Income Taxation, Wealth Effects, and Uncertainty: Portfolio Adjustments with Isoelastic Utility and Discrete Probability

Theodore S. Sims
Boston University School of Law

Follow this and additional works at: https://scholarship.law.bu.edu/faculty_scholarship

Part of the Tax Law Commons

Recommended Citation

This Article is brought to you for free and open access by Scholarly Commons at Boston University School of Law. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarly Commons at Boston University School of Law. For more information, please contact lawlessa@bu.edu.
INCOME TAXATION, WEALTH EFFECTS, AND UNCERTAINTY: PORTFOLIO ADJUSTMENTS WITH ISOELASTIC UTILITY AND DISCRETE PROBABILITY (v.2)

Boston University School of Law Working Paper No. 14-47
(August 7, 2014)

Theodore S. Sims
Boston University School of Law

This paper can be downloaded without charge at:

The expected utility formulation of the problem of a risk-averse agent’s allocating a portfolio between a safe and a risky asset is widely taken as standing for the proposition that if $a^*$ is the optimal holding of the risky asset in the absence of tax, $a^*/(1-t)$ is the optimal allocation in the presence of a proportional tax at rate $t$, a finding obtained on the assumption that the return $r$ to the riskless asset is (or is taxed as though it were) zero. In this paper I model the agent as exhibiting constant relative risk aversion and the probability distribution of the risky asset as binomial, and take $r$ to be greater than zero. With those assumptions the solution $a^*$ depends on the all the parameters of the problem. The key finding of the paper, however, is that the optimal adjustment to taxation does not. But it differs from $a^*/(1-t)$. It depends on $r$, as well as $t$, and reflects in a natural way the response of the agent to the wealth effect of taxation.

Keywords: Taxation and risk; uncertainty; portfolio choice; cash-flow taxation; income taxation.

JEL Classification: D80, G11, H20, H21, H24, H25, K34.
1. Introduction

Domar & Musgrave’s classic (1944) paper explored the impact of a proportional tax on capital income on the optimal allocation of a portfolio between a safe and a risky asset. They concluded that such a tax would typically increase aggregate (social) risk-taking, including that portion assumed by the government through taxation of the returns from and allowance of deductions for losses on risky investments. The subsequent reformulation of the problem using expected utility by Mossin (1968), Stiglitz (1969), and Sandmo (1977), yields the insight that if the risk free rate of return \( r \) (or tax on that return) is zero, then for any risk-averse actor the optimal holding of the risky asset in the presence of tax at rate \( t \) is \( a^*/(1-t) \), where \( a^* \) is the optimal holding of the risky asset in the absence of tax. With that "gross-up" adjustment the agent restores after tax his original (pre-tax) portfolio risk; and, in the absence of offsetting action by the government, social risk-taking is increased. See Sandmo (1985) and Poterba (2002) for surveys.

For Domar & Musgrave (and subsequently Tobin (1958)), the assumption \( r = 0 \) was a natural byproduct of the fact that they took the riskless asset to be cash. Even though that is not generally an innocuous assumption -- in effect it allows the agent costlessly to recover after-tax their original pre-tax holding of the risky asset -- the strikingly simple adjustment it induces has been influential, leading a generation of legal analysts (and more than a few public finance economists) to conclude that an income tax effectively does not (and cannot) tax returns to risk, and so is equivalent to a tax levied on the risk free return to capital;\(^1\) and from that to infer that the

