Why are municipal bonds issued at a premium?*

Mattia Landoni
Columbia Business School
mattia.landoni@gmail.com
March 12, 2014

Abstract

After-tax pricing of tax-exempt bonds suggests an explanation for two previously unexplained stylized facts of the 4-trillion-dollar municipal bond market: the fact that many bonds are issued at a premium, and the fact that coupons are “sticky”, i.e. they rise and fall slower than yields do. I propose a simple explanation: because of a quirk of tax law, it is the coupon, not the yield, that is tax-exempt; a high coupon rate is efficient, because it ensures that the bond stays fully tax exempt even if later it trades at a higher yield. This same mechanism also predicts a third, previously undocumented, stylized fact: issue prices of noncallable tax-exempt bonds are increasing in time to maturity. I show that this upward-sloping “term structure of issue prices” exists, using a collection of 200,000 unique securities held by insurance companies.

*PRELIMINARY—PLEASE DO NOT CITE WITHOUT THE AUTHOR’S PERMISSION. COMMENTS ARE WELCOME. Please address correspondence to mattia.landoni@gmail.com. Thanks to John Levin, Charles Jones, Dan Amiram, Andrew Ang, Saki Bigio, Andy Kalotay, and Doron Nissim for more than one useful discussion; to JK Auh, for listening to frequent rants about the importance of taxes for asset pricing; and to Alex Raskolnikov, for encouraging me with his enthusiasm for my early results. There are no mistakes left.
1 Introduction

The U.S. municipal bond market is the world’s largest market for tax-exempt assets. With a total capitalization of about four trillion dollars, it constitutes about 8 percent of the net wealth of U.S. households and nonprofits, and about 2 percent of the world’s financial assets.¹

Few of the municipal bond market’s many quirks have been examined in the finance literature.² A well-established stylized fact is that many municipal bonds are issued at a premium, i.e. with a coupon rate significantly larger than the yield; yet another fact is that coupons are “sticky”: recently, as required returns have fallen to record lows, coupon rates have remained high around 5 percent. To the best of my knowledge, neither one of these facts has been previously considered, let alone explained, in the academic finance literature.

In this paper I propose that after-tax pricing of tax-exempt bonds can account for both these stylized facts. The mechanism is simple: for bonds issued at par, it is the coupon, not the yield, that is tax-exempt; if later the bond is bought in the secondary market at a yield larger than the coupon, the additional market discount yield is taxable. Thus, a high coupon rate is efficient, because it ensures that the bond stays fully tax exempt even if later it trades at a yield higher than the original issue yield, as long as this yield is not larger than the coupon. Of course, a bond issued with a coupon rate in excess of the required yield will fetch a premium price, like the ones observed in reality.

After-tax pricing also yields a novel prediction, which I show to be supported by the data: noncallable bond issue prices are increasing in the maturity of the bond. Once again, the mechanism is simple. Long-term bonds have more volatile prices, and a larger issue price is needed in order to accommodate this greater expected volatility, so that the price does not dip below par (i.e., the yield does not exceed the coupon). Taken all together, these predictions and the corresponding empirical facts constitute strong

¹Total net wealth of U.S. households is about $50 trillion, according to the Federal Reserve Flow of Funds statistics. An estimate by McKinsey Global Institute places the stock of all debt and equity across the globe at $212 trillion in 2010.
²A recent notable exception is Ang et al.’s (2010) examination of the “de minimis” effect, i.e. the sudden jump in yields as bond prices fall below an artificial boundary created by the applicable tax rules.
evidence that taxes are capitalized in asset prices when the marginal investor is taxable.³

A useful byproduct of deriving all these predictions about the issuer’s choice of an optimal coupon rate is the investor’s optimal trading strategy (taking the bond’s coupon as given). Constantinides and Ingersoll (1984) already consider the optimal trading of (taxable and tax-exempt) bonds with taxes. Perhaps because their main interest is in the measurement of the taxable yield curve, they make the apparently innocuous decision to focus on bonds issued at par. After describing several optimal trading rules, they conclude that “no simple characterization of the optimal trading policy is possible,” and that the “tax option” value is negligible for tax-exempt bonds.⁴

The optimal strategy is indeed not simple to characterize, as it involves realizing either losses or gains, depending on all tax rates, transaction costs, time to maturity, and prevailing interest rate. However, in Section 4 I show a form of visual characterization that is complete and human-readable.

More importantly, the value of trading optimally is far from negligible for an investor in tax-exempt bonds. Simple observation of the tax code suggests that focusing on bonds issued at par needlessly constrains the tax trading opportunities available to the investor; and in reality, tax exempt bonds are often issued at a premium. For tax-exempt bonds issued optimally, the tax option value can easily be worth several percentage points of the issue price. In terms of Constantinides and Ingersoll’s original purpose, ignoring the tax option creates a serious bias in the measurement of the tax-exempt yield curve.⁵

However, discussing the magnitude of the tax option value embedded in the price

³Several authors in the past have made valiant attempts at establishing similar evidence in the context of the U.S. bond market, but have encountered several difficulties. It is hard to find evidence that taxes affect the prices of U.S. Treasury bonds, perhaps because of the prevalence of foreign, sovereign and just tax-exempt investors (Litzenberger and Rolfo, 1984; Jordan and Jordan, 1991; Green and Ødegaard, 1997; Elton and Green, 1998). It is also hard to reconcile the relative levels of taxable and tax-exempt yields, for a myriad of reasons including investor clienteles, different liquidity, and different security features such as call options (for an incomplete and impressionistic list, see Trzcinka, 1982; Green, 1993; Chalmers, 1998; Longstaff, 2011; Jordan, 2012).

⁴In the asset pricing literature, “tax option” generally indicates the ex-ante value of optimally harvesting future gains or losses for tax purposes, as compared to a buy-and-hold strategy. Because gains and losses can be harvested only when it’s profitable to do so, tax trading has an option-like payoff, hence the name.

⁵Interestingly, the bias is an underestimate of the true yield; correcting it would further deepen the “muni puzzle”, i.e. the fact that long-term tax-exempt yields appear “too high” compared to their taxable counterparts.
of tax-exempt bonds is neither the purpose, nor the most novel contribution of this paper.\textsuperscript{6} The main positive contribution is to link the existence of tax options to the otherwise puzzling issuance patterns mentioned above. The main normative contribution is to provide valuable policy advice for tax-exempt bond issuers—those who don’t already behave optimally, and those who behave optimally without knowing why. When investors trade optimally and bid competitively, issuing a bond with the “right” coupon can be worth one or two percentage points of the issue price, compared with issuing at par. \textit{This magnitude is likely to exceed the cost of issuance itself,} which is rarely more than one percent of principal, and often much less.

The paper is structured as follows. In Section 2, I describe the taxation of bonds to the minimum extent necessary to understand the economic forces at work. In Section 3, I build a stylized model to gain some intuition: whether there is a unique optimal coupon, and how it is affected by the expected level and the uncertainty in future interest rates. In Section 4, I reproduce the same intuition with much more realism: in order to make the model’s predictions comparable with the data, I use dynamic programming to derive issue prices as a function of coupon rates when investors trade optimally. In Section 5, I document the rather well-known fact that indeed, tax-exempt municipal bonds are issued at significant premia, while corporate bonds, foreign government bonds, and even taxable municipal bonds are not. I also verify the one novel prediction derived in the previous section: noncallable bond issue prices are increasing and concave in maturity. This is striking evidence that even in a market as illiquid and opaque as the municipal market, prices transmit information in a remarkably efficient way. Section 6 concludes.

\section{The taxation of bonds}

This section is meant to familiarize the reader with the U.S. federal income tax rules that are relevant in order to understand the modeling choices of Sections 3 and 4 (through-

\textsuperscript{6}The magnitude of the tax option value is already noted (and independently verified) by Kalotay and Howard (2014), who use a slightly different methodology. In a working paper, Kalotay (2013) also gives a more intuitive description of certain features of the optimal trading strategy.
Figure 1: Two types of bonds, five tax regimes. (See text for an explanation).
out the paper, I ignore state and local taxes for simplicity). While this paper is primarily concerned with tax exempt bonds, discussing the taxation of taxable and tax exempt bonds together entails little extra cost and helps to more clearly highlight the differences that drive all the results in this paper.

