Response: A Defensible and Workable Approach to the Income Tax Treatment of Financial Instruments

Tim Edgar*

KEYWORDS: FINANCIAL INSTRUMENTS ■ HEDGING ■ EQUITY FINANCING ■ DERIVATIVES ■ DEBT FINANCING ■ FINANCING

It is a privilege to have the opportunity to respond to the comments on my monograph provided by Mark Gergen, Glenn May, and Gordon Longhouse. Their comments, which are inevitably coloured by their very different perspectives, reflect the considerable expertise that each one of them has in the area of the income taxation of financial instruments. Indeed, it is with some hesitation that I offer a response in defence of various portions of the analysis presented in my monograph in support of some pretty modest proposals in this extremely difficult area of income tax law. Although I spent considerable time exploring some necessary first principles and their implications for the design of a system for the income taxation of financial instruments, I made several concessions to certain practical constraints that led me to support, in some measure, the status quo reflected in certain of the existing literature, as well as the legislation in a select group of countries.

On the assumption that many readers may be unfamiliar with the monograph, I propose to respond by outlining much of my analysis in the monograph and the proposals that are the logical outcome. Throughout the outline, I will highlight and respond to what I see as the important points of difference emphasized by Gergen, May, and Longhouse.

THE CORE ELEMENTS OF A COMPREHENSIVE ACCRUAL REGIME

At the risk of repetition, let me begin my response by restating the goal of the monograph and the broad aspects of the analysis offered in support of its modest proposals. The goal was to explore, in a relatively comprehensive way, the state of

* Faculty of Law, The University of Western Ontario, London, and a senior research fellow, Taxation Law and Policy Research Institute, Deakin University, Melbourne.
the theoretical work on the income taxation of financial instruments and to see how that work might be translated into an administrable legislative regime. I drew extensively on the experience in Australia, New Zealand, and the United States, since policy makers in these three countries have been the most ambitious in developing such a regime. The framework for the inquiry was explicitly second-best in nature. That is, I accepted that a comprehensive system of accretion or mark-to-market taxation applied to all financial instruments, including shares, is not about to be adopted in the short or even the medium term. Given this framework, I sought to flesh out the broad policy case and design details of a “first-best, second-best approach” to the income taxation of financial instruments.

As a legislative paradigm, the approach I advocated was that of expected-return taxation, which has been articulated most clearly by certain US writers. The principal competitor to this approach is that of retrospective taxation, which also has been developed at length in the academic literature. I rejected this alternative approach primarily because of its unfamiliarity. Given its marginal substantive improvements over a system of expected-return taxation, I do not think that policy makers should embark any time soon on a legislative elaboration of retrospective taxation.

Expected-return taxation is based on the observation that all financial instruments, beginning with the basic building-block instruments (fixed-payment debt, forwards, futures, swaps, and options) and extending to sophisticated combinations of these basic instruments, consist of expected returns attributable to the time value of money and unexpected gains or losses attributable to a bet on a specified contingency. I argued that the development of a comprehensive accrual regime as the legislative expression of a system of expected-return taxation is not predicated on any attempt to implement a comprehensive income tax base approximating the Haig-Simons ideal. In fact, the principal goal of policy makers should be the consistent taxation of equivalent cash flows identified in terms of expected and unexpected gains and losses. This goal ensures that perfectly substitutable transactions are taxed equivalently. The imperative underlying this goal is the need to block revenue leakage from avoidance transactions and minimize the associated welfare loss. I emphasized this imperative as the normative justification for a system of expected-return taxation. Unlike May, I did not construct a hierarchy of broad policy norms that would otherwise obscure my analysis with unwanted policy noise. In this respect, I was probably unrealistic, but I believe it is important to frame the fundamental design issues as a critical first step in the development of a theoretically defensible and administratively workable system for the income tax treatment of financial instruments.

