DOI: 10.1111/1475-5890.12255

Fiscal Studies
The Journal of Applied Public Economics
IFS Institute for Fiscal Studies

ORIGINAL ARTICLE

Taxation of capital gains upon accrual: is it really more efficient than realisation?

Giampaolo Arachi[1] | **Massimo D'Antoni**[2]

[1] Università del Salento;
Dondena Centre, Università Bocconi

[2] Università di Siena

Correspondence
Massimo D'Antoni, Università di Siena, Piazza San Francesco 7, 53100 Siena, Italy.
Email: dantoni@unisi.it

Submitted March 2020

Abstract

Taxation of capital gains upon realisation is known to discourage the liquidation of an appreciated asset (lock-in effect), thus resulting in a distortion in portfolio choices. However, compared with taxation upon accrual, realisation-based taxation may imply a lower distortion in intertemporal allocation of consumption. In particular, we show that this is always the case when preferences are intertemporally weakly separable and homothetic. Using a simple model of intertemporal choice, we assess and compare the distortions induced by the two tax regimes of accrual-based and realisation-based taxation of capital gains. Our numerical simulations show that, for reasonable values of the parameters, realisation-based taxation can in fact imply a lower overall distortion of investment choices, and hence no clear ranking of the two systems is possible in general. Additionally, we find that the differences in efficiency, in either direction, are likely to be very small.

KEYWORDS
capital gains taxation, taxation upon accrual, taxation upon realisation

JEL CLASSIFICATION NUMBERS
H21, H25

1 | INTRODUCTION

The taxation of capital gains is one of the most critical areas in modern tax systems and its design raises a number of difficulties. The definition of capital gains itself is one of the most contentious and controversial issue in tax policy,[1] as the concept may have a different meaning and play a different

[1] Evans and Krever, 2017; Avi-Yonah, Sartori and Marian, 2011.

role from country to country.[2] In particular, Commonwealth countries tend to include in the definition a wide array of non-recurring receipts that in other jurisdictions are usually classified as business income, investment income or income from services and labour. European continental countries traditionally apply a much narrower definition, limited to gains on particular assets.[3] Furthermore, the same income may be treated as a capital gain or not, depending on the legal form through which individuals save for the future.[4] The uncertainties and inconsistencies surrounding the definition of capital gains for tax purposes have been increasingly exploited in tax planning through innovative financial instruments, placing considerable strain on the tax system.[5]

Once an income stream is classified as a capital gain, the key issue is whether the tax should be levied upon accrual or upon realisation. Under an accrual system, capital gains and losses are assessed and taxed with the same timing of ordinary income, usually on a yearly basis, whereas under a realisation system the tax is levied when the asset is disposed of through sale, gift, exchange or other means.[6]

The economic literature has long argued that accrual should be preferred to realisation-based taxation, on both equity and efficiency grounds. From the point of view of equity, taxation of accruing capital gains is consistent with the Haig–Simons definition of comprehensive income as the sum of consumption and the change in value of taxpayer's wealth over a given period of time. From an efficiency perspective, taxation upon accrual would avoid a number of distortions in capital markets that arise when the tax is levied on realised capital gains. In particular, taxation upon realisation provides a timing option that allows investors to defer capital gains (and realise capital losses immediately). The opportunity to reduce the effective tax burden by deferring realisation will distort the optimal liquidation policy, by leading investors to defer the disposal of an appreciated asset.[7] This distortion is usually referred to as the lock-in effect, and it can result in an inefficient portfolio allocation because investors find it optimal to hold an appreciated asset even if alternative investments provide higher before-tax returns or a more efficient portfolio diversification.[8]

Despite the suggestion of the economic literature, no country applies accrual taxation on a broad scale.[9] The usual explanation for this gap between theory and practice is that the actual implementation of accrual taxation is hindered by a series of problems: (a) it involves high compliance costs for taxpayers; (b) 'marking to market' is difficult for non-traded assets; (c) in some extreme cases, unrealised gains may force liquidation at the time of tax payment.[10] To summarise, the common wisdom is that an accrual system is more equitable and efficient than taxation upon realisation, but it

[2] Ault, Arnold and Cooper, 2019.

[3] Evans and Krever, 2017.

[4] For example, interest from bonds may be treated as capital gain if the investment is made through a mutual fund that distributes annual income by repurchasing mutual fund shares from fundholders; the share repurchase triggers capital gains treatment on an amount equal to the income on the underlying bonds held in the fund (Erickson et al., 2020).

[5] See Alworth (1998). Regarding the distortions determined by an exemption of capital gains, see Stiglitz (1985).

[6] As well as for the concept of capital gain, there is also a large cross-country variability in the definition of the realisation event (Avi-Yonah et al., 2011).

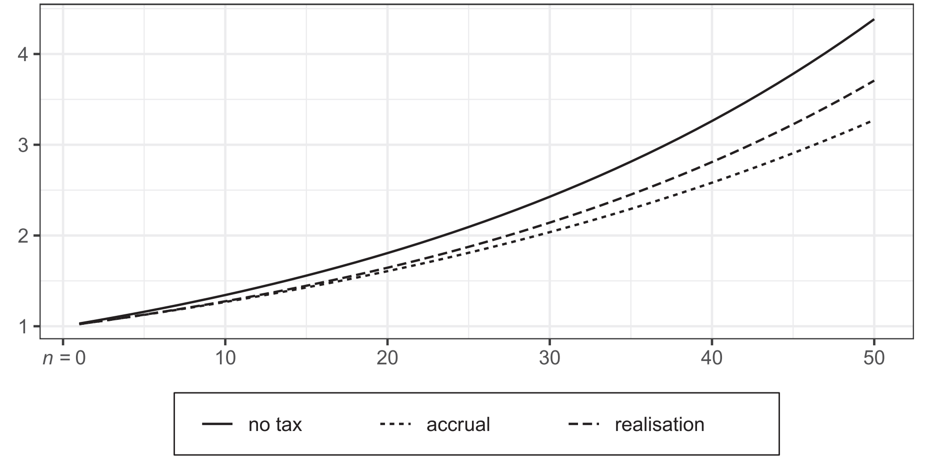
[7] Constantinides, 1983.

[8] See Auerbach (1991) and Kovenock and Rothschild (1987). Additionally, if capital losses can be fully offset against other types of income, the overall tax burden can be further reduced (down to nil) through the selective recognition of losses (Constantinides, 1983; Stiglitz, 1983, 1985). The traditional response to these problems has been to limit loss offsets and implement anti-avoidance rules, but all these remedies have ultimately been unsatisfactory and have resulted in very complicated provisions in the tax code (Alworth, Arachi and Hamaui, 2003).

[9] Harding (2013) finds that all OECD countries that tax capital gains do so on realisation. Only Italy made an attempt to introduce an accrual system in its 1998 tax reform, but this solution was abandoned a few years later (Alworth et al., 2003).

[10] See Auerbach (1991), Alworth (1998) and Alworth et al. (2003). The presumption of the superiority of accrual taxation and the acknowledgement of the practical difficulties in its implementation have prompted a long-lived debate among economists, as well as tax practitioners, on viable alternative approaches. In particular, attention has been focused on imitating an accrual tax by retrospective taxation on a realisation basis in order to circumvent the lock-in effect without running into the problems of liquidity and valuation. The basic idea was first proposed by Vickrey (1939): with full information regarding the path of asset prices and holdings, it is possible to adjust taxes paid on realisation and to equalise them to those that would have been paid if portfolios had been marked to market. The adjustment is determined by calculating ex post accrued gains in each tax year between purchase and realisation, and capitalising the implicit or 'virtual' tax payment (or tax

FIGURE 1 Value of one unit of consumption after n periods under different tax regimes (with t = 20 per cent and r = 3 per cent)



is difficult to implement in practice. This paper offers an alternative perspective: we show that there are instances where a tax on capital gains levied upon realisation can result in a lower distortion than a tax levied upon accrual, so that the efficiency-based argument for taxation upon accrual appears weaker than is commonly held.

Our argument is based on the fact that the lock-in effect induced by taxation upon realisation affects not only portfolio choices, but also the intertemporal allocation of consumption, in a way that may reduce the overall distortion compared with taxation upon accrual. This claim can be illustrated using an argument proposed by Banks and Diamond (2010) to interpret the Chamley–Judd result of a zero optimal capital income tax in the long run. Taxation of capital income and capital gains increases the relative price of future consumption, resulting in a distortion that grows with the time horizon of the investment. With a constant rate of return r and no taxes, one unit of consumption today is worth $(1 + r)^n$ units of consumption after n years. When an investor is subject to an annual tax on accrued returns at rate t, one unit of consumption today is converted into $(1 + (1 - t)r)^n$ units after n years. As a consequence, the distortion in the relative price of future consumption will be very high when n is large, even for a small t. When, instead, returns are taxed upon realisation, one unit of consumption today is converted into $1 + (1 - t)[(1 + r)^n - 1]$ units after n years, implying a lower distortion of future consumption price when compared with taxation upon accrual (the different cases are illustrated in Figure 1).[11]

This suggests that, especially when people plan over a long horizon, a shift from realisation to accrual taxation will have ambiguous effects in terms of efficiency: if, on the one hand, by removing the lock-in effect, it prevents distortions in portfolio allocation, then, on the other hand, it increases the distortion on intertemporal allocation of consumption. As a consequence, even leaving aside practical implementation issues, there is no warranty that accrual taxation would always be superior in terms of efficiency.[12] This point has been noticed by Auerbach (1992), who, however, has not further

credit) for each year using the net-of-tax rate of return. Several alternatives, with varying informational requirements, have been subsequently developed (for a survey, see Sahm, 2009). The introduction of retrospective taxation in Italy in 1998 raised strong protest and was revoked in 2001 (Alworth et al., 2003).