Figure 1 is meant as a guide and summary to this section. In each of the six charts, a point represents a transaction executed at the current market price. The horizontal axis represents the time of the sale (0 = time of issue), and the vertical axis represents the sale price (100 = face value). Bonds can be taxable (left) or tax exempt (right), and they are issued at a price equal to face value (par), below (discount) or above (premium).

The solid black line is the deterministic time path of the tax basis for a buy-and-hold investor who buys the bond at the issue price, and receives 100 (the bond’s face value) when the bond matures. Accordingly, the tax basis starts at the issue price (a point on the vertical axis) and gradually converges to 100 at the time of maturity (a point on the horizontal axis). The dashed lines around the tax basis are confidence bounds for the random path of the actual market price; unlike the basis, the market price is not known in advance.

Finally, the coloring of each area indicates what tax regime applies to the transaction. Thicker polka dots represent a greater tax incentive to sell, assuming that the current holder bought the bond at the time of issue. As the Figure shows, the taxation of bonds creates five distinct tax regimes! Moreover, as explained below, more than one regime could apply to a given sale.

2.1 The wash sale approach

In this paper, the tax consequences of trading are studied by focusing on a wash sale, i.e. a transaction in which the owner of the bond sells the bond and buys it back at the same prevailing price. Focusing on a wash sale, as opposed to a “good faith sale”, is without loss of generality. To understand why, consider the following thought experiment. Investor S holds a bond; her tax basis is lower than the current prevailing price,

7The tax basis is the book value of the asset for tax purposes: upon a sale, the seller realizes a capital gain or loss calculated as the difference between the sale price and the tax basis.
so by selling the bond $S$ will realize a taxable gain. A buyer, $B$, comes along and makes an offer:

- $S$ may sell the bond to $B$ directly (a “good faith sale”). A gain will be realized and the bond will change hands.
- Or, $S$ can execute a wash sale (gain will be realized), then sell the bond to $B$ (the bond will change hands with no tax consequences, because the asset’s tax basis is now equal to the market price).

Because of this “separability” of the tax and ownership consequences of a sale, if a wash sale is advantageous for the investor from a tax standpoint, so is any sale; and viceversa. Thus, focusing on a wash sale permits us to examine the tax consequences of trading through a simple partial equilibrium argument.

### 2.2 Two types of bonds, five tax regimes

Income from bonds can be subject to five distinct regimes, which I call “lock-in”, “quasi accrual”, “de minimis”, “market discount” and “tax option”. Under all regimes, the tax bill assessed upon a wash sale is identical. The investor pays capital gains taxes (if realizing a gain) or receives a loss deduction (if realizing a loss). For simplicity, I assume that gains and losses are taxed at the same rate $\tau_G$, so realizing a one-dollar gain causes the investor to face a tax bill of $\tau_G$, while realizing one dollar of loss the investor receives a rebate of $-\tau_G$.

The difference between the regimes is entirely determined by what happens after a wash sale. Tax law is such that every dollar of gain (loss) realized today is balanced by a one-dollar reduction (increase) in the taxable part of income flowing from the asset in the future. Each regime specifies a different combination of (i) timing and (ii) applicable tax rates.

It is useful to classify the regimes depending on their reliance on an “economic” or a “cash” criterion.\(^8\) For instance, take a bond issued at a price of 90, paying no coupon,
and paying off 100 in two years. According to a cash criterion, interest income of 10 is realized in the second year, when the cash is received. According to an economic criterion, the bond yields a constant return of 5.41% per year, and thus the investor will record an income of 4.74 (5.41% of 90) in the first year, and 5.26 (5.41% of 94.74) in the second year. This is meant to more closely reflect the underlying economics of the situation.

**Quasi accrual (economic)** This regime applies to taxable bonds only—to all gains, and to any losses above face value (as in the top-left panel of Figure 1). Suppose an investor holds a bond with a tax basis of 120; the market price is 110. By executing a wash sale, the investor posts an immediate taxable loss of 10, and future taxable income from the bond is increased by 10. This extra income is spread over the remaining life of the bond as if it were extra yield. Under this regime it can be advantageous to realize gains or losses, depending on the specific rates faced by the investor, and the time remaining to maturity; but the advantage or disadvantage will be small.

**De minimis (cash)** This regime applies to all bonds whose current price is just below face value (for bonds issued at premium or par) or just below the original tax basis (for bonds issued at discount). Suppose the bond price is now 98.5, while the tax basis 120 as before. The investor posts a taxable loss of 21.5 (120-98.5). Of this, the first 20 (120-100) falls into the quasi accrual area so the investor will have 20 of extra ordinary income spread over the remaining life of the bond. The remaining 1.5 (100-98.5) falls into the de minimis area, so the investor will realize 1.5 of capital gain, but only when she disposes of the bond, or the bond matures. Realizing de minimis losses with a wash sale creates a taxable loss today in exchange for a taxable gain in the future. Because losses and gains are assumed to be taxed at the same rate, realizing de minimis losses is always advantageous.

---

9The de minimis treatment applies to bonds whose discount to face value is no greater than 0.25 times the number of complete years between the bond’s acquisition date and its maturity date. For instance, a price of 98.5 qualifies for de minimis treatment only if the bond has at least six whole years left to maturity: 

\[(100 - 98.5)/0.25 = 6\].
**Market discount (cash)** This regime applies to all bonds whose current price is anywhere below the *de minimis* threshold. Suppose the bond price is now 90, and the tax basis 120 as always. The investor posts a taxable loss of 30 (120-90). As before, the first 20 (120-100) falls into the quasi accrual area so the investor will have 20 of extra ordinary income spread over the remaining life of the bond. The remaining 10 (100-90) falls into the market discount area: when the bond matures or is disposed of, the investor will realize 10 of *ordinary income*. Thus, a wash sale creates a taxable loss today in exchange for ordinary income in the future. Compared to the *de minimis* regime, income is taxed at the same time but at a higher tax rate, and thus the market discount regime is *less* advantageous; compared to the quasi accrual regime, income is taxed at the same tax rate but later, and thus the market discount regime is *more* advantageous. Realizing market discount losses may or may not be advantageous.

The next two regimes apply only to tax exempt bonds. Together, they are the counterpart of the “quasi accrual” regime of taxable bonds.

**Lock-in (economic)** This regime applies to all gains on tax-exempt bonds. Suppose an investor holds a bond with a tax basis of 120; the market price is 130. Upon a wash sale, the investor realized 10 of capital gains. As in the quasi accrual regime, future interest income is reduced by 10. However, interest income is already tax-exempt! The economic benefit from reducing future tax-exempt interest income is either null or very small, depending on the applicable rules, while the present cost of paying the capital gains tax is large and immediate. This regime makes selling very costly, and creates a so-called “lock-in effect”.

**Tax option (economic)** This regime applies to tax-exempt bonds, and only to losses above face value: therefore, it appears only in the top-right panel of Figure 1. It is the opposite of “lock-in”: realizing losses creates an immediate tax benefit, and increases the amount of future tax-exempt interest income, which does not entail a significant cost. As in the market discount example, suppose the bond price is 90 and the tax basis 120. The investor posts a tax loss of 30 (120-90). Future tax-exempt
income increases by 20 (120-100), spread over the remaining life of the bond as extra tax-exempt yield. The remaining 10 (100-90) falls into the market discount area, so the investor will realize 10 of ordinary income when she disposes of the bond or the bond matures.10

3 Optimal security design for tax-exempt bonds

The “quasi accrual” tax regime of taxable bonds described in Section 2 creates little opportunities for profitable tax trading because the tax rules try to mirror the underlying economics of the situation: income is recorded approximately when it accrues to the investor. If the taxation of tax-exempt bonds were equally consistent with the “accrual” benchmark, capital gains and losses on tax-exempt bonds would also be treated as tax-exempt. It is the stark departure from this benchmark that creates opportunities for tax trading.

Figure 1 highlights an important asymmetry: while a strong disincentive to realize gains exists for all tax-exempt bonds, a valuable tax option to realize losses exists only for bonds issued at a premium. Constantinides and Ingersoll (1984) find that the value of tax trading is not a meaningful part of the market price of tax-exempt bonds, because they restrict themselves to examining bonds issued at par. Allowing for premium issuance turns the result on its head: now the tax option value in tax-exempt bonds is much larger than that found in taxable bonds.