In terms of a preferable second-best approach, policy makers can realize the paramount goal of consistency within a system that eliminates both the corporate income tax (with its required distinction between debt and equity instruments) and the preferential treatment of capital gains and losses (with its required distinction between non-capital and capital amounts). In addition, given a world in which different taxpayers are taxed at different rates, accrual recognition of expected gains and losses is required to address the problem of anticipated deferral, whereby
expected gains are taxed in a period subsequent to that in which they are considered to arise. Unexpected gains and losses can be ignored because, from a strictly ex ante perspective, there is an equal chance of realizing gain or loss from a specified bet. Ignoring both unexpected gains and losses eliminates the problem of selective realization presented by the realization-based recognition of such amounts measured ex post.

The monograph does not propose, however, the adoption of this pure form of expected-return taxation. Indeed, the principal story told in the monograph is the move from the theoretical ideal to a model that can be made to work tolerably well in the real world of practical policy making and administration. In this respect, I accept that radical reform of the corporate income tax, such as the imputation of interest on equity or the non-deductibility of corporate interest expense, is simply not feasible in the short to the medium term. I also accept that we are not about to ignore ex post results where those results can be measured in a readily administrable manner. Accordingly, unexpected gain or loss must continue to be brought to account on a realization basis, but with some attempt to control the problem of selective realization. I do not accept as readily the need to maintain the preferential treatment of capital gains, and I would certainly prefer that the distinction be eliminated for gain or loss on financial instruments, with the exception perhaps of shares. I also accept that interest need not be imputed wherever we find a time-value-of-money return; nor must it be measured in an overly precise manner. At some point, which is not always very clear, the administrative and compliance costs associated with such an inquiry are not warranted, at least in the absence of compelling evidence that concessions made for administrative reasons induce instances of tax-motivated behaviour. As Gergen notes, acceptance of this administrative imperative is the apparent reason that US policy makers have chosen not to require accrual of the expected-return portion of secondary-market discount, which otherwise would have to be distinguished from unexpected gain attributable to improvement in the creditworthiness of a debtor. Such acceptance also underlies a general reluctance to account for the term structure of interest rates and the weighted average life of an instrument in measuring expected returns.

The clearest legislative expression of a necessarily modified system of expected-return taxation is a comprehensive accrual regime, such as that adopted in New Zealand and proposed in Australia. At its core, such a system is nothing more than an extension of the existing treatment of fixed-payment debt to a broad range of expected returns associated with other financial instruments. In effect, a comprehensive accrual regime requires some massaging of the status quo in many countries in a way that shifts the familiar boundaries that are often the focal point of tax-driven financial innovation. Gergen describes accurately the basic elements of the kind of massaging that I envision. In keeping with many existing income tax systems, a comprehensive accrual regime consists of (1) the realization-based recognition of unexpected gains and losses associated with forwards, futures, and swap contracts; and (2) accrual recognition of the expected gains and losses on fixed-payment debt. The modifications proposed in the monograph are:
the elimination of the distinction between ordinary income and capital amounts;
the extension of accrual recognition to the expected return on debt with contingent payments, as well as the embedded debt in options and prepaid and off-market forwards; and
the application of accretion recognition to unexpected gains and losses on traded instruments that give rise to problems of selective realization.

The massaging is far from complete, however, because practical policy constraints prevent the consistent treatment of all financial cash flows. This incomplete application of expected-return taxation leaves troublesome boundaries that are all too familiar. Indeed, my acceptance of the significance of practical policy constraints requires the development of what Gergen refers to as a “mongrel” accrual system and what I labelled a “modified” system of expected-return taxation. This system is reflected, in part, in the legislation in New Zealand and the United States and the legislative proposals in Australia. As several US writers have pointed out, the boundaries that such a system entails create “discontinuities,” whereby small changes in the composition of cash flows result in disproportionate differences in tax treatment. The difficult policy issue is how to define and then police these boundaries in a manner that suppresses their significance.

In the absence of any normative significance, the accepted boundaries must be defined in a somewhat arbitrary and even messy fashion. Nonetheless, the policy focus should be the substitutability of instruments. This focus has been explored most thoroughly in the academic literature by David Weisbach; yet one has a sense that policy makers have always focused on substitutability, although not as systematically as Weisbach has. Perhaps most important, Gergen is entirely correct in emphasizing the lack of empirical work on the significance of different tax treatments as an impetus to tax-driven substitutions. Unfortunately, in the absence of this important work, policy makers must develop legislation in a state of blissful ignorance that is illuminated only by a very casual empiricism.