[11] A widely discussed issue is the impact of inflation on capital gains taxation. Because capital gains taxes (along with capital income taxes) are usually levied on nominal quantities, the effective tax rate on the real return is higher than the statutory tax rate. In the case of accrual taxation, a tax rate t on nominal return is equivalent to an effective tax rate $\hat{t} = t[1 + (1/r)\mathrm{CPI}/(1 + \mathrm{CPI})] > t$ on the real return, where CPI is the inflation rate. Under realisation-based taxation, the effective tax rate decreases with the length of the investment, so inflation contributes to increase the profitability of deferred liquidation and reinforces the lock-in effect (Brinner, 1973). In the rest of the paper, we do not explicitly consider inflation.

[12] Sahm (2008) also argues that a regime shift from realisation to accrual may not produce a Pareto improvement. He disregards the intertemporal dimension and focuses on a model where the only distortion induced by taxation upon realisation is on portfolio choices. When the distortion is eliminated by a shift from realisation to accrual, the asset equilibrium prices will change, causing distributive effects among heterogeneous

investigated the ensuing potential trade-off. In his analysis, where only some values of the parameters are taken into account, a move from realisation to accrual-based taxation results in an unequivocal efficiency gain.

In our paper, we argue that when the trade-off between different distortionary effects is properly taken into account, there may be instances where the inefficiency induced by realisation (in portfolio allocation) is lower than the one induced by accrual (in the intertemporal allocation of consumption). We show that the relative efficiency of one tax regime compared with the other depends, as expected, on the time horizon of the investments and, when considered, on risk (both the risk profile of alternative assets and the individual's risk attitude).

The paper is organised as follows. In the next section, using a simple three-period framework with a representative consumer and no uncertainty, we provide a recap of the lock-in effect and discuss how the intertemporal distortion is affected by the choice of the tax regime. In this setting, absent any incentive to adjust the portfolio, we identify some sufficient condition on preferences (i.e., intertemporal weak separability and homotheticity) for taxation upon realisation to be more efficient than taxation upon accrual. In Section 3, we provide a preliminary exploration of the trade-off between the distortion in portfolio choice and the intertemporal distortion in a deterministic setting, by assuming that in the second period there is an opportunity to invest in an alternative asset that yields a higher return. However, in this context, the portfolio allocation is an all-or-nothing choice; consequently, it is not possible to consider how the optimal mix of assets is affected by the tax regime, and the trade-off implied by the different distortions cannot be fully appreciated. For this reason, in Section 4, we introduce risk and consider the case of a consumer choosing the optimal mix of assets with different return/risk profiles. Using a numerical simulation, we show that, for a sizeable interval of the relevant parameters, the intertemporal distortion may be more relevant than the distortion on portfolio choice; therefore, we are able to conclude that, in some circumstances, a tax on capital gains levied upon realisation may be more efficient than a tax levied upon accrual.

2 | LOCK-IN EFFECT, PORTFOLIO CHOICE AND INTERTEMPORAL ALLOCATION OF CONSUMPTION

It is well known that when capital gains are taxed upon realisation, the tax system is not *holding-period neutral*;[13] that is, the net-of-tax return on savings depends on the length of the holding period of the assets in the consumer's portfolio.

This can be easily illustrated in a three-period framework. Consider the case in which all assets appreciate at rate r_1 between the first and second periods and at rate r_2 between the second and third periods. We indicate by $R = r_1 + r_2 + r_1 r_2$ the total appreciation rate over the three periods.

Let $\theta = 1 - t$ be the net-of-tax rate for a given tax rate t on capital gains. When the tax is levied upon realisation, an investor that holds the same assets for two periods will obtain an after-tax return θR. In contrast, if the assets are disposed of after one period and the net of tax proceeds are reinvested for another period, the same consumer will earn an after-tax return equal to $(1 + \theta r_1)(1 + \theta r_2) - 1$. The after-tax return is larger under the holding strategy, as

$$\theta R - [(1 + \theta r_1)(1 + \theta r_2) - 1] = r_1 r_2 (1 - \theta)\theta > 0. \tag{1}$$

By violating holding period neutrality, taxation upon realisation will distort liquidation choices. Namely, the investor will have an incentive to defer the realisation of the first period gain.[14] This

agents. As a result, even if distortions are reduced, the regime shift will not bring about a Pareto improvement. In contrast, we argue that when both the intertemporal and the portfolio distortions are taken into account, taxation upon realisation may result in higher welfare even when agents are homogeneous.

[13] Auerbach, 1991.

[14] See Constantinides (1983). This is true with a (positive) capital gain. Throughout the paper, we consider an appreciation in the second and third periods. The reason for this is that, in the case of a capital loss, under realisation-based taxation with full offset of losses, the investor would find

distortion, usually referred to as the lock-in effect, may produce an inefficient portfolio allocation when the owner of an appreciated asset chooses to forgo alternative investment opportunities in order to reap the benefit of tax deferral. To illustrate, assume that in the second period a new asset is available, yielding a return $g > r_2$. The difference between investing the proceeds from the liquidation of the original asset in the new asset and holding the original asset for the two periods is

$$(1 + \theta r_1)(1 + \theta g) - [1 + \theta R] = \theta\left[g(1 + \theta r_1) - r_2(1 + r_1)\right]. \tag{2}$$

When this difference is negative, which is the case if

$$g\frac{1 + \theta r_1}{1 + r_1} < r_2, \tag{3}$$

the investor will not find it profitable to adjust the portfolio, giving up the more rewarding alternative investment. Condition 3 shows that the lock-in effect is more likely to arise the higher the past appreciation of the asset, which affects the tax to be paid when the asset is liquidated.

The lock-in effect disappears if holding period neutrality is restored by applying the tax on capital gains upon accrual. Under accrual taxation, the timing of the tax is not affected by the timing of liquidation and the alternative investment will be preferred if and only if $g > r_2$. A shift from realisation to accrual taxation may thus produce an efficiency gain.

However, the distortion in portfolio allocation is only one of the consequences of the lock-in effect. As the proceeds of the liquidation of the assets may also be used to finance consumption, taxation upon realisation would also affect intertemporal consumption choices. Assume that no better alternative investment materialises in the second period, so that, regardless of the tax system, there is no benefit from changing the portfolio composition. In this case, any capital gains tax will result in a distortion in the intertemporal allocation of consumption that discourages longer-term savings. However, a counteracting incentive to defer consumption, such as the one determined by the lock-in effect when taxes are levied upon realisation, may result in a lower distortion.

This is best analysed by considering how capital gains taxes affect the implicit prices of future consumption under the two tax regimes.

In our three-period framework, consider a representative individual allocating a given amount of wealth M across three periods to finance consumption, with no bequest. The individual saves by investing in assets whose returns r_1 and r_2 in the two periods are taxed as capital gains.[15] Considering the intertemporal budget constraint of the consumer/investor $M = C_1 + q_2C_2 + q_3C_3$, we can analyse the effect of taxes on consumer prices. Under accrual-based taxation, such prices are

$$q_2' = \frac{1}{1 + \theta r_1}, \qquad q_3' = \frac{1}{(1 + \theta r_1)(1 + \theta r_2)}, \tag{4}$$

while under realisation-based taxation, we have

$$q_2'' = \frac{1}{1 + \theta r_1}, \qquad q_3'' = \frac{1}{1 + \theta R}. \tag{5}$$

it profitable to liquidate the asset, thus eliminating any difference between the two tax regimes. In more complex cases, where there are both gains and losses, this may create room for tax avoidance (see also footnote 8), but we do not consider this aspect in our analysis.

[15] Usually, financial instruments with fixed returns, such as bonds issued at deep discount, are taxed upon accrual following the ‘yield-to-maturity’ approach (Sahm, 2009). However, to give more realism to our assumption, the riskless asset can be thought of as a mutual fund investing in bonds or a synthetic position created by combining different option contracts, as explained, for example, by Alworth (1998). The gains or losses arising from each of the assets within the synthetic portfolio are usually treated as capital gains and taxed upon realisation.

In both cases, prices are higher than in the case of no taxation:

$$p_2 = \frac{1}{1+r_1}, \qquad p_3 = \frac{1}{(1+r_1)(1+r_2)} = \frac{1}{1+R}; \tag{6}$$

hence, a tax on capital gains introduces a wedge between the marginal rate of substitution and the marginal rate of transformation of consumption in different periods. The tax wedge between consumption in the first and second periods is the same under realisation-based and accrual-based taxation,

$$\frac{q_2' - p_2}{q_2'} = \frac{q_2'' - p_2}{q_2''} = (1-\theta)\frac{r_1}{1+r_1}, \tag{7}$$

while the tax wedge between consumption in the first and third periods under accrual-based taxation is

$$\frac{q_3' - p_3}{q_3'} = (1-\theta)\frac{R + r_1 r_2 \theta}{1+R}. \tag{8}$$

The two-period tax wedge (equation 8) is clearly higher than the single-period one (equation 7). This difference can be quite large if we take our model as a stylised representation of choices over a long time horizon, with decisions spanning decades (this translates into correspondingly high appreciation rates r_1 and r_2). As long an increasing tax wedge is not justified on efficiency grounds, the resulting welfare loss can be substantial. As remarked by Banks and Diamond (2010) and Diamond and Saez (2011), it is difficult to reconcile growing implicit tax wedges over long horizons with optimality, thus suggesting a possible role for taxation that varies with the time lapse between saving and later consumption (an example of this is tax-favoured retirement savings).