In this Section, I show that this “inconsistency” of the tax code has important implications for the way securities are designed. An issuer of tax-exempt bonds (a state, local government or nonprofit institution) designs the bond in a way that maximizes the investor’s future ability to generate valuable tax losses for federal tax purposes. A competitive investor pays the issuer for this “service”, and in equilibrium this payment translates to an implicit fiscal subsidy from the national government to the subnational

---

10 Generally, in the finance literature, every opportunity for profitable tax trading is called a “tax option”. In the case of tax-exempt bonds, however, the value of realizing losses above par dwarfs all other opportunities for profitable tax arbitrage. Thus, for simplicity, throughout the paper “tax option” refers only to this option to realize losses above par.
Beside the obvious conclusion that issuing premium bonds is optimal, the two-period model in this Section reveals not-so-obvious insights. First, as shown in Figure 2, often there is a unique optimal coupon (or issue price). A “too high” coupon reduces the bond’s duration, dampening price volatility and ultimately hurting the value of the investor’s tax option. Second, the optimal issue price (as a fraction of face value) should be increasing in the expected volatility of the bonds and the direction of future interest rate movements. These qualitative predictions are useful to build economic intuition. In Section 4, I use dynamic programming to calibrate quantitative predictions that can be compared with the actual features of real-world bonds.
3.1 Definitions and assumptions

There are two time periods \((t \in \{0, 1, 2\})\) and two agents (an issuer and an investor). The agents are risk-neutral and competitive, and there are no transaction costs. At time 0, the issuer issues a two-period bond with face value 1. In this simplified setup, I assume that the only security design choice available to the issuer is the coupon, \(c\).\(^{11}\)

**Time value of money** Both agents discount future payments at the same rate. One dollar received at time \(t + 1\) is worth \(\delta_t\) at time \(t\). I assume \(\delta_t < 1\), i.e. the discount rate is positive. The discount factor is random: while \(\delta_0\) is known at time 0, \(\delta_1\) becomes known only at time 1:

\[
\delta_1 = \begin{cases} 
\bar{\delta} + \sigma & \text{with probability } p \\
\bar{\delta} - \sigma & \text{with probability } 1 - p
\end{cases}
\]

where \(\sigma\) can be interpreted as the volatility of the discount factor. Thus, the bond price at time 1 will be

\[
P_1 = \delta_1 (1 + c) = \begin{cases} 
P_{\text{high}} \equiv (1 + c) (\bar{\delta} + \sigma) & \text{with probability } p \\
\bar{P}_{\text{low}} \equiv (1 + c) (\bar{\delta} - \sigma) & \text{with probability } 1 - p
\end{cases}
\]

**Taxes** The issuer is tax exempt, and therefore the investor will not pay any tax on the coupon income. However, the investor does pay capital gains taxes. If at time 1 the investor sells the bond for a price above its book value (“tax basis”), she will realize a gain and pay tax at a rate \(\tau_G\). If the investor realizes a loss, conversely, she will receive a subsidy at the same rate \(\tau_G\). However, the subsidy applies only to the part of loss that is above the bond’s face value.\(^{12}\)

Define \(T_1\) as the time-1 tax:

\(^{11}\)Note that the *yield* is determined by market forces, but the *coupon* is the issuer’s choice; if the coupon is higher than the yield, the bond will be issued at a premium to face value, and vice versa.

\(^{12}\)Based on the tax rules reviewed in Section 2, the value of realizing losses above par is much larger than the value of tax trading under the quasi accrual, de minimis, and market discount regimes. To keep the math as simple as possible, I assume that the value of tax trading under all these other regimes is exactly zero.
\[ T_1 = -\tau_G \max(1, P_1) - B_1 \]  

where \( P_1 \) is the sale price at time 1, and \( B_1 \) is the tax basis at time 1. (A negative number means the investor pays a positive tax). For simplicity, \( B_1 \) is determined using straight line amortization:

\[ B_1 = 1 + \frac{P_0 - 1}{2} = \frac{P_0 + 1}{2} \]  

Also for simplicity, I assume that \( \sigma \) is large enough, so that \( P_{\text{low}} < B_1 \) and \( P_{\text{high}} > B_1 \). Finally, (2) contains the implicit assumption that premium bonds are optimal \( (P_0 \geq 1) \). Later this assumption will turn out to be superfluous.

To clarify, consider the following three examples. Assume that the bond’s face value is 100, but the bond was issued at a premium price of 110 at time 0. At time 1, the tax basis is \((110 + 100)/2 = 105\).

- If the bond is sold at 109, the gain is fully taxable and the investor pays a tax of \((109 - 105)\tau_G = 4\tau_G\).

- If the bond is sold at 101, the loss is fully taxable and the investor receives a subsidy of \((105 - 101)\tau_G = 4\tau_G\).

- If the bond is sold at 98, only part of the loss is taxable, because the price is now below face value. The investor receives a subsidy of \((105 - 100)\tau_G = 5\tau_G\) (instead of \((105 - 98)\tau_G = 7\tau_G\), if the loss were fully taxable).

This unusual set of rules is a stylized version of the actual U.S. federal tax code. Under the actual rules, the seller always get to deduct the full loss, as discussed in the previous Section. However, in the third example, the two dollars of discount to face value are called “market discount interest” and they will be taxable to the buyer as interest income sometime in the future. Because the buyer is likely a taxable entity, she will capitalize the future taxes into the price that she is willing to offer to the seller. This will make the price lower, negating part of the benefit of realizing losses for the seller.
3.2 The investor’s problem

At time 1, when \( P_1 = P_{\text{high}} \) (with probability \( p \)), the investor may be forced to sell the bond with an exogenously given probability \( \lambda \). Moreover, when \( P_1 = P_{\text{low}} \) (with probability \( 1-p \)), the investor will sell and buy back the bond, because realizing losses is advantageous. The expected tax at time 1 is then

\[
\mathbb{E}[T_1] = \tau_g \left[ p \lambda (B_1 - P_{\text{high}}) + (1-p) (B_1 - \max(1, P_{\text{low}})) \right]
\]  

(4)

And thus the investor will offer the following price for the bond:

\[
P_0 = \delta_0 \left( \bar{\delta} (1 + c) + c + \mathbb{E}[T_1] \right)
\]

Use (3) and (4) to replace \( \mathbb{E}[T_1] \) and solve for \( P_0 \):

\[
P_0 = \delta_0 \frac{\bar{\delta} (1 + c) + c + \tau_g L(c)}{1 - \delta_0 \tau_g \frac{\rho + 1 - p}{2}}
\]  

(5)

with

\[
L(c) = \frac{p \lambda + 1 - p}{2} - p \lambda P_{\text{high}} - (1-p) \max(1, P_{\text{low}}).
\]

3.3 The issuer’s problem

The issuer is a state or local government, while the investor pays only Federal taxes. Thus, the issuer wants to design the security (i.e., choose \( c \)) in a way that minimizes the present value of expected taxes paid by the investor over the life of the bond. Because in practice the tax paid at time 1 will often turn out to be negative, I will frame the problem in terms of maximizing the tax subsidy. It then becomes clear that what matters is not the dollar tax \( \mathbb{E}[T_1] \), but rather the subsidy/principal ratio \( W \):

\[
\max_c W \equiv \frac{\delta_0 \mathbb{E}[T_1]}{P_0}
\]
In other words, for every dollar of capital raised by issuing the bond, a fraction $W$ comes from a federal tax subsidy, and does not entail any future cash outflows for the issuer.