\(^1\) This insight is often and incorrectly attributed to Domar & Musgrave themselves. It appears to have originated with Warren (1980, 1996) with an influential extension by Kaplow (1991, 1994). See Sims (2013) for a critical survey. The foundation for the insight is simple: Suppose that the returns to the risky and riskless asset are \( \bar{R} \) and \( r \), respectively, and that the optimal allocation to the risky asset in the absence of tax is \( \alpha^* \), so that the return to the optimal portfolio in the absence of tax is \( \alpha^* \bar{R} + (1-\alpha^*)r \). If, in the presence of tax at rate \( t \), the allocation to the risky asset is adjusted to \( \alpha^*/(1-t) \), then the after-tax portfolio return is

\[
\frac{\alpha^*}{(1-t)} \bar{R}(1-t) + \left[ 1 - \frac{\alpha^*}{(1-t)} \right] r(1-t) = \alpha^* \bar{R} + (1-\alpha^*)r - rt,
\]

equivalent to the (pre-tax) return to the original portfolio, reduced by tax on the return to the riskless asset as though earned on the entire portfolio. The effect, however, is implicit: what appears as a restoration to the after-tax portfolio of the pre-tax return on (and risk of) the risky asset is accomplished by enlarging the pre-tax holding of that asset. Hence, nominal private pre-tax risk-taking increases, as (through taxation) does social risk-taking, unless and except to the extent that the government engages in offsetting conduct of the sort essential to the conclusions of Kaplow (1991, 1994).
difference between income and cash flow taxation is just tax on the risk free return. While the first of those two propositions requires only that economic agents exhibit risk tolerance that is consistent under the two alternative systems, the second requires that they respond to an income tax by adjusting their portfolio composition exactly as indicated above.²

These insights were obtained, however, only by putting aside the possible wealth effects of capital income taxation on the tolerance for risk. In the expected utility setting, when the assumption \( r = 0 \) is relaxed, the optimal holding of the risky asset is no longer \( a^* / (1-t) \), but is instead defined (see Mossin) by

\[
\frac{\partial a}{\partial t} = \frac{a}{1-t} - \frac{r W}{1 + r(1-t)} \frac{\partial a}{\partial W},
\]

where \( W \) is wealth and \( \partial a / \partial W \) is the wealth derivative of the risky asset. Setting \( r = 0 \) recovers the basic insight of Mossin and Stiglitz; when \( r \neq 0 \) the optimal holding is \( a^* / (1-t) \) only if \( \partial a^* / \partial W = 0 \), that is, the optimal holding is independent of wealth. When neither is 0, however, Mossin and Stiglitz were unable to obtain a solution to (1). They were, however, able indirectly to draw inferences about the properties of \( a^* \), suggesting, consistent with Domar & Musgrave’s original findings, that the presence of a proportional income tax plausibly would lead agents to enlarge their holdings of the risky asset, thereby increasing aggregate social risk, but to something less than \( a^* / (1-t) \).

2. DESIGN

This paper takes a different approach, imposing additional structure on the problem by parameterizing the agent’s preferences as isoelastic (constant relative risk aversion, or CRRA) and the probability distribution as binomial. Although these assumptions are restrictive, the former is consistent with widely held priors about attitudes towards risk. Even with those

² Under a cash-flow tax, which allows investments to be expensed, the tax savings from expensing would enable the agent simultaneously to gross up their holdings of both the risky and the riskless asset, financing the enlargements out of the tax savings from expensing. Then the after-tax portfolio return would be

\[
\frac{a^*}{(1-t)} \bar{R}(1-t) + \frac{(1-a^*)}{(1-t)} r(1-t) = a^* \bar{R} + (1-a^*) r,
\]

so that the entire portfolio is effectively tax exempt, and the apparent difference between income taxation (see footnote 1) and cash-flow taxation is then just \(-rt\). The result is obtained, however, only if in response to income taxation the agent adjusts their optimal holding to \( a^*/(1-t) \).
assumptions, moreover, the optimal holding of the risky asset in the absence of tax depends on all the parameters. What is unexpected and striking, however, is that the adjustments to taxation do not. The departure from a simple gross-up turns out to be independent of both the probability distribution of the risky asset and the parameters of the utility function. With binomial probability the optimal adjustment to taxation by an actor with CRRA preferences depends only on \( r \) and \( t \), in a manner that reflects in a natural way the impact of income taxation on wealth.