As far as matters of process are concerned, the monograph suggests that the responsibility for developing a sensible approach to the treatment of financial instruments lies with policy makers, who should consider adopting a modified system of expected-return taxation to eliminate inconsistencies and suppress troublesome discontinuities more completely. In short, in terms of institutional competency and associated efficiency considerations, it is the task of policy makers to explore the extent to which tax legislation is constrained by a need to tax on the basis of legal form, and/or to explore the alternative approaches that are available as a means to tax on the basis of economic substance defined in terms of cash flows giving rise to expected and unexpected gains and losses. In this respect, May is correct in emphasizing that financial accounting practice can be relied on when it comes to issues of measurement, but not issues of characterization. My disagreement with his proposition is not one of kind but simply one of degree. Even on issues of measurement, financial accounting practice should be used only as a convenient starting point. Because of the different goals of financial and tax accounting, policy makers must
reserve the right to deviate from financial accounting practice in designing measurement rules for financial instruments. The monograph describes in some detail the kinds of parameters that might be set for an acceptance of such practices.\textsuperscript{11}

\section*{DEFINING BOUNDARIES IN A SECOND-BEST WORLD}

Gergen, May, and Longhouse all emphasize, to varying degrees, many of the significant problems of a modified system of expected-return taxation. These problems are, in general, attributable to a necessary acceptance of the different taxation of interest and dividends and the application of realization-based recognition to shares and other non-traded financial instruments and non-financial assets. The principal pressure points arise in the following areas: the boundary between debt and derivatives; the boundary between debt and equity; the boundary between synthetic replication and hedge accounting; and the boundary between financial instruments and non-financial assets. A modified system of expected-return taxation also creates a boundary between traded instruments subject to accretion recognition and non-traded instruments subject to a combination of accrual recognition of expected returns and realization-based recognition of unexpected gains and losses. I proposed, however, to relieve the pressure on this particular boundary by defining traded instruments broadly in a functional manner that targets the features associated with opportunities to trade strategically.\textsuperscript{12} Moreover, I suggested that, despite possible efficiency losses, a loss-limitation rule should probably be adopted for non-traded instruments.\textsuperscript{13}

\subsection*{Debt-Equity Boundary}

With respect to the boundary between debt and equity instruments, I agree with Longhouse that, ideally, some form of imputation system should be adopted in an effort to equate, to some extent, the tax treatment of debt and equity returns. In a sense, the corporate income tax is nothing more than a proxy for taxing expected returns on corporate assets to shareholders on an accrual basis and accounting for unexpected gains and losses on realization. I have never been convinced, however, that the complexities associated with a system of full imputation are worth the candle, particularly when, for revenue reasons, the benefit of full imputation is not extended to tax-exempt entities and non-residents. In this respect, I believe the Canadian system, which provides unfunded (but only partial) dividend relief on common shares and limits a distributions tax to taxable preferred shares that are close debt substitutes, has much to recommend it.

Even under the Canadian approach, as well as under a full imputation system such as that adopted in Australia and New Zealand, the returns on debt and equity are not equated and there remain problematic instances of tax-driven substitution. These limitations of an imputation system arise primarily with respect to cross-border investment and investment by domestic tax-exempt entities. I argued in the monograph that a necessarily indeterminate factors approach is the preferable way
to police the boundary that remains between debt and equity under a comprehensive accrual regime for financial instruments other than shares. The use of such an approach is especially important in the area of cross-border investment, where differences in national tax regimes provide tax-avoidance opportunities based on the different classifications by source and residence countries of financial instruments. Until greater coordination of national tax regimes is achieved, a “facts and circumstances” approach to the classification of financial instruments may be the only effective means of addressing avoidance opportunities. The provision of certainty along the debt-equity border is a breeding ground for cross-border avoidance.14

As Longhouse points out, my premise for the adoption of a factors approach is an assumption that much financial innovation along the debt-equity boundary is tax-driven. Here again, systematic empirical work is somewhat lacking. However, I am not sure that the examples of benign innovation that Longhouse describes cause me to reject my basic premise; nor am I sure that they would be especially problematic under a factors approach for characterization purposes. In fact, a factors approach can accommodate benign instances of innovation and treat the tax-driven ones in a manner consistent with their closest substitute. In whatever way it is constructed, the alternative of a bright-line test inevitably creates stark discontinuities that become the focal point for tax planning. I do not think this fundamental dynamic would change much under the “tax-gap” approach suggested by Longhouse. Indeed, I suspect that a broad range of instruments that would fall into the purgatory of his “tax gap” could be dealt with more appropriately under a factors approach.