Using again (3) and (4), substitute and simplify:

$$W = \delta_0 \tau G \left[ \frac{p\lambda + 1 - p}{2} + p\lambda \left( \frac{1}{2} - P_{\text{high}} \right) + \left( 1 - p \right) \left( \frac{1}{2} - \max \left( 1, P_{\text{low}} \right) \right) \right]$$

(6)

Recognize that the numerator of the second term inside the brackets is just $L(c)$:

$$W = \delta_0 \tau G \left[ \frac{p\lambda + 1 - p}{2} + L(c) \right]$$

Use the definition of $P_0$ and simplify:

$$W = K + (1 - K) \frac{L(c)}{\delta (1 + c) + \tau G L(c)}$$

$$K \equiv \delta_0 \tau G \frac{p\lambda + 1 - p}{2} \in [0, 1]$$

Because $K$ and $1 - K$ are positive constants,

$$\frac{\partial W}{\partial c} > 0 \iff \frac{\partial}{\partial c} \frac{L(c)}{\delta (1 + c) + \tau G L(c)} > 0$$

which simplifies to

$$\left( c + \frac{\delta}{1 + \delta} \right) L'(c) > L(c) .$$

(7)

$L(c)$ contains a max operator, and thus it needs to be examined case by case.

3.3.1 Case I: $P_{\text{low}} < 1$

Intuitively, $P_{\text{low}} < 1$ should not be optimal, because only a part of the losses is tax deductible. This is easily verified by showing that the tax subsidy ratio is increasing in the coupon rate $c$. This would also imply that some premium is optimal, since par issuance ($P_0 = 1$) is a subcase of this case.

The definition of $L(c)$ in this case is
\[ L(c) = \frac{p\lambda + 1 - p}{2} - p\lambda(\delta + \sigma)(1 + c) - (1 - p) \]

\[ L'(c) = -p\lambda(\delta + \sigma) \]

Substituting the definitions of \( L, L' \) into (7) and simplifying, we obtain

\[ \lambda \frac{p}{1 - p} \frac{1 - \delta - 2\sigma}{1 + \delta} < 1 \]  

(8)

If \( 1 - \delta - 2\sigma < 0 \), \( W \) is always increasing in \( c \). Otherwise, \( W \) is increasing in \( c \) under a very weak condition:

\[ \lambda \frac{p}{1 - p} < \frac{1 + \delta}{1 - \delta - 2\sigma} \]  

(9)

For instance, with \( p = \frac{1}{2} \), (9) requires that \( \lambda < \frac{1 + \delta}{1 - \delta - 2\sigma} \), which is always true because the right-hand side is greater than 1.

The result that \( W \) is almost always increasing in \( c \) makes things easier. So far we have assumed that issuing at a premium is optimal \((P^*_0 \geq 1)\). A consequence of this result is that the assumption is essentially not needed, because

\[ \frac{\partial W}{\partial c} \bigg|_{P_0=1} > 0 \]

and therefore issuing at par is suboptimal, and we can ignore all issue prices below par.  

3.3.2 Case II: \( P_{\text{low}} \geq 1 \)

After showing that some premium is optimal, the remaining question is whether there is a unique, well-defined optimal coupon. If

\[ \frac{\partial W}{\partial c} \bigg|_{P_{\text{low}}=1} < 0 \]

13Showing rigorously that below-par issuance is suboptimal would require specifying the OID tax rules mentioned in Section 2, which would make the mathematics more complicated without adding any extra insight; especially given that the optimality of premium issuance should be obvious by just looking at Figure 1.
then the optimal coupon is the one that makes $P_{\text{low}} = 1$ exactly. In the opposite case, there is no unique optimum, and an infinite coupon is optimal.

An infinite coupon would be no problem in practice: all the arguments in this section are scale-invariant, because the issuer maximizes the fraction of issued principal that is a tax subsidy. For $c \to \infty$, this fraction is a well-defined number. An $n$-period bond with an infinite coupon is just like a mortgage with $n$ payments, up to a scaling constant. Nonetheless, a closed-form, finite expression for the optimal coupon is desirable, because it makes it possible to make additional testable predictions.

The definition of $L(c)$ in this case is

$$L(c) = \frac{p\lambda + 1 - p}{2} - p\lambda \left(\bar{\delta} + \sigma\right) (1 + c) - (1 - p) (\bar{\delta} - \sigma) (1 + c)$$

$$L'(c) = -p\lambda (\bar{\delta} + \sigma) - (1 - p) (\bar{\delta} - \sigma)$$

Substituting the definitions of $L, L'$ into (7) and simplifying, we obtain

$$\frac{p\lambda}{1 - p} \frac{1 - \bar{\delta} - 2\sigma}{1 + \bar{\delta}} > \frac{1 - \bar{\delta} + 2\sigma}{1 + \bar{\delta}}$$

(10)

If $1 - \bar{\delta} - 2\sigma > 0$ $W$ is never increasing in $c$, because the right-hand side expression is positive, but the left-hand side is negative, and the condition is not verified. If $1 - \bar{\delta} - 2\sigma < 0$, however, $W$ is increasing in $c$ under a hard-to-verify condition:

$$\frac{\lambda - p}{1 - p} > \frac{1 - \bar{\delta} + 2\sigma}{\bar{\delta} + 2\sigma - 1}$$

(11)

However, this condition may occasionally hold. Thus, a situation where an infinite coupon is optimal is possible, as well as one in which there is a well-defined unique optimum.\(^{14}\)

\(^{14}\)First, $1 - \bar{\delta} - 2\sigma > 0$ is likely to be verified. Rewrite the condition $1 - \bar{\delta} - 2\sigma > 0$ as $1 - (\bar{\delta} + \sigma) > \sigma$. Note that $\bar{\delta} + \sigma < 1$ by assumption, because it is a discount factor. In terms of the discount rate, this means that the expected discount rate at time 1 is at least two standard deviations above zero. Negative discount rates are possible in theory and negative interest rates have been observed in practice. However, both in theory and practice this is a rare occurrence; hence, for most realistic calibrations, this condition will be verified.

Second, if it’s not verified, it is easy to show that the right-hand side expression is larger than 4. Because $\lambda$ on the left-hand side is less than one, one should choose a very high $p$ to make this condition verified.
3.4 Solution and comparative statics

When it exists (i.e. most of the time), the optimal coupon is found by setting $P_{\text{low}} = 1$, a “kink” solution like the one pictured in Figure (2).

\[ P_{\text{low}} = (\bar{\delta} - \sigma) (1 + c^*) = 1 \]

\[ c^* = \frac{1 - \bar{\delta} + \sigma}{\delta - \sigma} \tag{12} \]

The intuition is simple: the price should never go below par. If $P_{\text{low}} < 1$, when there are losses, some losses will not be deductible. Thus, increasing $c$ is optimal because increasing $c$ increases $P_{\text{low}}$ and with it, the amount of losses that can be deducted. When $P_{\text{low}} > 1$, all losses are already deductible. As with all bonds, increasing the coupon reduces the duration of the bond because a larger share of the bond’s cash flow is received at time 1 when the first coupon is paid. Thus, price volatility decreases, and with it decreases the value of the option to realize losses. The optimal price is

\[ P^*_0 = \frac{1 - \tau_G \left( \bar{\delta} \frac{p^{\lambda+1-p}}{2} \right) + \sigma \left( 1 - \tau_G \frac{(1-p)-3p\lambda}{2} \right)}{\left( \bar{\delta} - \sigma \right) \left( \delta_0^{-1} - \tau_G \frac{p^{\lambda+1-p}}{2} \right)} \tag{13} \]

Finally, rewrite by replacing the next period expected discount factor ($\bar{\delta}$) with the current discount factor plus expected change ($\delta_0 + \Delta$):

\[ P^*_0 = \frac{1 - \tau_G \left( \delta_0 + \Delta \right) \frac{p^{\lambda+1-p}}{2} + \sigma \left( 1 - \tau_G \frac{(1-p)-3p\lambda}{2} \right)}{\left( \delta_0 + \Delta - \sigma \right) \left( \delta_0^{-1} - \tau_G \frac{p^{\lambda+1-p}}{2} \right)} \tag{14} \]

This expression will facilitate the interpretation of the dynamic programming results in the next Section.

4 Dynamic programming

Equation (14) of Section 3 implies that the optimal issue price $P^*_0$ is:
• Greater than one. A large coupon (i.e., a premium issue price) is optimal.

• Increasing in price volatility $\sigma$. For instance, longer term bonds have a more volatile price. To accommodate this volatility and prevent the price from ever going below par, a higher issue price is needed.\footnote{Note that a higher issue price does not require a higher coupon rate. For instance, a 5\% coupon with a 3\% yield will result in a higher price for a 10-year bond compared to a 5-year bond, because the same excess yield (5\% - 3\% = 2\%) is paid to the investor for a longer time.}

• Decreasing in $\Delta$, the expected change in the discount factor. If the price is expected to fall, a higher price is needed everything else equals, in order to prevent the price from falling below par.