With a tax levied upon realisation, the tax wedge between consumption in the first period and consumption in the third period is

$$\frac{q_3'' - p_3}{q_3''} = (1-\theta)\frac{R}{1+R}. \tag{9}$$

Compared with taxation upon accrual in equation 8, we see that a realisation-based tax implies a lower tax wedge for consumption in the third period. Hence, a change of tax regime from accrual to realisation, which leaves the tax rate t constant, would not affect the price of second-period consumption, while it would reduce the price of third-period consumption, relative to the price of both first- and second-period consumption. Although it cannot eliminate the impact of the tax on long-term investment decisions, a move from accrual-based to realisation-based taxation could mitigate it.

In fact, the realisation-based tax could be replicated by an accrual-based tax with a tax rate that declines with the time horizon of the investment. After modifying equation 4 to allow for time-varying tax rates, it can be readily shown that a realisation-based tax with constant net-of-tax rate θ is equivalent to an accrual-based tax with time declining rates θ_1 and θ_2:[16]

$$\theta_1 = \theta, \qquad \theta_2 = \frac{1+r_1}{1+\theta r_1}\theta > \theta. \tag{10}$$

[16] Setting $\theta_1 = \theta$ to secure $q_2' = q_2''$, the expression for θ_2 follows from $q_3' = q_3''$, where q_3' is now

$$q_3' = \frac{1}{(1+\theta r_1)(1+\theta_2 r_2)}.$$

The optimal taxation literature usually finds that the optimal tax on capital income is determined by a balance between distortions of labour supply and savings decisions, on one hand, and advantages of redistribution, insurance or limiting income shifting, on the other.[17] In our analysis, we focus on a narrower issue and take as solved the problem of justifying a capital gain tax when other tax instruments are available: that is, we simply assume that a tax on capital gains must be used to raise some revenue from a representative individual,[18] and we focus on the ensuing choice between realisation and accrual on the basis of the efficiency of these two solutions.[19]

For the purpose of comparing the efficiency of the two tax regimes, we compare the tax revenue obtained when the tax rate in each case is set so that the same level of utility is granted to the taxpayer. Measured in units of the untaxed numeraire (i.e., consumption in period 1), the tax revenue is[20]

$$G = (q_2 - p_2)C_2(q_2, q_3, M) + (q_3 - p_3)C_3(q_2, q_3, M), \tag{11}$$

where $C_i(q_2, q_3, M)$ is consumer demand in period i.

From our analysis of the tax wedges, it should be clear that, in order to secure the same level of utility, the tax rate under realisation-based taxation must be higher than under accrual-based taxation. Moreover, it must be $q_2' < q_2''$ and $q_3' > q_3''$.[21] In other words, a change of the tax regime from accrual-based to realisation-based taxation, which leaves the consumer's utility constant, involves an increase in the price of second-period consumption and a decrease in the price of third-period consumption. Given that the impact on tax revenue depends on the response of consumption to price changes going in opposite directions, no general conclusion can be drawn without imposing further restrictions on the shape of the utility function. However, some clear-cut results can be reached under conditions of weak separability between consumption in the first and following periods and homotheticity.

Proposition 1. Consider an individual investing in an asset whose returns are taxed as capital gains, and let the individual preferences be represented by a utility function $u(C_1, \Phi(C_2, C_3))$ with Φ homothetic. For any tax levied upon accrual, a move to a realisation-based tax with the rate adjusted so that the utility of the individual is unaffected will increase tax revenue.

For the proof, see the Online Appendix.

The intuition is the following. Any capital gain tax will produce a wedge between present and future consumption. When preferences are weakly separable it is inefficient to further distort the allocation of the consumption among different periods in the future. In our framework, with a time-invariant statutory tax rate, it is not possible to eliminate completely the tax wedge between second- and third-period consumption (i.e., we cannot have $q_3/q_2 = p_3/p_2$). However, as discussed above, by replacing accrual with realisation taxation and appropriately increasing the tax rate, we can make the cumulative

[17] A comprehensive survey of the literature can be found in Banks and Diamond (2010). For an account of recent contributions, see Saez and Stantcheva (2018). With regard to the insurance role of capital taxation, see Golosov, Tsyvinski and Werning (2006). The case for capital taxation based on income shifting is developed by Christiansen and Tuomala (2008).

[18] Notice that our analysis does not apply to an expenditure tax, which would imply the exemption of the risk-free return to savings, thus removing the intertemporal distortion. Under an expenditure tax, an individual is taxed on the difference between the total sales and total purchase of asset in the year. However, this is different from summing up capital gains, which requires that sales are matched with past purchases occurring in a number of different years (Institute for Fiscal Studies, 1978).

[19] In a way, we are following the logic of the standard Ramsey approach, where the optimal design of a given tax instrument is discussed from the point of view of efficiency alone, imposing some restrictions on the tax instruments available.

[20] As in Ramsey's standard problem with fixed producer prices, we evaluate revenue in terms of the untaxed numeraire, namely first-period consumption. Therefore, we assume that the government can freely adjust its portfolio, without affecting before-tax prices, in order to achieve the desired allocation of expenditure across periods. In other words, in our analysis we disregard possible general equilibrium effects arising from the implementation of such allocation.

[21] From the fact that, with the same rate t (the same θ), $q_2' = q_2''$ while $q_3' > q_3''$, it follows that a slightly higher tax rate under realisation, which increases both q_2'' and q_3'', will be necessary to have the same utility in the two cases. However, because a higher t under realisation implies $q_2' < q_2''$, q_3'' must remain below q_3'.

tax relative price q_3/q_2 closer to p_3/p_2, and reduce the distortion; this is shown to be sufficient to produce an efficiency gain. This result can be thought of as a special case of the well-known result on the optimality of uniform commodity taxation[22] where we do not require the income tax rate to be set optimally.[23]

It is not difficult to show that the superiority of realisation in terms of intertemporal efficiency is not guaranteed in the absence of the specified restrictions on consumer preferences. Consider, for example, the following additively separable but non-homothetic utility function:

$$U(C_1, C_2, C_3) = (C_1/\kappa) + (C_2/\kappa)^{1-\mu} + (C_3/\kappa)^{1-\nu}. \tag{12}$$

This functional form implies no cross price effect, so the optimal tax wedges will be inversely proportional to the demand elasticities. If ν is sufficiently higher than μ, then the optimal tax structure will require $(q_3 - p_3)/q_3$ to be larger than $(q_2 - p_2)/q_2$. It is then possible that accrual-based taxation is closer to the optimal tax structure than realisation-based taxation.[24]

Although weak separability and homotheticity are quite restrictive assumptions, which cannot be expected to hold in general, they are usually satisfied by commonly used utility function specifications, which allow a simple representation of attitude towards risk and intertemporal consumption substitution. From our point of view, though based on specific assumptions on preferences, this result can be considered as a useful benchmark that neatly illustrates the potential trade-off mentioned in the introduction: while accrual-based taxation is known to be more efficient with respect to portfolio choices, realisation-based taxation is more efficient with respect to the intertemporal margin. As a consequence, we cannot take for granted that taxation upon accrual, even absent problems of implementation, would always produce a net efficiency gain.

In the rest of this paper, we explore such a trade-off, and try to assess its magnitude through numerical simulation. We start from a simple deterministic model, which extends the one used in this section.

3 | EFFICIENCY COMPARISON IN A SIMPLE DETERMINISTIC MODEL

In order to identify cases where realisation-based taxation is more efficient in spite of the fact that it may induce a distortion in portfolio choice, we provide a numerical simulation with our three-period deterministic model.

To this purpose, it is useful to explicitly specify the consumer's/investor's opportunities in each period and under the two tax regimes. Let B_1 and B_2 be the amounts invested in the available asset and liquidated after one and two periods, respectively. We assume for simplicity that such an asset gives a constant return, so that $r_1 = r_2 = r$. Let A indicate an alternative investment opportunity, available only in the second period, whose return is $g > r$. Under accrual-based taxation, we have:

$$M = C_1 + B_1 + B_2; \tag{13}$$

$$B_1(1 + \theta r) = C_2 + A; \tag{14}$$

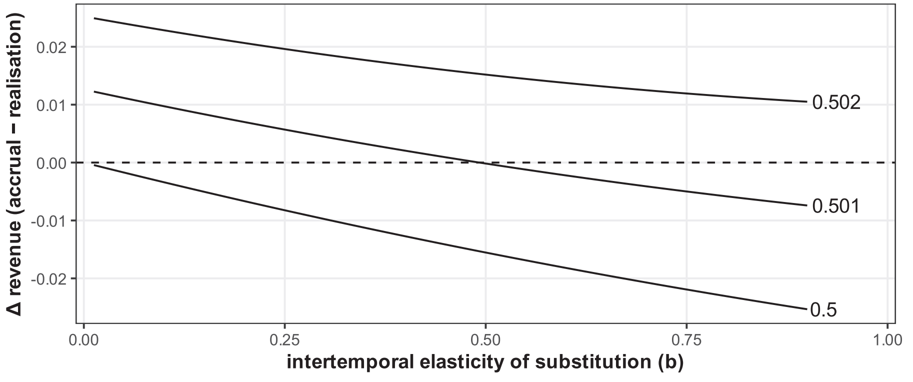
$$C_3 = B_2(1 + \theta r)^2 + A(1 + \theta g). \tag{15}$$

[22] Deaton, 1981.

[23] This is in the spirit of Laroque (2005) and Kaplow (2006).

[24] Assuming that $\kappa = 100$ with $M = 100$ and $\nu = 0.7$, accrual is more efficient than realisation for $\mu < 0.4$ (while realisation is better for $\mu > 0.4$).