3. **Analysis**

To see this, write terminal wealth as

\[
\hat{W} = W_0 (1 + r (1-t)) + a (\hat{X} - r) (1-t),
\]

where \( a \) is the amount invested in the risky asset, \( \hat{X} \) is distributed according to

\[
\hat{X} = \begin{cases} 
X_H > r, & \text{probability } = p \\
-1 < X_L < r, & \text{probability } = 1 - p, \\
E[\hat{X}] > r.
\end{cases}
\]

and utility is isoelastic

\[
U(\hat{W}) = \frac{\hat{W}^{1-\gamma}}{1-\gamma}, \quad (\gamma \neq 1),
\]

so that expected utility is

\[
E \left[ U(\hat{W}) \right] = p \frac{W_H^{1-\gamma}}{1-\gamma} + (1-p) \frac{W_L^{1-\gamma}}{1-\gamma} =
\]

\[
p \left\{ W_0 (1 + r (1-t)) + a (X_H - r) (1-t) \right\}^{1-\gamma} +
\]

\[
(1 - p) \left\{ W_0 (1 + r (1-t)) + a (X_L - r) (1-t) \right\}^{1-\gamma}.
\]

Then the first order condition for expected utility maximization is

\[
\frac{\partial}{\partial a} E \left[ U(\hat{W}) \right] =
\]

\[
p W_H^{-\gamma} \frac{1}{X_H - r} + (1-p) W_L^{-\gamma} \frac{1}{X_L - r} = 0,
\]

3
(after eliminating \((1-t)\) from both terms), or

\[
\frac{W_L^\gamma}{W_H^\gamma} = -\frac{(1-p)(X_L - r)}{p(X_H - r)} \in (0, 1)
\]

since \(E[\tilde{X}] > r\). Writing

\[
\frac{(1-p)(r - X_L)}{p(X_H - r)} = \gamma,
\]

and exponentiating (F2) gives

\[
\frac{W_L}{W_H} = \gamma^{1/\gamma}
\]

and on expanding \(W_L\) and \(W_H\) and solving for \(a^*\) yields

\[
a^* = \frac{W_0[1 + r(1-t)]\{1 - \gamma^{1/\gamma}\}}{\left\{\left[\gamma^{1/\gamma}(X_H - r) - (X_L - r)\right]\right\}(1-t)} = \frac{W_0[1 + r(1-t)]K_N}{\hat{K}_D(1-t)} \bigg|_{t=0}
\]

\[\text{(a*)} \]

\[
-\frac{W_0[1 + r]K_N}{\hat{K}_D},
\]

where in the last two versions the factors in braces involving \(K\) in the numerator and denominator of the initial version have for convenience been denoted \(K_N\) and \(\hat{K}_D\)3.

Since the solution \(a^*\) emerges from a specialization of the problem as formulated by Mossin, we would expect it to satisfy (1). That it does is verified in the appendix. Observe also that as a general matter \(a^*\) depends on all the parameters of the problem, \(r, \gamma, t, W_0,\) and the probability distribution of \(\tilde{X}\). But the key feature of the solution here is that the adjustments of the optimal holdings in response to taxation do not. Note, first, that since \(t\) does not enter into \(K\) (or \(K_N\) or \(\hat{K}_D\)), then for \(r = 0\) (constraining \(X_L < 0\)) the optimal holding \(a^*\) under a proportional tax is just the non-taxable optimal holding divided by \((1-t)\), replicating the findings of Mossin and Stiglitz. Next, the optimal holding in a no-tax world with \(r > 0\) is given by evaluating \(a^*\) at \(t = 0\), and its grossed-up counterpart is simply that divided by \((1-t)\):

\[\text{3 The factor in the denominator has been denoted } \hat{K}_D \text{ to distinguish it from } K_D \text{ in v.1 of this working paper (No. 14-23), where the problem was modelled using as the choice variable the share invested in } \tilde{X}, \text{ as a result of which the solution (there } \alpha^* \text{) took a slightly different form.}\]
From this, on dividing the optimized after-tax holding $a^*$ by the grossed-up optimized pre-tax holding it follows that

\[
\frac{a^* |_{i=0}}{1-t} = \frac{W_0 (1+r) K_N}{\hat{K}_D (1-t)}.
\]

So we have the following:

**THEOREM:** If utility is isoelastic, then for any $\gamma$ and $W_0$ and a risky asset distributed according to a binomial probability satisfying $X \sim$, the ratio of the optimized after-tax holding to the grossed-up optimized pre-tax holding is given by $R_{a^*} < 1$, and depends only on $r$ and $t$.