These qualitative predictions are useful to build some economic intuition, but they are not precise enough to be meaningfully compared with data. In order to make more quantitatively precise predictions, in this Section I price bonds of arbitrary maturity using backward induction to characterize the exact optimal trading strategy for a taxable investor. I use a Markov process with 21 states for the one-period interest rate, and I model transaction costs and a more “realistic” tax code (described below). Given the coupon $c$ and the optimal trading strategy, it is easy to solve for the exact market price of the bond at any time $t$ prior to maturity $T$ as a function of the current state $r_t$. This market price is also the issue price of a bond with maturity $M = T - t$. Finally, assuming that the issuer and the investor agree on the discount rate process, I can calculate the subsidy ratio $W$ and solve for the optimal coupon for each maturity $M$.

4.1 The interest rate process

Similar to Constantinides and Ingersoll (1984), I assume that the long-run distribution of the interest rate is uniform. In particular, the interest rate $r$ assumes one of $N_r = 21$ discrete values ($r \in \{0\%, \ 0.5\%, \ldots, r_{\text{max}} = 10\%\}$). Time is divided into discrete one-year ticks. Each year, the interest rate has a positive probability to stay at the same level, or to jump up or down by up to two percentage points (e.g., if $r_t = 5\%$, $r_{t+1} \in$
Why are municipal bonds issued at a premium? \{3\%, 3.5\%, \ldots, 6.5\%, 7\%\}. The probability of switching is described by the following table.\footnote{When the current rate is less than two percentage points away from the reflecting boundaries of 0\% and 10\%, an adjustment is needed because some events in the table are impossible. For instance, if the interest rate is 8.5\%, a +2\% change would bring the interest rate to 10.5\%, which is higher than the upper bound. In this case, the vector is simply truncated and rescaled to sum to one. This adjustment leaves the long-run distribution uniform.}

<table>
<thead>
<tr>
<th>Change</th>
<th>-2%</th>
<th>-1%</th>
<th>-1/2%</th>
<th>0</th>
<th>+1/2%</th>
<th>+1%</th>
<th>+1%</th>
<th>+2%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr (Change)</td>
<td>.05</td>
<td>.07</td>
<td>.11</td>
<td>.15</td>
<td>.24</td>
<td>.15</td>
<td>.11</td>
<td>.07</td>
</tr>
</tbody>
</table>

This simple calibration is meant to roughly mimic the dynamic of the Fed Funds rate without requiring a lengthy explanation. A more sophisticated calibration would, for instance, try to mimic the short-term tax exempt rate that actually prevails on the U.S. municipal market. Moreover, a yield curve model that yields on average an upward sloping curve would produce more realistic results. However, I do not believe that adding any of these features would change the result substantially.

4.2 Security features, trading, taxes, transaction costs, and investor types

I make several assumptions for simplicity. First, bonds are assumed to be noncallable and pay a coupon only once a year. While most municipal bonds are in fact callable, the presence of a call option makes things more complicated, but does not change the main intuition. For instance, premium bonds are still optimal, and the premium is still increasing in price volatility.\footnote{However, both volatility and the tax basis are likely to be affected by the time to first call, as opposed to the time to maturity—but everything depends on the specific details of the call feature.}

Second, investors can trade the bonds once a year, immediately after the coupon has been paid. As in the simple model of Section 3, there is an exogenous probability \( \lambda \) that the investor must trade; however, if realizing gains or losses is optimal, the investor will trade with certainty.

Third, the tax code is still a simplified version of the actual U.S. federal income tax code, but this time it is much more “realistic” than in Section 3, and I argue that it essentially captures the economics of the actual rules. Unlike in Section 3, market discount income is explicitly taxed as ordinary income. Ordinary income, capital gains/losses,
<table>
<thead>
<tr>
<th>Investor Type</th>
<th>$\tau$</th>
<th>$\tau_G$</th>
<th>$\tau_E$</th>
<th>$\lambda$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individuals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benchmark</td>
<td>40%</td>
<td>20%</td>
<td>0%</td>
<td>0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Frequent Trading, Low Cost</td>
<td>40%</td>
<td>20%</td>
<td>0%</td>
<td>5%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Rare Trading, Low Cost</td>
<td>40%</td>
<td>20%</td>
<td>0%</td>
<td>1%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Rare Trading, High Cost</td>
<td>40%</td>
<td>20%</td>
<td>0%</td>
<td>1%</td>
<td>1.0%</td>
</tr>
<tr>
<td><strong>Nonlife Insurance Companies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benchmark</td>
<td>35%</td>
<td>35%</td>
<td>5.25%</td>
<td>0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Rare Trading, Low Cost</td>
<td>35%</td>
<td>35%</td>
<td>5.25%</td>
<td>5%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Medium Trading, Low Cost</td>
<td>35%</td>
<td>35%</td>
<td>5.25%</td>
<td>10%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Frequent Trading, Low Cost</td>
<td>35%</td>
<td>35%</td>
<td>5.25%</td>
<td>15%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Medium Trading, High Cost</td>
<td>35%</td>
<td>35%</td>
<td>5.25%</td>
<td>10%</td>
<td>0.4%</td>
</tr>
</tbody>
</table>

Table 1: Different investor types face different tax codes ($\tau$, $\tau_G$, $\tau_E$), propensities to trade $\lambda$, and costs of trading $K$. Note that the tax rate on tax-exempt income $\tau_E$ is nonzero for some investors. 

and tax-exempt income are taxed respectively at tax rates $\tau$, $\tau_G$, and $\tau_E$. More details and simplifying assumptions are discussed in detail in Appendix A for the interested reader.

Finally, trading may or may not be cost-free. Depending on several attributes of the sale, the seller, and the bond, the transaction cost $K$ could be as high as a whole percentage point of the bond’s market value. On the other hand, transaction costs may not matter at all: if an investor is deciding which of two bonds to sell, incurring a transaction cost is inevitable, and it is not part of the decision. Similarly, if the bid-ask spread is due mainly to inventory risk and asymmetric information, the cost of trading for tax purposes (e.g. selling a bond to realize a gain and buying it back immediately) could be very low, because tax trading does not create any of these burdens for the dealer.

For all these reasons, different investors may face different tax codes ($\tau$, $\tau_G$, $\tau_E$), propensities to trade $\lambda$, and costs of trading $K$. In order to study the sensitivity of the results to several assumptions, I solve the dynamic programming problem for nine different investor types. Table 1 lists the types. In the rest of the paper, however, for brevity’s sake, I will focus only on the four cases that I deem to be most representative of actual price-setting investors in tax-exempt bonds: the benchmark case ($\lambda = 0$ and $K = 0$ for both individuals and nonlife insurance companies), and a case with moderate trading.

---

18 Note that $\tau_E$ can be nonzero for certain investors, as shown in Table 1.
and transaction costs ($\lambda = 1\%$ and $K = 0.5\%$ for individuals, and $\lambda = 5\%$ and $K = 0.2\%$ for nonlife insurance companies).

### 4.3 The investor’s problem, or pricing the bond

The bond matures at time $T$. At time $t < T$, a risk-neutral investor maximizes the present value of after-tax final wealth. The investor’s utility is

$$V_t(c, T - t, B_t, r_t),$$

a function of the constant coupon rate $c$ and three time-varying state variables: the short rate $r_t$, the time to maturity $T - t$, and the investor’s tax basis $B_t$.