FIGURE 2 Realisation versus accrual: differences in tax revenue for various levels of return g ($r = 0.5$)



Under realisation-based taxation, equation 15 must be replaced by

$$C_3 = B_2(1 + \theta R) + A(1 + \theta g). \tag{16}$$

Although, in principle, the investor can choose any mix of the two assets, when $g > r$ it will never be optimal for the investor to choose $B_2 > 0$ under accrual-based taxation, as the return will be higher investing all second-period wealth in A. However, as discussed in Section 2, we may have $1 + \theta R > (1 + \theta r)(1 + \theta g)$; in this case, when taxation is upon realisation, the investor will choose $A = 0$ and the composition of the portfolio will be suboptimal. In the following, we focus on this case. Differences in investment strategies under the two tax regimes result in different prices in the intertemporal budget constraint:

$$q_2' = q_2'' = \frac{1}{1 + \theta r}, \qquad q_3' = \frac{1}{(1 + \theta r)(1 + \theta g)}, \qquad q_3'' = \frac{1}{1 + \theta R}. \tag{17}$$

For the purpose of providing a numerical comparison, we consider the following specification of the three-period utility function:

$$U(C_1, C_2, C_3) = C_1^{1-\beta} + C_2^{1-\beta} + C_3^{1-\beta}. \tag{18}$$

Such a function is separable and homothetic, so the conditions of Proposition 1 are satisfied.

We proceed as follows. First, we calculate the utility associated with a ‘benchmark’ lump-sum tax of given revenue G (we set $G = 10$ and initial wealth $M = 100$). Then, for each capital gains tax regime, we compute the tax rate that provides the same level of utility of the lump-sum tax. Finally, given the tax rate, we calculate consumptions in each period and the resulting revenue G from the tax on capital gains, whose expression is given by equation 11. The difference with respect to the revenue raised by the lump-sum tax provides a measure of the dead-weight loss associated with each capital gains tax regime. Indeed, the relative efficiency of the two tax regimes can be directly compared by looking at their revenues G.

We conduct our numerical simulation assuming that r corresponds to a yearly risk-free rate of around 2 per cent for a period of 20 years (i.e., $r = 0.5$). We consider how the difference in revenue responds to the intertemporal elasticity $b = 1/\beta$, letting b vary between 0 and 1; we repeat the exercise at increasing values of the alternative asset return g, starting from $g = r$. The results are summarised in Figure 2.

We see that taxation upon realisation always dominates in the limiting case of $g = r$ (the lowest curve in the figure), which corresponds to the absence of any distortion in portfolio choice.[25] When $g > r$, such distortion materialises and accrual-based taxation is more efficient when the intertemporal distortion is low (i.e., for low values of b), while realisation-based taxation dominates when the intertemporal distortion is high (i.e., for high values of b).

However, we note that taxation upon realisation is desirable only for very small differences in returns between activities, while taxation upon accrual is more efficient for any b starting from values of g as low as 0.502. This might suggest that, except in extreme cases, the intertemporal distortion is less important than the distortion in the portfolio.[26] However, it should be noted that, in this simple context, portfolio reallocation is extreme even when alternative investment opportunities are only marginally more profitable: under accrual-based taxation, individuals in the second period are induced to invest all their wealth in the new asset, while under realisation-based taxation no reallocation takes place at all. In other words, we only have corner solutions.

This feature of the model, which magnifies the portfolio distortion, prevents us from fully appreciating the trade-off involved. In a more realistic model with risky assets, an investor would always hold a mix of the 'old' and the 'new' asset, in a proportion that will depend, in a continuous way, on the individual's degree of risk aversion and on the return/risk profile of the assets. Such a model is developed in the next section.

4 | EFFICIENCY COMPARISON IN A MODEL WITH RISKY ASSETS

In order to model the effect of different taxation regimes on the optimal portfolio allocation, we consider the case that the first-period asset and the new asset available in the second period have a different risk profile. Namely, we assume that the appreciation of one or the other asset is uncertain.[27] Our investor then has the possibility to modify their portfolio in order to choose the desired combination of return and risk. We expect to observe that taxation upon realisation implies a less sizeable adjustment in the portfolio.

However, the new framework also introduces the issue of taxation and risk-tasking. Since the influential paper by Domar and Musgrave (1944), economists have long studied the effect of taxation on the investment in risky assets, usually under the assumption that taxes are levied upon accrual.[28] In general, the results depends on the assumptions made on how revenue are spent, but the main insight of this literature is that fully refundable capital gains may reduce the risk of individual investments and therefore increase risk-taking. As we focus on the direct effect of taxation and ignore the additional effect produced by government expenditure, this is what we expect to observe also in our model, when comparing investment choices under a tax on capital gains and under a lump-sum tax.

When we compare taxation upon realisation with taxation upon accrual, we see two effects at work. On the one hand, because the statutory tax rate is higher under a realisation-based tax (recall that the the utility of the individual is assumed to be the same under the two regimes), the government will absorb more risk, thus encouraging the taxpayer to increase the share of wealth invested in risky assets.

[25] Because, with $g = r$, there is no difference between the asset already in the portfolio and the alternative asset, this result follows from Proposition 1.

[26] In our simulation, the highest value for which we have portfolio misallocation (i.e., the highest value compatible with condition 3), is around $g = 0.56$, with the exact value depending on the tax rate. For higher values of g, the portfolio reallocation will take place regardless of the tax regime.

[27] An asset (e.g., a share of stock) can appreciate in two different ways. First, the agent may add to the asset by not withdrawing the return, for example, by retaining earnings in a corporation that is reflected in the share price. Second, it may appreciate for some exogenous reason (e.g., changing expectation in the stock market). In either case, we can refer to the appreciation as a capital gain according to standard terminology.

[28] See, for example, Mossin (1968), Stiglitz (1969), Sandmo (1977), Gordon (1985) and Kaplow (1994).

On the other hand, under taxation upon realisation the lock-in effect will discourage the liquidation of the appreciated assets in the portfolio. This will have an ambiguous effect on risk-taking: depending on whether the asset in the portfolio is more or less risky than the alternative asset, it could lead either to more or to less risk-taking compared with accrual-based taxation.

4.1 | Model set-up

Our formalisation of the choice under risk will mostly follow Auerbach (1992). We assume that only two values of the return of the risky asset are possible, a high return g_1 and a low return g_2, occurring with probability π_1 and π_2 ($\pi_1 + \pi_2 = 1$), respectively. Accordingly, we redefine the choice set of the model in Section 3 by considering consumption in the third period contingent on the two corresponding states of the world. Namely, C_{31} and C_{32} represent consumption in the third period conditional on states 1 or 2, respectively, where the two states are defined with reference to the returns from the risky asset.

Consumer preferences will be represented by a constant elasticity of substitution (CES) utility function:[29]

$$\tilde{U}(C_1, C_2, C_{31}, C_{32}) = \left(C_1^{1-\beta} + C_2^{1-\beta} + \left(\pi_1 C_{31}^{1-\gamma} + \pi_2 C_{32}^{1-\gamma} \right)^{(1-\beta)/(1-\gamma)} \right)^{1/(1-\beta)}. \tag{19}$$

Here, $\beta, \gamma > 0$; observe that $b = 1/\beta$ is the intertemporal elasticity of substitution and $c = 1/\gamma$ is the elasticity of substitution between states, where γ is the coefficient of relative risk aversion. This functional form, which departs from the more common representation in terms of expected utility, allows us to vary independently the sizes of these two elasticities, in order to disentangle the different distortions originating from the tax and determine their magnitude.[30]

Because, in principle, a portfolio adjustment can go either way (i.e., towards more or less risk) and the effects of taxation are not necessarily the same, we consider two different cases.[31] In the first case, we maintain the assumption that the asset available in the first period is riskless, and we assume that the alternative asset A, available only in the second period, is risky. In the second case, we assume that the individual invests in the first period in a risky asset (although, for simplicity, we assume that the risk shows only in the second period), while in the second period the individual has the opportunity to invest in an alternative riskless asset. Hence, the adjustment is in the direction of a higher exposure to risk in the first case, and of a lower exposure to risk in the second. Once again, the returns from all assets, risky and safe, are taxed as capital gains, and we evaluate the efficiency of the two tax regimes, accrual and realisation, by comparing the tax revenue when the tax rate in each regime is adjusted in order to secure a given level of the taxpayer's utility.

In all cases considered, the budget constraints for the first two periods are represented by equations 13 and 14, as in the deterministic case of the preceding section, with the variables having the same interpretation as previously. Regarding consumption in the third period and its price q_3, we must consider separately the two cases of adjustment towards more or less risk.

[29] To help reproduction of the results, the expressions of the demand functions are derived in the online Appendix.

[30] A representation in terms of expected utility would require that $\beta = \gamma$. By allowing β to be different from γ, we impose less restriction on these parameters. It has been noted that the restriction implied by the expected utility assumption is often not compatible with empirical observation on consumption smoothing and risk behaviour; the use of general functional forms allowing more flexible relation between the two behavioural parameters is a possible response to this puzzle (Epstein and Zin, 1991).

[31] In this regard, our analysis departs significantly from Auerbach (1992), who considers only the case in which the adjustment is aimed at reducing the riskiness of the portfolio.