Debt-Derivative Boundary

The different treatment of debt and derivatives requires the maintenance of a boundary between these two categories of instruments. The boundary falls along two lines, each of which raises the issue of the extent to which expected returns are to be recognized on an accrual basis by imputing interest irrespective of the legal form of a particular instrument. The first line involves the embedded derivative (or bet) element in contingent-payment debt. The second involves the embedded debt element in certain derivative financial instruments. I argued in the monograph that, in principle, bifurcation can be applied to both these categories of instruments to realize consistent treatment of the expected return associated with the debt element and the unexpected gain or loss associated with the bet element. As applied to the taxation of hybrid instruments, which consist of two or more basic instruments combined in one legally distinct instrument, the process of bifurcation is conventionally understood as the attempt to break down the instrument into its components and apply a specified tax treatment to each part. It has long been recognized that bifurcation is virtually impossible to apply as a general approach to the taxation of financial instruments under an income tax system that is inconsistent in its treatment of the basic building-block instruments of financial innovation. The varieties of financial equivalences and the sophistication needed to bifurcate complex financial instruments mean that bifurcation is difficult to administer. More fundamentally, the different taxation of the basic components of financial instruments renders
bifurcation inconclusive in many instances, since different combinations of the basic parts can produce different tax treatments even after a complex instrument has been bifurcated.

I argued, however, that policy makers can alleviate these difficulties somewhat by identifying the embedded debt element in all derivatives and imputing interest, whether the derivative is issued on a stand-alone basis or is embedded in a contingent-payment debt instrument. I do acknowledge, though, that under this approach the debt-derivative boundary remains uncertain and troublesome as a result of possible revenue loss from increased interest deductions as well as perceived administrative and compliance problems. The monograph tentatively suggests some ways of identifying those instances in which the expected return associated with the embedded debt in a stand-alone, non-traded derivative is significant enough to indicate that the instrument can be considered a substitute for a contingent-payment debt instrument and, therefore, should be subject to accrual treatment. This proposition accepts, of course, that the expected return associated with the fixed-payment debt and embedded derivative elements in contingent-payment debt are significant enough to require accrual treatment along the lines imposed by the US regulations for contingent-payment debt.

**Synthetics and Hedge Accounting**

Because accretion recognition under a modified system of expected-return taxation must be limited to traded financial instruments, it remains possible to create financial equivalences using simple and complex synthetic instruments, which combine two or more legally distinct instruments that replicate the cash flow pattern associated with another legally distinct instrument. For example, taxpayers can create offsetting positions with financial instruments that are subject to accretion taxation and others that are subject to expected-return taxation. Similarly, financial instruments subject to expected-return taxation can be used to replicate other instruments subject to accretion taxation. These kinds of equivalences are advantageous because expected-return taxation requires the recognition of expected gains and losses on an accrual basis, while accretion taxation captures both expected and unexpected gains and losses. Avoidance opportunities may also be available because of rate differences among taxpayers and any exclusion from an accrual regime of instruments held by individuals. Other problematic areas are (1) the exclusion of non-traded shares from an accrual regime; (2) the provision of exemptions from non-resident withholding tax for portfolio interest payments and payments on derivatives; and (3) the denial of dividend imputation credits to non-resident shareholders and tax-exempt entities. In these instances, long and short positions in financial instruments can be combined to (1) constructively dispose of shares without attracting a disposition transaction for income tax purposes; (2) avoid non-resident withholding tax on interest and dividends; and (3) transfer the benefit of dividend imputation credits and foreign tax credits.