At every period, the investor can choose to do nothing (“hold”), or to realize the position’s unrealized gain or loss by selling the bond and buying back (“sell”). If the investor decides to hold, however, the bond will be sold (and the gain or loss realized) with a probability $\lambda$. This yields the following definition for the value function:

$$V_t(c, T - t, B_t, r_t) = \max_{i \in \{\text{sell, hold}\}} \left[ \frac{c + \mathbb{E}_t[V_{t+1}^i]}{1 + r_t} - \tau_G (P_t - B_t) 1_{(i = \text{sell})} \right]$$

(15)

where $P_t$ is the market price of the bond, $c$ is the coupon, $\mathbb{E}_t[V_{t+1}^i]$ the next-period residual value, and $1_{(i = \text{sell})}$ is an indicator function that is 1 if the investor decides to sell and 0 otherwise. Moreover,

$$V_{t+1}^{\text{sell}} = V_{t+1} (c, T - t + 1, B (P_t, c, T - t), r_{t+1})$$

(16)

$$V_{t+1}^{\text{hold}} = \lambda V_{t+1}^{\text{sell}} + (1 - \lambda) V_{t+1} (c, T - t + 1, B (B_t, c, T - t), r_{t+1})$$

(17)

and finally

$$B (x, c, T - t)$$

is the constant yield method function that gives the next-periods tax basis given the current basis $x$, the coupon $c$, and the time to maturity $T - t$. Because of the need to
solve numerically for the yield, this function cannot be expressed analytically, but it is altogether very simple. Note that in \( V_{t+1} \) the first argument to \( B(\cdot) \) is \( P_t \), because selling the bond resets the basis equal to the current price.

So far we have taken the existence of a market price \( P_t \) for granted. However, \( V_t \) depends on \( P_t \); before we solve for the optimal policy (hold or sell) at the current period, we need to calculate the current period market price. The price is defined recursively as a special case of the value function:

\[
P_t(c, T - t, r_t) = V_t(c, T - t, P_t, r_t).
\]

(18)

Note that \( P_t \) appears on both sides! If the investor buys the bond now, the basis now will be equal to the price. Of course, the value of the position depends on the tax basis, so the price is a function of itself. As show in Appendix B, this presents no practical problem even if no analytical solution exists, because \( V(\cdot) \) is a contraction map with respect to its second argument. Any guess for \( P_t \) converges in a few iterations to the unique fixed point, and thus the price \( P_t \) can be recovered by just knowing the interest rate \( r_t \), the time to maturity \( T - t \), and the value function at the next period \( V_{t+1} \). Once obtained \( P_t \), it is easy to calculate \( V_t \) for all other values of the tax basis \( B_t \) using (15).

### 4.4 The issuer’s problem

The issuer’s problem is simple. \( P_t(c, T - t, r_t) \) as defined in the previous subsection is the issue price of a bond maturing in \( T - t \) years with coupon \( c \) when the current price is \( r_t \). Define \( \hat{P}_t \) as the cost to the issuer of all the future cash flows from the bond:

\[
\hat{P}_t(c, T - t, r_t) = \mathbb{E}_t \left[ \frac{c}{1 + r_t} + \frac{c}{(1 + r_t)(1 + r_{t+1})} + \cdots + \frac{1 + c}{(1 + r_t) \cdots (1 + r_{T-1})} \right]
\]

Like in Section 3, the issuer wants to maximize the subsidy ratio \( W \), i.e. the share of issued amount that is not compensation for future cash outflows:

\[
\max_{c \geq 0} W = \frac{P_t - \hat{P}_t}{P_t}
\]
The difference now is that $W$ also depends on the cost of trading $K$ (which the issuer won’t have to bear directly) and all three tax rates $(\tau, \tau_G, \tau_E)$.

The optimal coupon $c^*$ is simply found by solving the investor’s problem for a set values of $c$ (.01, .02, .03, .04, .045, .05, .055, .06, .065, .07, .075, .08, .085, .09, .095, .10, .11, .15, .2) and picking the one that gives the highest $W$. This rough “grid search” method yields a less-precise optimum compared to a hill-climbing or annealing algorithm. However, given the nature of the problem, one- or two-digit precision is good enough. Moreover, the intuition from Section 3 suggests that a finite optimum may not exist sometimes, at least the way the problem is defined. In these cases, “grid search” will simply yield $c^* = 20\%$, the maximum value available. In practice, a finite optimal coupon below 20% exists very often.

**4.5 Results**

In this subsection, I use the dynamic programming results to describe several interesting aspects of the problem.

**4.5.1 Qualitative predictions**

The properties of optimal issue prices are comparable with the qualitative predictions of Section 3.

**Premium is optimal**  Figure 3 plots $W$ (in percentage points) over the set of $c$ values enumerated in Subsection 4.4 above (except for those values of $c$ that are known to be suboptimal because the issue price is below par). The leftmost point of each curve represents par or near-par issuance. The marginal value of increasing the coupon is very high at first (a very steep hill), as more and more future potential losses become tax-deductible. At the optimum, most future losses have become tax deductible. From now on, in most cases, the marginal value of increasing the coupon is negative: increasing the coupon lowers the duration of the bond, decreasing price volatility, and with it, the value of the investor’s tax option. (All this is in relative terms: both dollar price
Figure 3: Value of the subsidy ratio $W$ (the value of tax options as a fraction of total amount issued) for a 10-year noncallable tax-exempt bond when the marginal issue buyer is an individual (left) or a nonlife insurance company (right). The leftmost point of each curve represents par or near-par issuance. The value increases rapidly at first, as more and more future potential losses become tax-deductible; then it tapers off (in most cases), as nearly all losses are already tax deductible, but further increasing the coupon decreases price volatility and hence option value.

and dollar option value are increasing in the coupon, but after a while price increases faster than option value because of the effect described). It is worth noting that the optimal coupon rate is often below 10%, the maximum allowed interest rate; therefore, the existence of a unique optimum does not depend on the mechanical fact that the interest rate is bounded above for computational purposes.

**Coupons are “sticky”**  This is an interesting case of the prediction from Section 3 that the optimal coupon is increasing in the expected change in yields. The interest rate process assumed for this Section’s dynamic programming exercise exhibits a degree of mean reversion, like most widely used interest rate processes. Thus, when the interest rate is low, one expects it to rise, i.e., one expects the bond to have more losses than average, and a higher issue price (a higher coupon) is needed to make sure all these losses are deductible. The reverse happens when interest rates are high. The net result is that even when interest rates fall to a very low level (e.g., 0%), the optimal coupon stays high (e.g., 5%). Figure (4) shows how “stickiness” varies with maturity and the level of interest rates. For the high-interest rate scenario ($r_0 = 8\%$) for insurance companies
Figure 4: Optimal coupon as a function of bond maturity. For individuals, the optimal coupon is a finite number in every scenario and for every maturity. For nonlife insurers, however, the optimal coupon is always larger than 20%. As the bond maturity increases, the sensitivity of the optimal coupon to the current interest rate drops considerably, i.e. coupons are “sticky”. The presence of “steps” in the graph is an artifact of the coarse grid search optimization described in Subsection 4.4. (right), the optimal coupon is higher than 20%—and possibly infinite—and therefore no stickiness exists.

**Term structure of issue prices** Figure (5) displays the “term structure of issue prices.” For both individual and corporate investors, the optimal issue price is an increasing and concave function of bond maturity. This is a consequence of the simple principle highlighted by the model of Section 3: in order to ensure that potential future losses are fully tax deductible, price should never fall below par. Other things equal, a longer-term bond has a more volatile price, and a higher issue price is needed to accommodate this volatility.

### 4.5.2 Value of the tax option for the investor

Figure (6) plots the investor’s gain from following the optimal trading strategy (realizing gains and losses optimally) compared to a buy-and-hold strategy. The gains is approximately 15 basis points per year for individual investors, and 25 for insurers. This gain is large for a risk-free fixed-income strategy. Stock trading strategies such as value or momentum promise annual “alpha” that is an order of magnitude larger; tax trading
Figure 5: The “term structure of issue prices.” Optimal issue prices (expressed in percentage points of par) are an increasing and concave function of bond maturity. To ensure that potential future losses are fully tax deductible, price should never fall below par. Other things equal, a longer-term bond has a more volatile price, and hence requires more “protection” in the form of a higher issue price. The presence of “steps” in the graph is an artifact of the coarse grid search optimization described in Subsection 4.4.

of tax exempt bonds produces “alpha” that is based on ex ante theory, and therefore not subject to “model risk,” and risk free, because it is generated through tax arbitrage. Quantitative stock strategies are actually very risky, even if the model is correct, and even though they may compensate very generously for this risk. Finally, tax trading does not require short selling, or the purchase of a myriad securities. The strategy can be carried out by the owner of even a single bond.