4.2 | First case: the asset in the portfolio is safe, the alternative asset is risky

With taxation upon accrual, consumption in the third period in the state of the world i will be

$$C_{3i} = B_2(1+\theta r)^2 + A(1+\theta g_i), \qquad i \in \{1,2\}. \tag{20}$$

By substituting from equations 13 and 14, we can calculate the intertemporal budget constraint and determine the third-period consumer prices:

$$q'_{31} = \frac{1}{(1+\theta r)^2} \cdot \frac{r-g_2}{g_1-g_2}, \qquad q'_{32} = \frac{1}{(1+\theta r)^2} \cdot \frac{g_1-r}{g_1-g_2}. \tag{21}$$

Under realisation-based taxation, we have

$$C_{3i} = B_2(1+\theta R) + A(1+\theta g_i), \qquad i \in \{1,2\} \tag{22}$$

so that prices are

$$\begin{aligned} q''_{31} &= \frac{1}{\theta g_1 - \theta g_2}\left(\frac{1}{1+\theta r} - \frac{1+\theta g_2}{1+\theta R}\right), \\ q''_{32} &= \frac{1}{\theta g_1 - \theta g_2}\left(\frac{1+\theta g_1}{1+\theta R} - \frac{1}{1+\theta r}\right). \end{aligned} \tag{23}$$

The effect of taxation under different regimes can be better appreciated by comparing equations 21 and 23 with the expression of consumer prices with no taxes ($t = 0$):

$$p_{31} = \frac{1}{(1+r)^2} \cdot \frac{r-g_2}{g_1-g_2}, \qquad p_{32} = \frac{1}{(1+r)^2} \cdot \frac{g_1-r}{g_1-g_2}. \tag{24}$$

We see that the relative price between consumption in the two states of the world in the third period is not affected by taxation upon accrual,

$$\frac{q'_{31}}{q'_{32}} = \frac{r-g_2}{g_1-r} = \frac{p_{31}}{p_{32}}, \tag{25}$$

while in the case of realisation, for $t > 0$, we have

$$\frac{q''_{31}}{q''_{32}} = \frac{(1+\theta R)-(1+\theta g_2)(1+\theta r)}{(1+\theta g_1)(1+\theta r)-(1+\theta R)} > \frac{r-g_2}{g_1-r} = \frac{p_{31}}{p_{32}}. \tag{26}$$

In other words, the disincentive to adjust the portfolio towards more risk-taking can be thought of as an implicit tax incentive to increase consumption in the low-return state, C_{32}.

4.3 | Second case: the asset in the portfolio is risky, the alternative asset is safe

In this case, the asset purchased in the first period is risky (even though such risk arises only in the second period) and the new asset A is safe. With taxation upon accrual, we have

$$C_{3i} = B_2(1+\theta r)(1+\theta g_i) + A(1+\theta r). \tag{27}$$

We see that in this case the intertemporal budget constraint is the same as in the previous case, so prices are as in equation 21.

With taxation upon realisation, consumption in the third period in state i is

$$C_{3i} = B_2(1+\theta R_i) + A(1+\theta r), \qquad i \in \{1,2\}, \tag{28}$$

where $R_i = (1+r)(1+g_i) - 1 = g_i + r + g_i r$ indicates the return of the risky asset over the two periods in state i. Therefore, consumer prices in the third period are

$$q''_{31} = \frac{1}{\theta R_1 - \theta R_2}\left[1 - \frac{1+\theta R_2}{(1+\theta r)^2}\right],$$
$$q''_{32} = \frac{1}{\theta R_1 - \theta R_2}\left[\frac{1+\theta R_1}{(1+\theta r)^2} - 1\right]. \tag{29}$$

It follows that relative prices are

$$\frac{q''_{31}}{q''_{32}} = \frac{(1+\theta r)^2 - (1+\theta R_2)}{(1+\theta R_1) - (1+\theta r)^2} < \frac{r-g_2}{g_1-r} = \frac{p_{31}}{p_{32}}. \tag{30}$$

Hence, realisation-based taxation, by discouraging the adjustment towards less risk-taking (it makes the alternative safe asset A less attractive), creates an incentive to consume more in the high return state.

4.4 | Simulation results

In order to compare accrual-based and realisation-based taxation, we follow the same steps as in Section 3. Namely, we take as benchmark the utility of an individual with wealth $M = 100$ subject to a lump-sum tax $G = 10$, and we calculate the tax rate and the corresponding tax revenue, which guarantees the same taxpayer's utility under different tax regimes. Tax revenue is now given by[32]

$$G = (q_2 - p_2)C_2 + (q_{31} - p_{31})C_{31} + (q_{32} - p_{32})C_{32}. \tag{31}$$

Once again, the two tax regimes can be directly compared with each other and with the case of lump-sum taxation, looking at their revenues G.

In order to carry out the simulation, it is necessary to specify reasonable values of the rates of return (and variance) of the assets involved. According to the estimates in Jordà et al. (2019), the

[32] This simply extends equation 11 to the case of two states of the world in the third period: the government discounts tax revenue at different times and states of the world using state-contingent prices (see also footnote 20). It is worth noting that the government has no preferences on the time or risk profile of its revenue, as we assume that it can freely transfer resources across periods and states through financial markets.

TABLE 1 Summary of the assumptions about the rates of risky and safe assets

Period length	r	g_1	g_2	$\pi_1 g_1 + \pi_2 g_2$
20 years	0.50	5.8	0	2.9
25 years	0.65	9.0	0	4.5

yearly rate of return on safe assets in peace time in the last century has been in the range of 1–3 per cent,[33] while riskier investments in housing and equity have yielded on average around 7–8 per cent per year in the post-1950 period, with standard deviations ranging from 10 per cent (housing) to 24 per cent (equity). Hence, we assume for our exercise a 2 per cent return for the safe asset and a 7 per cent average return for the risky asset; because each of our periods covers many years, these rates translate into different cumulative returns r and g_i depending on the assumed time span. For example, considering periods of 20 years, we have a cumulative safe return $r = 0.5$ and (assuming independence of returns over time) an average risky return of around 2.9. In our two-state framework, such a return is obtained by assuming $g_1 = 5.8$ and $g_2 = 0$ with probabilities $\pi_1 = \pi_2 = 0.5$; these values imply a standard deviation of 2.9 for the cumulative return, corresponding to 0.16 on a yearly basis, which lies within the range of the estimates for the yearly standard deviation mentioned above. Similarly, we find values of r and g_2 for the case of 25-year periods. Table 1 summarises our assumptions.

4.4.1 | First case: the asset in the portfolio is safe, the alternative asset is risky

The results of the simulations for the case where the asset in the portfolio is safe and the portfolio adjustment in the second period is towards more risk-taking are shown in Tables 2 and 3 for the cases of 20 and 25 years, respectively. We present our results for different values of the intertemporal elasticity of substitution (measured by b) and the elasticity of substitution of consumption between states (measured by c, which is the inverse of the coefficient of risk aversion). For each simulation, the columns LUMP, ACCR and REAL refer to lump-sum taxation (considered as a benchmark), accrual and realisation, respectively.

For each value of $b = 1/\beta$ and $c = 1/\gamma$, the tables show consumption in the second and third periods. In order to analyse the effect on portfolio composition, we also calculated the ratio a between the alternative asset (A) and total wealth at the end of the second period, $B_2(1 + \theta r) + A$ (with $\theta = 1$ in the LUMP case).[34]

As expected, consumption in the second period increases with b and third-period consumption is more equally distributed among the two possible states the lower is c (i.e., the higher risk aversion is). The tax rate is always higher under realisation and, for all combinations of the parameters c and b we have considered, the share of the risky asset a is lower in the LUMP case than in the ACCR and REAL cases. This is the result of the implicit insurance component of capital gains taxation we have highlighted above: the tax absorbs some of the risk and therefore encourages private risk-taking. The insurance effect is stronger in the REAL case, given that the tax rate is higher. However, the lock-in effect counteracts the incentive to risk-taking by discouraging a switch to the alternative (risky) asset in the second period. As a consequence, a is usually higher in the ACCR case; an exception is the leftmost column in Table 3, where the resulting distortion in the portfolio is minimum due to the high-level risk aversion, which makes the acquisition of the risky asset unattractive in any case.

[33] Estimates by King and Low (2014) show that the real interest rate has been declining from more than 4 per cent at their peak in the late 1980s to less than 1 per cent in the early 2010s.

[34] Note that, to allow comparison, we consider the value of B_2 at the end of the second period, net of taxes, both for accrual and for realisation; in the case of realisation, this is the value the investor would obtain by liquidating all their wealth at that time. Thus, a varies between 0 (no change in the portfolio) to 1 (all wealth is liquidated and reinvested in the new asset).

TABLE 2 Asset in the portfolio is safe, alternative asset is risky (one period = 20 years)

		$c = 0.025$			$c = 0.2$			$c = 0.4$			$c = 0.6$		
		LUMP	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**
$b = 0.025$	G	**10**	**9.986**	**9.986**	10	9.985	9.979	10	9.985	9.968	10	9.985	9.956
	t	0	0.3423	0.3724	0	0.3492	0.3737	0	0.3581	0.375	0	0.368	0.3761
	C_2	45.86	45.82	45.8	46.4	46.35	46.34	47.06	47.01	47	47.76	47.7	47.69
	C_{31}	48.21	47.99	47.82	62.55	62.25	60.43	81.69	81.28	76.9	102.6	102.1	94.68
	C_{32}	45.78	45.57	45.62	41.35	41.15	41.42	35.69	35.52	36.14	29.64	29.48	30.51
	a	0.009438	0.01976	0.01951	0.08502	0.1828	0.176	0.1922	0.4286	0.4001	0.3192	0.7427	0.6719
$b = 0.2$	G	**10**	**9.896**	**9.904**	**10**	**9.894**	**9.894**	10	9.891	9.88	10	9.889	9.863
	t	0	0.3238	0.3538	0	0.3293	0.3539	0	0.3363	0.3535	0	0.3442	0.3528
	C_2	47.02	46.73	46.6	47.53	47.21	47.11	48.17	47.81	47.74	48.84	48.45	48.41
	C_{31}	54.04	52.26	52.26	70.85	68.45	66.75	93.82	90.52	86.12	119.7	115.4	107.7
	C_{32}	51.31	49.62	49.83	46.84	45.25	45.66	41	39.56	40.29	34.58	33.32	34.44
	a	0.009833	0.01938	0.01918	0.08902	0.179	0.173	0.2027	0.4187	0.3932	0.3395	0.7236	0.6603
$b = 0.4$	G	**10**	**9.813**	**9.829**	**10**	**9.81**	**9.817**	10	9.806	9.8	10	9.802	9.78
	t	0	0.3041	0.3339	0	0.3079	0.3326	0	0.313	0.3304	0	0.3186	0.3277
	C_2	48.1	47.68	47.42	48.55	48.08	47.87	49.11	48.6	48.44	49.72	49.15	49.07
	C_{31}	61.21	57.65	57.84	81.18	76.35	74.79	109.1	102.5	98.01	141.7	132.7	124.7
	C_{32}	58.12	54.74	55.14	53.66	50.47	51.04	47.68	44.77	45.63	40.92	38.34	39.57
	a	0.01025	0.019	0.01885	0.09323	0.1752	0.17	0.2137	0.4087	0.3863	0.3608	0.7045	0.6487
$b = 0.6$	G	**10**	**9.749**	**9.771**	**10**	**9.746**	**9.759**	10	9.742	9.741	10	9.738	9.719
	t	0	0.2858	0.3154	0	0.2883	0.3128	0	0.2914	0.309	0	0.2949	0.3044
	C_2	48.89	48.48	48.1	49.23	48.78	48.46	49.66	49.17	48.94	50.14	49.6	49.47
	C_{31}	68.88	63.58	63.96	92.35	85.13	83.7	125.9	115.9	111.4	166.1	152.6	144.1
	C_{32}	65.41	60.37	60.96	61.04	56.27	57.01	55.02	50.63	51.63	47.98	44.07	45.42
	a	0.01062	0.01866	0.01855	0.09703	0.1718	0.1673	0.2236	0.4001	0.3803	0.38	0.6881	0.6387