4. **Implication**

The ratio $R_{a^*}$ is simply the percentage reduction due to taxation of the yield to the riskless asset, and can be viewed as a natural measure of the wealth effect on the portfolio of proportional taxation. It induces a departure from the adjustment that has been widely deployed to characterize the effects of taxation as equivalent (implicitly) to exempting from tax the returns to risk. With the adjustment given by $R_{a^*}$, the after-tax return to the portfolio as optimized in response to taxation (see footnote 1) is given by

\[
a^* \left[ \frac{1 + r (1-t)}{1+r} \right] \hat{R} + \left[ 1 - a^* \left[ \frac{1 + r (1-t)}{1+r} \right] \right] r - r t,
\]

so that the allocation to the risky asset is effectively reduced, and, compared to the pre-tax portfolio, after-tax returns are reduced by more than just a tax as though levied on the riskless return. As such, the conclusion reached here crystallizes the notion that the income tax is (even implicitly) more than just a tax confined to the riskless return. More importantly, it contradicts the stronger claim, premised on the belief that the optimal adjustment to income taxation is to enlarge $a^*$ to $a^*/(1-t)$, that an income tax differs from a cash-flow tax only by taxation of the riskless return.
APPENDIX

In Mossin’s (1968) formulation, the after-tax optimal allocation satisfies the partial differential equation

\[
\frac{\partial a^*}{\partial t} = \frac{a}{1 - t} - \frac{rW}{1 + r(1-t)} \frac{\partial a}{\partial W^*},
\]

so that we should expect the solution \(a^*\) to satisfy (1).

Dividing \(a^*\) by \((1-t)\) yields

\[
(1a) \quad \frac{a^*}{1 - t} = \frac{W_0 [1 + r(1-t)] K_N}{\hat{K}_D (1-t)^2},
\]

Differentiating \(a^*\) with respect to \(W_0\) gives

\[
(1b) \quad \frac{\partial a^*}{\partial W_0} = \frac{[1 + r(1-t)] K_N}{\hat{K}_D (1-t)},
\]

and multiplying that by \(rW_0/[1+r(1-t)]\) produces

\[
(1c) \quad \frac{\partial a^*}{\partial W_0} \frac{rW_0}{1 + r(1-t)} = \frac{rW_0 K_N}{\hat{K}_D (1-t)}.
\]

So the right-hand side of (1) is

\[
\frac{a^*}{1 - t} = \frac{rW_0}{1 + r(1-t)} \frac{\partial a^*}{\partial W_0}
- \frac{W_0 [1 + r(1-t)] K_N}{\hat{K}_D (1-t)^2} - \frac{rW_0 K_N}{\hat{K}_D (1-t)}
- \frac{W_0 K_N}{\hat{K}_D (1-t)^2}.
\]

On the other hand, differentiating \(a^*\) with respect to \(t\) produces

\[
\frac{\partial a^*}{\partial t} = - \frac{W_0 rK_N \hat{K}_D (1-t)}{\hat{K}_D^2 (1-t)^2} - \frac{W_0 [1 + r(1-t)] K_N \hat{K}_D}{\hat{K}_D^2 (1-t)^2}
- \frac{W_0 K_N \hat{K}_D}{\hat{K}_D^2 (1-t)^2} = \frac{W_0 K_N}{\hat{K}_D (1-t)^2}.
\]

Hence, \((a^*)\) satisfies (1).
REFERENCES


