$, and for one year and one day $e^{\frac{366}{365}}$. If we then compare the relative discount rate between one year from now and one year and one day from now, we have $\frac{e^{\frac{366}{365}}}{e^{r}} = e^{\frac{1}{365}}$, which is the same discount rate as tomorrow.

For discrete time periods, we note that if each time period has the same level and it is always the case that $\frac{d_{n+1}}{d_n} = r$ for all $n$, where $d_n$ represents the rate of discount between the $(n-1)$st and the $n$th period, then $d_{n+1} = rd_n$, and then $d_{n+1} = r^{n-1}d_1$. 


The inconsistency of intertemporal discounting under the $\beta - \delta$ model is most easily seen in a three period model. If we start off with a constant level of capital to spend, the individual will optimize the equation:

$$\max U(C) = U(c_1) + \beta \left[ \delta U(c_2) + \delta^2 (c_3) \right] \text{ subject to } c_1 + c_2 + c_3 = K$$

The first order conditions are then:

$$U'(c_1) - \lambda = 0$$
$$\beta \delta U'(c_2) - \lambda = 0$$
$$\beta \delta^2 U'(c_3) - \lambda = 0$$

The dynamic inconsistency can be seen by the Euler equations for the substitution from the first period to the second $\frac{U'(c_1)}{U'(c_2)} = \beta \delta$, while the substitution from the second to the third period is $\frac{U'(c_2)}{U'(c_3)} = \delta$.

If we wish to add at return on investment equal to an interest rate $r$, so that the constraint is now $K = c_1 + \frac{c_2}{r} + \frac{c_3}{r^2}$ we have the similar result from the first order conditions of:

$$U'(c_1) - \lambda = 0$$
$$\beta \delta U'(c_2) - \frac{\lambda}{r} = 0$$
$$\beta \delta^2 U'(c_3) - \frac{\lambda}{r^2} = 0$$

This gives us the Euler equations of $\frac{U'(c_1)}{U'(c_2)} = \frac{\beta \delta}{r}$ and $\frac{U'(c_2)}{U'(c_3)} = \frac{\delta}{r}$, which again are different by a factor of $\beta$. 

26