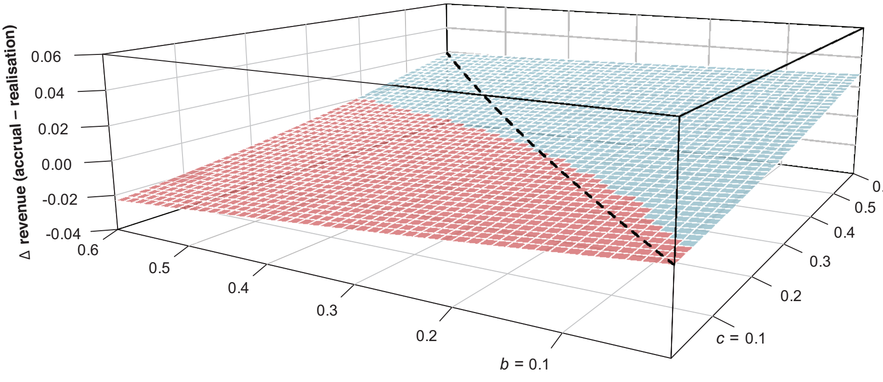
Note: $r = 0.65$, $g_1 = 5.8$ ($\pi_1 = 0.5$), $g_2 = 0$ ($\pi_2 = 0.5$), and $M = 100$.

TABLE 3 Asset in the porfolio is safe, alternative asset is risky (one period = 25 years)

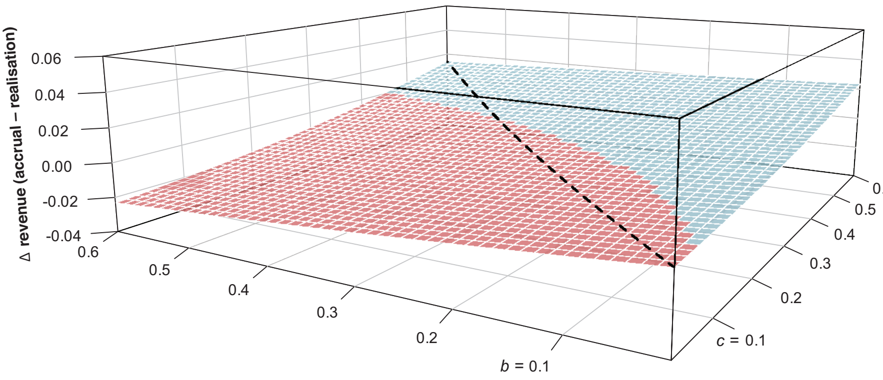
		$c = 0.025$			$c = 0.2$			$c = 0.4$			$c = 0.6$		
		LUMP	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**
$b = 0.025$	G	**10**	**9.986**	**9.986**	10	9.985	9.981	10	9.985	9.972	10	9.984	9.961
	t	0	0.3426	0.3728	0	0.3515	0.3767	0	0.3637	0.3813	0	0.3778	0.3861
	C_2	45.88	45.84	45.82	46.57	46.52	46.51	47.46	47.4	47.39	48.41	48.34	48.34
	C_{31}	48.76	48.54	48.37	68.29	67.96	65.94	96.28	95.79	90.46	128.3	127.6	118
	C_{32}	45.75	45.54	45.58	40.98	40.79	40.98	34.67	34.5	34.97	27.73	27.58	28.4
	a	0.007552	0.01582	0.01597	0.07134	0.1548	0.1524	0.1702	0.3886	0.37	0.2963	0.7208	0.662
$b = 0.2$	G	**10**	**9.896**	**9.904**	**10**	**9.893**	**9.896**	10	9.89	9.883	10	9.886	9.866
	t	0	0.324	0.3541	0	0.3311	0.3564	0	0.3407	0.3588	0	0.352	0.3609
	C_2	47.04	46.75	46.62	47.7	47.37	47.26	48.55	48.18	48.1	49.49	49.05	49.01
	C_{31}	54.68	52.87	52.87	77.64	74.98	73.09	111.6	107.5	102.2	152.1	146.3	136.3
	C_{32}	51.3	49.61	49.81	46.59	45	45.33	40.18	38.73	39.31	32.87	31.63	32.54
	a	0.007869	0.01552	0.01568	0.07483	0.1515	0.1495	0.1804	0.3791	0.3628	0.3181	0.7003	0.6485
$b = 0.4$	G	**10**	**9.813**	**9.829**	**10**	**9.809**	**9.819**	10	9.804	9.803	10	9.799	9.782
	t	0	0.3042	0.3341	0	0.3093	0.3346	0	0.3161	0.3345	0	0.3242	0.3337
	C_2	48.12	47.69	47.44	48.69	48.22	48	49.45	48.91	48.75	50.3	49.68	49.59
	C_{31}	61.96	58.35	58.55	89.3	83.95	82.22	131	122.9	117.4	183.1	171.2	160.6
	C_{32}	58.13	54.74	55.14	53.59	50.38	50.88	47.18	44.25	44.96	39.58	37.01	38.04
	a	0.008202	0.01521	0.01538	0.0785	0.1482	0.1466	0.191	0.3695	0.3555	0.341	0.6801	0.635
$b = 0.6$	G	**10**	**9.748**	**9.771**	**10**	**9.745**	**9.761**	**10**	**9.74**	**9.743**	10	9.734	9.721
	t	0	0.2859	0.3156	0	0.2891	0.3144	0	0.2934	0.312	0	0.2985	0.3084
	C_2	48.9	48.49	48.1	49.34	48.89	48.55	49.93	49.41	49.17	50.6	50.01	49.88
	C_{31}	69.75	64.38	64.77	102	93.94	92.37	152.5	140.2	134.7	218.2	200	188.8
	C_{32}	65.44	60.4	60.98	61.19	56.38	57.04	54.93	50.5	51.35	47.16	43.24	44.39
	a	0.008504	0.01494	0.01512	0.08182	0.1453	0.144	0.2006	0.3613	0.3491	0.3615	0.6627	0.6234

Note: $r = 0.65$, $g_1 = 9$ $(\pi_1 = 0.5)$, $g_2 = 0$ $(\pi_2 = 0.5)$ and $M = 100$.

FIGURE 3 Difference in revenue between accrual and realisation, when the asset in the portfolio is safe, and the alternative asset is risky



(a) 20-year investment ($r = 0.5$, $g_2 = 5.8$ and $g_2 = 0$)



(b) 25-year investment ($r = 0.65$, $g_2 = 9$ and $g_2 = 0$)

In Tables 2–5, the bold font identifies the cases in which realisation dominates accrual in terms of revenue. The same comparison is proposed in Figure 3, where the surface shows the difference between revenue under accrual and under realisation, depending on the values of the elasticities. The values of the elasticities where the difference is negative are indicated in red on the surface; this is where, despite the portfolio distortion, taxation at realisation is more efficient than taxation upon accrual.

We see that realisation tends to dominate accrual where b is high and c is low. This is not surprising in the light of the discussion in previous sections: the higher b is, the greater the intertemporal distortion produced by taxation upon accrual; the lower c is, the smaller the distortion due to the portfolio distortion brought about by taxation upon realisation.

The dashed lines on the surfaces in the graphs of Figure 3 identify points where $b = c$, that is, the case where the coefficient of risk aversion ($\gamma = 1/c$) is the inverse of the intertemporal elasticity of substitution, consistently with the standard assumption of the expected utility approach. It is worth noting that $c < b$, where realisation tends to dominate accrual, is the case often assumed in finance in order to account for observed behaviour under risk (e.g., in the analysis of the risk equity premium).[35]

[35] As emphasised, for example, by Campbell (2000, p. 30), utility specifications where the elasticity of intertemporal substitution (EIS) differs from the reciprocal of risk aversion allow us to 'postulate high risk aversion to resolve the equity premium puzzle without driving the EIS to an unreasonably low value'.