These kinds of avoidance opportunities can be addressed, to some extent, with specific legislative responses applicable to transactions that involve synthetic
replication. Nonetheless, impossible questions surround the distinction between the tax-driven use of synthetics and legitimate hedging strategies. The distinction is an unprincipled one involving differences in degree. As a result, any attempt to draw a boundary and apply integrated tax treatment to otherwise independent financial instruments is an arbitrary exercise. I think the empirical assumption that policy makers must make along this boundary is that most synthetics are, in fact, the result of legitimate hedging strategies, and anti-avoidance responses to the tax-driven variety should be narrowly targeted. The monograph provides some examples of this suggested approach, drawing primarily on the US legislative experience. Although I agree with May that withholding tax regimes tend to be the outcome of messy tradeoffs between a defensible revenue claim and the need to attract foreign investment at an acceptable price, I fail to see, at least as a matter of principle, why attempts to avoid withholding tax by synthetically replicating the cash flows associated with a taxable instrument should be allowed to persist. As I suggested in the monograph, the lack of initiative by capital-importing countries may simply be the result of a combination of the lack of any serious commitment to the maintenance of withholding taxes and the existence of non-tax constraints on such transactions that keep the revenue loss to a tolerable level. Where these conditions do not hold, the restrained approach to synthetics described in the monograph can be applied relatively easily in an effort to maintain the withholding tax base.

When it comes to the adoption of a hedge accounting regime, I plead guilty to the charge of heresy levied by Gergen. My heresy is conditional, however, on the elimination of a capital gains preference and the recognition of the full amount of all gains and losses. Conceptually, a hedge accounting regime is nothing more than a limited form of integrated treatment whereby gain or loss on a financial instrument acquired as a hedge of the risk associated with an underlying asset or liability is taxed with reference to that underlying asset or liability. As the US experience reveals, such a regime presents serious administrative and compliance difficulties which may nonetheless be justified if preferential treatment of capital gains is maintained and character mismatches are created by the income tax system.

To the extent that the distinction between ordinary income and capital amounts can be eliminated, hedging rules may be unnecessary, or at least not worth the considerable administrative and compliance burdens. This appears to be the judgment of New Zealand policy makers, who have rejected a hedging regime in the context of their accrual regime. Although I tend to favour this position, the monograph concedes that a very limited hedging regime could be defended under a modified system of expected-return taxation that is based on the full recognition of gains and losses on financial instruments. This regime would permit realization-based recognition of unexpected gains and losses on traded derivatives otherwise subject to accretion taxation. However, the particular instruments should be suitably designated ex ante as a hedge of a non-traded underlying asset or liability.
Boundary Between Financial Instruments and Non-Financial Assets

Gergen also highlights the inconsistent treatment of financial instruments and non-financial assets that would arise because of a failure to extend a modified system of expected-return taxation to non-financial assets. The inconsistent treatment would be attributable primarily to the failure to impute interest for accrual purposes on non-financial assets. Although there is some discussion of such an extension in the literature, its application in practice has been very limited. Continued maintenance of inconsistent treatment requires the drawing of a boundary between financial and non-financial assets, with associated discontinuities. I argued in the monograph that the guiding principle along this boundary should be no different from that along the other boundaries. In particular, policy makers must decide when derivative financial instruments written on a non-financial asset can be considered to substitute for a contingent-payment debt instrument. Where the expected return is significant enough that it can be considered a debt substitute, the return on the instrument should be subject to accrual recognition.

Under a system of expected-return taxation that excludes non-traded shares, the focus of consistent treatment shifts to direct versus indirect ownership of a non-financial asset and the role of a dividend imputation regime in ensuring a measure of consistency. As Gergen points out, the principal inconsistency imposed by a system of expected-return taxation in this context arises from the application of accretion recognition to traded shares. This inconsistency is defensible if it is decided that the revenue and welfare loss from the opportunity to strategically trade such shares is more significant than the loss that would follow from the substitution of the non-corporate for the corporate form in an effort to avoid the application of accretion recognition. My intuition is that non-tax constraints on the latter form of substitutability tend to support the application of accretion recognition to traded shares as a first-best response to the problem of selective realization. Again, however, we have virtually no empirical evidence of the behavioural effects of the limited application of accretion recognition to traded shares.