4.5.3 Value of optimal issuance for an issuer

Figure (7) plots the issuer’s gain from issuing a bond with the optimal coupon, compared to the benchmark case of issuing at par. For instance, if the ex-ante value of the tax option is 3 percent of the issued amount when the bond is issued optimally, and I

19The optimal trading strategy does entail some change in the timing of cash flows, but an investor who is averse to interest rate risk can hedge it very cheaply using forward contracts. Moreover, tax rate risk, i.e. the possibility that the same individual faces different tax rates at different points in time, is a minor issue. Most of the option value comes from realizing losses, so that the benefits materialize immediately at a known tax rate, and the costs materialize in the future, if ever. For these future costs, I have assumed that the investor is in the top tax bracket, and therefore it is unlikely that the tax rate will be higher in the future, unless the government actually increases the tax rates. For the current benefits, for individuals, I have conservatively assumed that losses are deducted at the capital gain tax rate, while in actuality some losses can be deducted from ordinary income at a more favorable rate. In general, having time-varying tax rates creates more, not less, option value for taxpayers.
Why are municipal bonds issued at a premium?

Figure 6: Tax option value as a fraction of issue price. By realizing gains and losses optimally, investors can obtain a considerable improvement over a buy-and-hold strategy.
Figure 7: Tax option value as a fraction of issue price: optimal coupon improvement over par issuance.
Why are municipal bonds issued at a premium?

percent when the bond is issued at par, then the gain is 2 percent.

From the Figure it is evident that once again, the benefit is large. The gain from issuing optimally is between 0.5 and 2 percent of the principal amount issued. To put this number in context, the cost of issuance itself is in the order of 0.25 to 0.95 percent of the amount issued, and therefore setting the coupon right can turn bond issuance into a profitable business!

4.5.4 Optimal trading strategy

The optimal trading strategy is visualized in Figures 8 (no transaction costs) and 9 (with transaction costs). For simplicity, here I describe only Figure 8, the strategy in the case of no transaction costs and for individual investors. Each plot has time to maturity on the horizontal axis and tax basis on the vertical axis. Each plot corresponds to a unique value of the current interest rate \( r_t \) (2%, 4.5%, 5.5% and 8% respectively), and all plots refer to a 5 percent coupon bond. Thus, each point on a plot represents a state vector \((t, r_t, B_t)\) and can be associated with an optimal trading strategy: “sell” (white) or “don’t sell” (gray).

Without transaction costs, the strategy consists of up to three thresholds, separating “sell” areas from “don’t sell” areas. (“Sell” here indicates a wash sale, in which the investor realizes unrealized gains or losses but the bond does not actually change hands.) The first threshold is the price. Because the bond’s trading price is a function only of time to maturity \( T - t \) and current short rate \( r_t \), the price can be drawn on each plot as a function of time. Points above the line represent unrealized losses, because the investor’s tax basis is higher than the price; vice versa, points below the line represent unrealized gains.

The second threshold is the “Time Value of Money” (TVM) wall. Because of the specific tax rules, if maturity is far enough, it is profitable to realize losses below par, as the benefit is immediate and the cost is spread over a long period of time. However, if maturity is close enough, it is profitable to realize gains below par, precisely for the opposite reason - the cost is now, but the benefit will be realized quickly enough and it
Figure 8: Optimal trading strategy with no transaction costs (see explanation in text).
Why are municipal bonds issued at a premium?

Figure 9: Optimal trading strategy with transaction costs (see explanation in text).
is larger in present value.

The third threshold is the “Not-Enough” wall. To the right of the TVM wall, realizing gains (losses) below par is profitable (costly). However, realizing gains (losses) above par is costly (profitable). Sometimes, in order to realize gains below par, one has to realize gains above par as well. “Not-Enough” means that the gains or losses of the good type are not enough to make up for the cost of realizing the gains or losses of the bad type, and the investor should not sell.

5 Evidence: The features of actual tax-exempt bonds

This Section has evidence about the theoretical predictions of Section 3.

5.1 Data

The data are a combination of information from Bloomberg and the National Association of Insurance Commissioners (NAIC). NAIC filings provide a quarter-by-quarter history of all bond positions of all insurers. I selected the top 2,100 property and casualty insurers, corresponding to about 99.5% of all bonds by dollar value of positions, and I took all unique CUSIP numbers, about 200,000 unique bonds. The data was available from 2004q4 to 2012q3 at the time of acquisition. For each of these bonds, Bloomberg provides a near-official issue price, whether the bond has an issuer call option, and when it can be exercised; and the taxable status of the security (because multiple classes of municipal bonds are taxable).

5.2 Optimality of premium and the term structure of issue prices

A striking confirmation of the model’s predictions is found in Figures 10 and 11. Municipal bonds are indeed issued at large premia to par, in a sample including bonds issued throughout 20 years, and the term structure of issue prices is increasing and concave in the maturity of the bonds for noncallable bonds.
Why are municipal bonds issued at a premium?

Figure 10: Distribution of issue premium by issuer type in a sample of about 200,000 bonds held by P&C insurance companies between 2004 and 2012. The box contains 50% of the empirical distribution; only a few outlying datapoints are outside the area enclosed by the whiskers. Municipal bonds are well known for being issued at a premium. Such a phenomenon is consistent with maximization of the value of embedded tax options.
Figure 11: Distribution of issue premium by maturity. The box contains 50% of the empirical distribution; only a few outlying datapoints are outside the area enclosed by the whiskers. The stylized model of Section 3 predicts that issue price of noncallable tax exempt bonds is increasing in expected price volatility. This figure confirms that prediction. Bond maturity is a natural driver of ex-ante expected volatility for non-callable bonds. This pattern is very robust, and it emerges within every single vintage of issuance.
Actual issue prices (and coupons) are lower than the optimal ones predicted by the dynamic programming exercise. The model in this paper takes into account taxes only; perhaps, there are other factors that affect the optimal coupon, and further research is warranted. This is good news, however, because it means that taxes alone would be enough to explain even larger issue premia; the opposite pattern (actual coupons higher than optimal) would have been much harder to justify.

An important caveat regards the inherent sample selection problem. Insurance companies may be more tax-conscious than the average individual, and thus this sample of bonds may exaggerate the strength of the findings. In any case, this is either evidence that issuers issue intelligently designed securities, or that insurance companies choose carefully what securities to buy.

6 Conclusion

In this paper I show that after-tax pricing of tax-exempt bonds has security design implications for tax-exempt issuers. In a stylized model of optimal issuance, I derive that (i) tax-exempt bonds should be issued at a premium, and (ii) issue price should be increasing in the expected volatility of the bonds and in the expected change in yields—all to maximize the future gain and loss harvesting opportunities for the investor. I refine these qualitative predictions using dynamic programming to price tax-exempt bonds: with some degree of mean reversion in yields, I generate the additional prediction that coupons are “sticky”, i.e. they stay high when yields fall. I show empirical evidence consistent with tax options being priced, and with issuers issuing optimally designed securities. This is striking evidence that at least some institutional bond buyers are very sophisticated, and that market prices transmit information efficiently even in an opaque and illiquid market such as that for municipal bonds.
References


A Differences between the tax code used in the dynamic programming exercise and the actual tax code

The tax code used in the dynamic programming exercise is relatively similar to the actual U.S. federal income tax code. For an investor who buys the bond for a price $P_t$, there are three scenarios:

1. $P_t = 1$: the bond is bought at par. In this case, the tax basis is 1 as long as the bond is held. All coupon cash flows constitute taxable income and are taxed at a rate $\tau_E$. This rate is zero for individual investors, but it need not be zero for all investors.

2. $P_t > 1$: the bond is bought at a premium. In this case, the tax basis is equal to $P_t$ when the bond is bought. The excess over face value $P_t - 1$ is called “premium” and it is amortized over the remaining life of the bond using the constant yield method. The net income (coupon minus amortization amount) is taxed at a rate $\tau_E$.

3. $P_t < 1$: the bond is bought at a discount. In this case, the tax basis is equal to $P_t$ when the bond is bought. The discount to face value $1 - P_t$ is amortized over the remaining life of the bond using the constant yield method. The amortization amount is added to the basis; the same amount is also recorded as market discount interest income and taxed at a rate $\tau$. The entire coupon income is taxed at a rate $\tau_E$.