TABLE 4 Asset in the portfolio is risky, alternative asset is safe (one period = 20 years)

		$c = 0.025$			**$c = 0.2$**			**$c = 0.4$**			**$c = 0.6$**		
		LUMP	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**
$b = 0.025$	G	10	9.987	9.987	10	9.987	9.982	10	9.986	9.974	10	9.986	9.96
	t	0	0.3784	0.3789	0	0.3874	0.3926	0	0.3996	0.4122	0	0.4135	0.4362
	C_2	42.83	42.8	42.8	43.5	43.47	43.46	44.35	44.3	44.3	45.25	45.2	45.19
	C_{31}	45.21	45.03	45.18	61.51	61.25	62.91	84.26	83.88	88.39	109.9	109.4	117.9
	C_{32}	42.62	42.45	42.43	38.36	38.2	38.05	32.77	32.62	32.24	26.66	26.53	25.81
	a	0.9897	0.978	0.9766	0.9048	0.7891	0.7696	0.777	0.4834	0.4148	0.6178	0.0646	−0.1073
$b = 0.2$	G	10	9.905	9.905	10	9.903	9.9	10	9.9	9.89	10	9.897	9.877
	t	0	0.3625	0.3631	0	0.3697	0.3748	0	0.3794	0.3918	0	0.3905	0.413
	C_2	43.54	43.34	43.34	44.16	43.93	43.91	44.94	44.67	44.63	45.79	45.48	45.4
	C_{31}	49.43	47.95	48.11	68.06	65.95	67.68	94.77	91.7	96.48	126	121.7	130.9
	C_{32}	46.6	45.2	45.19	42.45	41.13	41	36.86	35.66	35.31	30.57	29.53	28.84
	a	0.9894	0.9784	0.9771	0.9014	0.7936	0.7755	0.7673	0.4963	0.4331	0.5972	0.09186	−0.06502
$b = 0.4$	G	10	9.826	9.825	10	9.822	9.82	10	9.818	9.811	10	9.813	9.799
	t	0	0.3454	0.3459	0	0.3505	0.3555	0	0.3575	0.3697	0	0.3655	0.3878
	C_2	44.19	43.9	43.89	44.71	44.37	44.34	45.38	44.99	44.91	46.12	45.67	45.52
	C_{31}	54.52	51.56	51.72	76.05	71.82	73.63	107.8	101.6	106.7	146.4	137.6	147.7
	C_{32}	51.39	48.61	48.6	47.43	44.79	44.68	41.93	39.52	39.19	35.52	33.39	32.73
	a	0.9891	0.9789	0.9776	0.8978	0.7983	0.7815	0.757	0.5094	0.4517	0.5756	0.1193	−0.02244
$b = 0.6$	G	10	9.76	9.76	10	9.757	9.756	10	9.753	9.749	10	9.748	9.738
	t	0	0.3292	0.3297	0	0.3325	0.3373	0	0.3369	0.3489	0	0.342	0.364
	C_2	44.66	44.37	44.36	45.04	44.72	44.67	45.55	45.17	45.05	46.11	45.68	45.45
	C_{31}	59.86	55.47	55.64	84.55	78.23	80.13	121.9	112.6	118	168.9	155.6	166.4
	C_{32}	56.43	52.29	52.28	52.73	48.79	48.69	47.41	43.78	43.48	40.96	37.73	37.1
	a	0.9888	0.9792	0.9781	0.8945	0.8024	0.7869	0.7476	0.5209	0.468	0.5559	0.1432	0.01461

Note: $r = 0.5$, $g_1 = 5.8$ ($\pi_1 = 0.5$), $g_2 = 0$ ($\pi_2 = 0.5$) and $M = 100$.

TABLE 5 Asset in the portfolio is risky, alternative asset is safe (one period = 25 years)

		c = 0.025			**c = 0.2**			**c = 0.4**			**c = 0.6**		
		LUMP	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**	**LUMP**	**ACCR**	**REAL**
$b = 0.025$	G	10	9.986	9.985	10	9.985	9.981	10	9.985	9.973	10	9.984	9.96
	t	0	0.3426	0.343	0	0.3515	0.3556	0	0.3637	0.374	0	0.3778	0.3966
	C_2	45.88	45.84	45.84	46.57	46.52	46.52	47.46	47.4	47.39	48.41	48.34	48.33
	C_{31}	48.76	48.54	48.71	68.29	67.96	69.96	96.28	95.79	101.4	128.3	127.6	138.4
	C_{32}	45.75	45.54	45.52	40.98	40.79	40.65	34.67	34.5	34.11	27.73	27.58	26.83
	a	0.9924	0.9842	0.9832	0.9287	0.8452	0.8311	0.8298	0.6114	0.5607	0.7037	0.2792	0.1491
$b = 0.2$	G	10	9.896	9.896	10	9.893	9.891	10	9.89	9.882	10	9.886	9.868
	t	0	0.324	0.3245	0	0.3311	0.3351	0	0.3407	0.3508	0	0.352	0.3707
	C_2	47.04	46.75	46.75	47.7	47.37	47.36	48.55	48.18	48.13	49.49	49.05	48.97
	C_{31}	54.68	52.87	53.06	77.64	74.98	77.08	111.6	107.5	113.6	152.1	146.3	158.3
	C_{32}	51.3	49.61	49.59	46.59	45	44.87	40.18	38.73	38.37	32.87	31.63	30.9
	a	0.9921	0.9845	0.9835	0.9252	0.8485	0.8356	0.8196	0.6209	0.5749	0.6819	0.2997	0.1825
$b = 0.4$	G	10	9.813	9.813	10	9.809	9.808	10	9.804	9.799	10	9.799	9.786
	t	0	0.3042	0.3046	0	0.3093	0.3132	0	0.3161	0.326	0	0.3242	0.3427
	C_2	48.12	47.69	47.69	48.69	48.22	48.18	49.45	48.91	48.82	50.3	49.68	49.51
	C_{31}	61.96	58.35	58.54	89.3	83.95	86.17	131	122.9	129.3	183.1	171.2	184.4
	C_{32}	58.13	54.74	54.74	53.59	50.38	50.27	47.18	44.25	43.91	39.58	37.01	36.31
	a	0.9918	0.9848	0.9839	0.9215	0.8518	0.8401	0.809	0.6305	0.5891	0.659	0.3199	0.2157
$b = 0.6$	G	**10**	**9.748**	**9.748**	10	9.745	9.745	10	9.74	9.737	10	9.734	9.727
	t	0	0.2859	0.2863	0	0.2891	0.2929	0	0.2934	0.3031	0	0.2985	0.3168
	C_2	48.9	48.49	48.48	49.34	48.89	48.84	49.93	49.41	49.28	50.6	50.01	49.76
	C_{31}	69.75	64.38	64.58	102	93.94	96.29	152.5	140.2	147.2	218.2	200	214.6
	C_{32}	65.44	60.4	60.39	61.19	56.38	56.29	54.93	50.5	50.19	47.16	43.24	42.57
	a	0.9915	0.9851	0.9843	0.9182	0.8547	0.8441	0.7994	0.6387	0.6014	0.6385	0.3373	0.2441

Note: $r = 0.65$, $g_1 = 9$ ($\pi_1 = 0.5$), $g_2 = 0$ ($\pi_2 = 0.5$) and $M = 100$.

By comparing the two simulations, we see that the range of values of the elasticities where taxation upon realisation is more efficient than taxation upon accrual is larger in the second simulation, where each period is 25 years long, than in the first, where it is only 20 years. Again, this is consistent with our initial observation that the intertemporal distortion of accrual taxation tends to increase with time relative to taxation at realisation.

We finally observe that differences in revenue are quite small in all cases considered: in our simulations, they barely exceed 0.02, corresponding to 0.2 per cent of the tax revenue.

4.4.2 | Second case: the asset in the portfolio is risky, the alternative asset is safe

In Tables 4 and 5, we consider the case in which the asset in the portfolio is risky, so that the reallocation in the second period is towards a reduction in risk exposure. The simulations show that in this case taxation upon accrual generally dominates taxation at realisation, except for extreme values of risk aversion (low c) and relatively high values of the intertemporal elasticity b. The wider superiority of accrual taxation in this case can be explained by the fact that the dead-weight loss produced by the portfolio distortion depends on two countervailing factors: the marginal distortion brought about by the change in relative prices and the size of the optimal portfolio adjustment. When risk aversion is high, the marginal distortion due to the tax levied upon realisation is low. However, when the asset in the portfolio is safe, the higher the risk aversion, the higher the portfolio adjustment, which is optimal in the second period. Indeed, we can see that in the previous case, in which the asset in portfolio was safe, the portfolio adjustment was smaller, the higher the risk aversion (the value of a in Tables 2 and 3 was decreasing in c). As a consequence, the total efficiency loss produced by taxation upon realisation was lower with a higher risk aversion, because there was both a reduction of the distortion at the margin and a reduction in the size of portfolio adjustment. In contrast, when the asset in the portfolio is risky, a higher risk aversion (a higher c) increases the portfolio adjustment, as shown by the value of a (see Tables 4 and 5). As a result, the efficiency loss determined by taxation upon realisation is larger in this case compared with the previous case.

Notice, now, that a is the share of the safe asset, so it is not surprising that it has lower values under LUMP than under REAL and ACCR, as taxation encourages more risk exposure, and hence a smaller portfolio adjustment. The fact that now under realisation-based taxation the higher tax rate and the lock-in effect both play against a switch to the safe asset explains why a is always lower (and sometimes notably lower) under REAL than under ACCR.[36]

Finally, we note that, although the inefficiency of accrual is almost always lower than that of realisation, the two inefficiencies tend to balance each other out. Indeed, the simulations show that in this case all differences in revenue are even smaller than in the previous case, almost everywhere below 0.01 (i.e., 0.1 per cent of tax revenue).

5 | CONCLUSION

In this paper, we have questioned the common wisdom that it is more efficient to tax capital gains when they accrue rather than when they are realised. The common view emphasises the efficiency loss due to the lock-in effect, which arises when taxes are levied upon realisation because the investor may find it optimal to hold an appreciated asset even if an alternative investment provides a higher before-tax rate of return. Such a view disregards the fact that, when people plan over a long horizon, the incentive to defer liquidation of assets due to realisation-based taxation can reduce the distortion in the intertemporal allocation of consumption.