CONCLUSION

I believe that even with the kinds of problems highlighted so well by Gergen, May, and Longhouse, the kind of modified system of expected-return taxation described in my monograph would be the most effective way forward in the short to medium term. Compared with the status quo in many countries, such a system would reduce much of the tax significance of distinctions among a broad range of financial instruments by eliminating opportunities to disguise expected gains as unexpected amounts. Because even a well thought out legislative regime cannot possibly deal with all circumstances, however, the courts must serve a critical supporting role, especially when the legislative framework is unavoidably second-best. That role is performed most effectively where there is a clear expression of general principles that tax administrators, practitioners, and the courts can focus on when working through
the details of the appropriate treatment of specific instruments. Some ambiguity in the difficult boundaries between debt and equity and between debt and derivatives, along with some anti-avoidance provisions that attempt to address the worst instances of synthetic replication, can be justified in order to introduce sufficient legal risk to deter taxpayers from unduly exploiting remaining discontinuities.

I continue to have some faith in specific, purpose-based anti-avoidance rules as an effective supplement to the selected use of bright-line definitions of some of the necessary boundaries. These kinds of rules focus the inquiry on the substitutability of transactions, with the characterization of purpose serving as a proxy for the identification of instances of tax-driven substitution. Policy makers should take care, however, not to respond to all apparently tax-driven financial instruments in an “anti-avoidance” mode. The saga of weak-currency borrowings in Canada, which is discussed by May, provides an excellent illustration of this point. Instead of enacting legislation that treats Canadian-dollar and foreign-currency debt consistently by accruing expected exchange gains and losses, the Department of Finance chose an anti-avoidance rule that maintains a measure of inconsistent taxation of equivalent cash flows and creates boundaries with associated discontinuities where none existed before. Systemic legislative responses that eliminate inconsistent treatment of equivalent cash flows should be chosen wherever possible. Boundaries and their associated discontinuities should be created only where compelling practical policy constraints prevent the realization of consistent taxation. Weak-currency borrowings arguably presented no such constraints. The task of designing a theoretically defensible and administratively workable system for the income tax treatment of financial instruments is difficult enough without making it more difficult than it needs to be.

NOTES


4 For a discussion of some of the deficiencies of a system of retrospective taxation, see Michael S. Knoll, “Financial Innovation, Tax Arbitrage, and Retrospective Taxation: The Problem with Passive Government Lending” (1997) vol. 52, no. 2 Tax Law Review 199-224; and Michael S. Knoll, “Tax Planning, Effective Marginal Tax Rates, and the Structure of the Income Tax” (2001) vol. 54, no. 4 Tax Law Review 555-83. I assume, perhaps incorrectly, that the elaboration of a workable system of expected-return taxation has sufficient intellectual content to occupy the attention of policy makers. If Sunley’s intuition (noted by Gergen) is, in fact, correct, retrospective taxation may be a more viable policy option than its substantive merits would suggest.

The mark-to-market legislation applicable to specified debt obligations held by specified financial institutions requires the accrual recognition of secondary-market discount. See the Income Tax Act, RSC 1985, c. 1 (5th Supp.), as amended (herein referred to as “the Act”), sections 142.2 to 142.6 and draft regulations 9100 to 9300. Unless otherwise stated, statutory references in this article are to the Act. It is not entirely clear on what basis, if any, the portion of a discount attributable to a decline in the creditworthiness of the debtor is to be excluded from accrual recognition. It may be, as Gergen suggests, that the difficulty in determining an appropriate repricing of a debt obligation to ensure the required exclusion effectively makes elective any accrual requirement for secondary-market discount. For taxpayers other than specified financial institutions, it is probably accurate to state that the balance of authority under the Act supports the proposition that secondary-market discount need not be accrued. It is not clear, however, that this approach has been adopted out of administrative concern with the distinction between the portion of discount attributable to changes in market conditions and the portion attributable to declines in creditworthiness.


Edgar, supra note 1, at 207-17.

Ibid., at 230-35.

Ibid., at 235-39.


Edgar, supra note 1, at 350-58.

Ibid., at 359-61.

Ibid., at 361-68.


22 Edgar, supra note 1, at 289-91.


24 Edgar, supra note 1, at 236-37.


26 Section 20.3.