---

20The constant yield method is described in detail by IRS Publication 550. In short, the taxpayer is to calculate the yield to maturity $y$ at the time of purchase. Every year, the taxpayer multiplies $y$ times the bond’s book value (the tax basis). The difference between this amount and the coupon is the (potentially negative) amortization amount. The amortization amount should be added both to the tax basis and to income, but which type of income depends on the rules explained in the text.
The three differences concern the timing of market discount interest, accounting for original issue discount, and so-called “de minimis” discount.

**Market discount**  Under the actual tax rules, market discount interest income is not taxed period by period as it accrues, but it is taxed when the bond matures or when the investor disposes of it. This adds one dimension to the dynamic programming problem, because one needs to keep track of both the tax basis and the purchase price. On the other hand, by assuming that market discount income is taxed as it accrues, it is sufficient to keep track of the basis. This can slightly change both the optimal trading strategy and the value of the bond. However, the intuition from Section 3 suggests that rarely will an optimally issued bond trade at a market discount.

**Original issue discount**  Under the actual tax rules, the threshold for market discount income is not always par (1). If the bond is issued with original issue discount (OID), i.e., the issue price is less than face value ($P_0 < 1$), the OID rules provide a way to calculate the boundary at every period. This is ignored, once again, to reduce the dimensionality of the dynamic programming problem; because otherwise one needs to keep track of the issue price. This assumption is especially harmless because, as always, issuing at a premium is optimal, and ignoring the OID rules affects only bonds issued at a discount, which are known ex-ante to be suboptimal.

**De minimis discount**  Under the actual tax rules, when market discount is very small (“de minimis”), it is taxed at the capital gains tax rate. While this would not require any extra state variables, it makes things more complicated and its quantitative effect is, effectively, “de minimis”. Therefore, this aspect can be safely ignored for the purposes of the present paper.
B  Asset pricing with taxes using dynamic programming

Description of the problem

In a problem with taxes like the one encountered in this paper, the following situation often arises: an investor needs to decide how much to pay for an asset, but the price paid today affects the tax basis tomorrow, and therefore it affects future after-tax cash flows and therefore it affects the price itself. This Appendix shows that in essentially all cases, it is possible to find the price numerically using value function iteration.

Some iterative method appears to be used in practice by the few people who actually had to solve this problem, both academics such as Constantinides and Ingersoll (1984), and practitioners such as Kalotay Analytics (for after-tax pricing of municipal bonds) and PORTAX (for after-tax mean-variance optimization). While none of these people can be blamed for its absence, I am not aware of any published proof that iterative techniques should work under very general conditions.

Simple example

To understand what this means, suppose there is a bond trading today at a price \( P \). Tomorrow, the bond pays back its face value plus a coupon: \( 1 + C \). Unfortunately, the interest income is taxable at a rate \( \tau \). Define interest income \( I \) as

\[
I \equiv 1 + C - P
\]

and thus the after-tax cash flow tomorrow is

\[
1 + C - \tau (1 + C - P) = (1 + C)(1 - \tau) + \tau P.
\]

In this instance, \( P \) is called the “tax basis” and it is directly evident that a higher tax basis today means more after-tax cash flow tomorrow.

However, \( P \) is also the price the investor pays today. If the applicable discount rate
is $r$, the price of the bond is

$$P = \frac{1}{1+r} [(1+C)(1-\tau) + \tau P]$$

In this case, it is easy to solve:

$$P = \frac{1 - \tau}{1 + r - \tau} (1 + C)$$

However, often it is not practical to solve analytically for $P$ (e.g., in the presence of uncertainty in discount rates, or in the timing or magnitude of cash flows). The next section defines and solves the problem in its most general form.

**General result**

Suppose that at time $t$ the value of holding an asset position is

$$V(t, B_t, s_t)$$

where $B_t$ is the tax basis and $s_t$ is a random state vector that does not depend on $P_t$. Suppose also that the tax basis is equal to the purchase price at the time of purchase, and then it evolves deterministically according to some rules. Thus, the price an investor is willing to pay for an asset is equal to

$$P_t = V(t, P_t, s_t) \quad (19)$$

The question is whether it is safe to start with a guess for $P_t$, plug it into the right-hand side, and iterate until convergence. Here I try to show that this procedure is guaranteed to work, because $V(\cdot)$ is a contraction mapping with respect to its second argument, $B_t$. In order to show this, it is necessary to be a little more explicit on the form of $V(\cdot)$:

$$V(t, B_t, s_t) \equiv \frac{1}{1+r_t} \left( \mathbb{E}^Q [V(t+1, B_{t+1}, s_{t+1}) + CF_{t+1} | s_t] \right) \quad (20)$$
Why are municipal bonds issued at a premium?

Where \( \mathbb{E}^Q \) is the risk-neutral expectation operator. Then, make the following technical assumptions:

1. \( CF_{t+1} \) is the next-period pretax cash flow promised by the asset, and it does not depend on how much an investor pays for the asset at the current period.

2. \( r_t \geq 0 \): the one-period risk-free rate is nonnegative at all periods.

3. The top marginal tax rate \( \tau \) is less than 100%.

4. The tax basis converges monotonically to a known finite number.

Then, it is easy to show

**Proposition** \( \frac{\partial V(t, P_t, s_t)}{\partial P_t} < 1 \)

**Proof** The derivative of (20) with respect to \( B_t \) is

\[
\frac{\partial V(t, B_t, s_t)}{\partial B_t} = \frac{\partial}{\partial B_t} \frac{1}{1 + r_t} \left( \mathbb{E}^Q [V(t+1, B_{t+1}, s_{t+1}) | s_t] + CF_{t+1} \right) =
\]

\[
= \frac{1}{1 + r_t} \mathbb{E}^Q \left[ V_B(t+1, B_{t+1}, s_{t+1}) \right] \cdot \frac{dB_{t+1}}{dB_t} | s_t | < 1
\]

(21)

Assumption 1 guarantees that \( CF_{t+1} \) drops out when we take derivatives. Assumption 2, 3 and 4 guarantee the three bracketed inequalities of (21). Assumption 2 is self-explanatory. Assumption 3 guarantees the second inequality because one dollar of tax basis by definition reduces income from the asset by one dollar at some point in the future. If all the benefit from a one-dollar higher basis is realized now and at the top marginal tax rate \( \tau \), the benefit is \( \tau \). Assumption 3 guarantees that \( \tau < 1 \). Assumption 4 guarantees the third inequality because increasing the tax basis by one dollar today can increase the tax basis tomorrow by at most one dollar.

Finally, since (21) holds for arbitrary \( B_t \), setting \( B_t = P_t \) we obtain

\[
\frac{\partial V(t, P_t, s_t)}{\partial P_t} < 1.
\]
Proposition B, together with the fact that \( \frac{\partial V(t, P_t, s_t)}{\partial P_t} > 0 \), directly implies that \( V(\cdot) \) is a contraction mapping, and therefore the unique fixed point can be found by iteration starting from any guess for \( P_t \). In practice, the convergence is very fast.

There is a very good economic intuition between this result. Assumptions 1-4 really boil down to a simple and intuitive condition: that paying one dollar more for an asset today generates less than one dollar in present value of tax benefits. If this condition were violated in a real-world tax code, the government would go bankrupt very quickly and there would be no asset market equilibrium, as paying arbitrarily high prices for assets would make people richer (as long as the government is not bankrupt, that is). So, even if one can find examples where each one of the assumptions fails to hold,\(^{21}\) the result is very general and robust.

**A final thought on Blackwell sufficient conditions**

One may try to construe this tax problem as a special case of policy function iteration; the Blackwell sufficient conditions are met (there is discounting, and the value function is strictly increasing in basis, hence in the price paid today). The hard-to-accept part is that the “policy” function would be a joint vector \((P_t, \gamma_t)\) where \( \gamma_t \) is 0 when the asset is held, and 1 when a wash sale is executed; and \( P_t \) is the price that the investor is willing to pay to buy an additional unit of the asset. But \( P_t \) is not “optimal policy” in any sense—optimal would be to pay exactly zero, if the investor could choose; and of course, the investor is assumed to be competitive, hence unable to choose.

---

\(^{21}\)This is left as exercise for the creative reader.