[36] Note that, for high values of c in Table 4, a takes negative values: because, in our simulation, we have not imposed a non-negativity constraint on the assets in portfolio, when risk aversion is low, the investor may have an incentive to sell the safe asset A to increase the return in the third period.

We have compared the two different tax regimes by taking into account both the portfolio and the intertemporal distortions. To this end, we have first identified a case where a trade-off may actually arise: when preferences are weakly separable and homothetic, if we take into account only the intertemporal allocation problem and the distortion induced by taxation along this dimension, realisation dominates accrual. Then, developing a numerical simulation that takes into account both the intertemporal distortion and the portfolio adjustment resulting from the lock-in effect, we have verified that realisation-based taxation may be preferable in some cases.

Namely, we have shown that there are cases in which the larger intertemporal distortion induced by accrual taxation dominates the portfolio distortion induced by realisation taxation. In particular, when optimal portfolio adjustment is towards more exposure to risk, realisation-based taxation will be more efficient than accrual-based taxation (i.e., the inefficiency from the intertemporal distortion will be larger than the inefficiency resulting from a suboptimal portfolio) when taxpayer's preferences display a high intertemporal elasticity of substitution and a high degree of risk aversion. This occurs for a broad range of values of the elasticities when the time span is long enough and when the optimal portfolio adjustment is towards more risky assets.

Our analysis also delivers an additional insight. Even when accrual dominates, the efficiency gain, measured in terms of additional revenue for a given utility level, is likely to be quite small – in our simulation, the magnitude is never larger, and usually much smaller, than 1 per cent of the revenue from the capital gains tax. This suggests that costs of compliance, even when they are low, are likely to offset the efficiency gain from accrual taxation. This may help explain why accrual or retrospective taxation of capital gains are not usually adopted. We have also highlighted that the two tax regimes do not have a clear differential impact on risk-taking. Because the statutory tax rate is higher under the realisation tax system, the government will absorb more risk, thus encouraging the taxpayer to increase the share of wealth invested in risky assets. However, the lock-in effect discourages the liquidation of appreciated assets and it may reinforce or counteract risk-taking, depending on whether the appreciated assets are relatively risky or safe.

Notice that the beneficial effect stemming from the mitigation of the intertemporal distortion of consumption suggests that it could be efficient to defer taxation even on capital income (e.g., interest and dividends), as long as it remunerates long-term savings.[37] This may help to rationalise the practice in several European countries (e.g., France, Ireland, Italy, Luxembourg and Spain), where the income of investment funds is taxed only on disposal or redemption of fund shares, even if the income is qualified as interests or dividends.[38]

A limitation of the Ramsey-type approach followed in this paper is that no justification is given for taxing capital gains in the first place. In particular, by assuming a representative household, we neglected any distributive advantage of levying taxes on capital income. Obviously, under these assumptions, it would be optimal to remove all intertemporal distortions: the tax wedge on capital income should be zero and revenue should be efficiently collected through a lump-sum levy. However, unless one argues that an optimal tax must involve a tax wedge increasing indefinitely with time, as implied by accrual taxation, the trade-off we have studied in our simplified framework would still be present in a more complex setting; such a trade-off depends on the relative costs of two different ways of implementing any optimal level of capital gains taxation. In a richer framework, the relative merits of realisation and accrual would also depend on any differential effects these two alternative approaches for taxing capital gains may have on the ability of governments to redistribute income, provide insurance or fight income shifting. The analysis of these effects is left for future research.

[37] There is no equally compelling reason to allow deferral in the case of short-term savings in liquid assets, such as a bank deposit.

[38] Branzoli et al., 2020.

ACKNOWLEDGEMENTS

In addition to the editor and three anonymous referees, we wish to thank Julian Alworth and Avraham Tabbach for their helpful comments. All responsibility for the analysis lies with the authors. The authors have no conflicts of interest with respect to the research reported in this paper.

REFERENCES

Alworth, J. (1998), Taxation and integrated financial markets: the challenges of derivatives and other financial innovations. *International Tax and Public Finance*, 5, 507–34.

Alworth, J., Arachi, G. & Hamaui, R. (2003), What's come to perfection perishes: adjusting capital gains taxation in Italy. *National Tax Journal*, 56, 197–219.

Auerbach, A. J. (1991), Retrospective capital gains taxation. *American Economic Review*, 81(1), 167–78.

Auerbach, A. J. (1992), On the design and reform of capital-gains taxation. *American Economic Review (Papers and Proceedings)*, 82, 263–7.

Ault, H. J., Arnold, B. J. & Cooper, G. S. (eds.) (2019), *Comparative Income Taxation. A Structural Analysis*, 4th edn, Alphen aan den Rijn: Kluwer Law International.

Avi-Yonah, R., Sartori, N. & Marian, O. (2011), *Global Perspectives on Income Taxation Law*, Oxford: Oxford University Press.

Banks, J. & Diamond, P. (2010), The base for direct taxation. In Institute for Fiscal Studies (ed.), *Dimensions of Tax Design: The Mirrlees Review*, chapter 6, Oxford: Oxford University Press, 548–648.

Branzoli, N., Pisano, E., Zangari, E., Messina, G. & Ricotti, G. (2020), The taxation of savings: the Italian system and international comparison. *Economia Pubblica*, 7–54.

Brinner, R. (1973), Inflation, deferral and the neutral taxation of capital gains. *National Tax Journal*, 26, 565–73.

Campbell, J. Y. (2000), Asset pricing at the Millennium. *Journal of Finance*, 55, 1515–67.

Christiansen, V. & Tuomala, M. (2008), On taxing capital income with income shifting. *International Tax and Public Finance*, 15, 527–45.

Constantinides, G. M. (1983), Capital market equilibrium with personal tax. *Econometrica*, 51, 611–36.

Deaton, A. (1981), Optimal taxes and the structure of preferences. *Econometrica*, 49, 1245–60.

Diamond, P. & Saez, E. (2011), The case for a progressive tax: from basic research to policy recommendation. *Journal of Economic Perspectives*, 25(4), 165–90.

Domar, E. D. & Musgrave, R. A. (1944), Proportional income taxation and risk-taking. *Quarterly Journal of Economics*, 58, 388–422.

Epstein, L. G. & Zin, S. E. (1991), Substitution, risk aversion, and the temporal behavior of consumption and asset returns: an empirical analysis. *Journal of Political Economy*, 99, 263–86.

Erickson, M., Hanlon, M., Maydew, E. L. & Shevlin, T. (eds.) (2020), *Scholes and Wolfson's Taxes and Business Strategy*, 6th edn, Cambridge: Cambridge Business Publishers.

Evans, C. & Krever, R. (2017), Taxing capital gains: a comparative analysis and lessons for New Zealand. *New Zealand Journal of Taxation Law and Policy*, 23, 486–515.

Golosov, M., Tsyvinski, A. & Werning, I. (2006), New dynamic public finance: a user's guide. In *NBER Macroeconomics Annual 2006*, Vol. 21, Cambridge, MA: National Bureau of Economic Research, 317–88.

Gordon, R. (1985), Taxation of corporate capital income: tax revenues versus tax distortions. *Quarterly Journal of Economics*, 100, 1–27.

Harding, M. (2013), Taxation of dividend, interest, and capital gain income. OECD Taxation Working Papers 19, Paris: OECD Publishing.

Institute for Fiscal Studies (1978), *The Structure and Reform of Direct Taxation. Report of a Committee chaired by Professor J. E. Meade*, London: Allen & Unwin.

Jordà, O., Knoll, K., Kuvshinov, D., Schularick, M. & Taylor, A. M. (2019), The rate of return on everything, 1870–2015. *Quarterly Journal of Economics*, 134, 1225–98.

Kaplow, L. (1994), Taxation and risk-taking: a general equilibrium perspective. *National Tax Journal*, 47, 789–98.

Kaplow, L. (2006), On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics*, 90, 1235–50.

King, M. & Low, D. (2014), Measuring the 'world' real interest rate. National Bureau of Economic Research (NBER) Working Paper 19887.

Kovenock, D. J. & Rothschild, M. (1987), Notes on the effect of capital gain taxation on non-Austrian assets. In A. Razin and E. Sadka (eds.), *Economic Policy in Theory and Practice*, Hong Kong: Macmillan, 309–39.

Laroque, G. R. (2005), Indirect taxation is superfluous under separability and taste homogeneity: a simple proof. *Economics Letters*, 87, 141–4.

Mossin, J. (1968), Taxation and risk-taking: an expected utility approach. *Economica*, 35, 74–82.

Saez, E. & Stantcheva, S. (2018), A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162, 120–42.

Sahm, M. (2008), Methods of capital gains taxation and the impact on asset prices and welfare. *National Tax Journal*, 61, 743–68.

Sahm, M. (2009), Imitating accrual taxation on a realization basis. *Journal of Economic Surveys*, 23, 734–61.

Sandmo, A. (1977), Portfolio theory, asset demand and taxation: comparative statics with many assets. *Review of Economic Studies*, 44, 369–79.

Stiglitz, J. E. (1969), The effects of income, wealth, and capital gains taxation on risk taking. *Quarterly Journal of Economics*, 83, 263–83.

Stiglitz, J. E. (1983), Some aspects of the taxation of capital gains. *Journal of Public Economics*, 21, 257–94.

Stiglitz, J. E. (1985), The general theory of tax avoidance. *National Tax Journal*, 38, 325–37.

Vickrey, W. (1939), Averaging of income for income tax purposes. *Journal of Political Economy*, 47, 379–97.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Arachi, G. & D'Antoni, M. (2022), Taxation of capital gains upon accrual: is it really more efficient than realisation?. *Fiscal Studies*, 43, 39–61.
https://doi.org/10.1111/1475-5890.12255