

Tax Heterogeneity, Duration Risk, and the Term Structure: An Application to the Municipal Bond Puzzle*

Kevin Zheyuan Cui[†]

October 19, 2026

Abstract

I develop a term-structure model that jointly prices taxable and tax-exempt bonds. Heterogeneous taxes and balance sheets endogenously segment ownership by maturity under short-sale constraints. The model explains the municipal bond puzzle: short municipal yields closely match the after-tax taxable benchmark of top-bracket households, who face the highest taxes, while long yields exceed it. Lower-tax investors whose liabilities offset interest-rate risk bear duration risk more cheaply and hold long municipal bonds despite their smaller exemption benefits. Yields reflect both holders' taxes and risk-bearing costs. Without targeting municipal yields, the calibrated model explains 83 percent of the 118-basis-point thirty-year gap. Empirically, household tax effects weaken with maturity while institutional effects strengthen, as the mechanism predicts. Through portfolio substitution, investor tax policies affect bond yields in both taxable and tax-exempt markets.

Keywords: Term Structure, Investor Heterogeneity, Taxes, Endogenous Segmentation, Asset Ownership, Municipal Bonds

JEL Codes: G12, G14, G22, H24, H63

*I am grateful to Moritz Lenel, Wei Xiong, and Motohiro Yogo for their guidance. I also thank Yacine Ait-Sahalia, Caio Almeida, Markus Brunnermeier, Fousseni Chabi-Yo, Aditya Chaudhry, Daniel Chen, Zhiyu (Julie) Fu, Nicolae Garleanu, Sergei Glebkin, Zhouzhou Gu, Nicolas Hommel, Paymon Khorrami, Jane Li, Shiqi Li, Ziang Li, Jonathan Payne, Yinan Qiu, Julia Selgrad, Amy Huber Wang, Lingxuan (Sean) Wu, Yinjie Yu, Yao Zeng, William (Yinhong) Zhao, Haonan Zhou, and Haoxing Zhu, and participants at the MPSS Conference and Princeton University Seminars for constructive feedback.

[†]Princeton University, Department of Economics.

1 Introduction

Whose taxes determine bond yields when investors pay different tax rates on the same bonds? U.S. municipal bonds provide a natural setting to study this question because the same tax exemption offers different benefits to investors facing different tax rates. These bonds finance state and local governments through a market of about \$4 trillion. Municipal bond interest is exempt from federal income tax, and the exemption is worth most to top-bracket households, who face the highest tax rates. The standard benchmark in the literature therefore uses these households' tax rate to determine municipal yields: at every maturity, the municipal yield should equal their after-tax yield on a comparable taxable bond (Schwert, 2017; Fleckenstein and Longstaff, 2025). The household benchmark fits short maturities, but thirty-year municipal yields exceed it by 118 basis points. The marginal tax rate implied by observed municipal and taxable yields falls from about 38% at two years, close to the top-bracket household rate, to just 21% at thirty years. This municipal bond puzzle has been documented for over thirty years (Feenberg and Poterba, 1991). Under the U.S. tax code, however, investors face very different tax rates, and the data suggest that top-bracket households hold only about half the municipal market. Lower-tax households, banks and insurers hold large amounts even though the exemption is worth less to them. Understanding relative yields therefore calls for a model that accounts for this heterogeneity.

Motivated by this heterogeneity in ownership, I explain the puzzle with an equilibrium term-structure model in which multiple investor groups differ in their taxes and the duration of their liabilities. They choose portfolios jointly across taxable and tax-exempt markets and maturities, subject to short-sale constraints. Their portfolio choices endogenously segment ownership and determine both yield curves. Tax differences draw high-tax investors toward tax-exempt municipal bonds and lower-tax investors toward taxable bonds. However, long-term liabilities allow some lower-tax investors to bear interest-rate risk more cheaply than high-tax households. When rates rise, the value of these liabilities falls, offsetting losses on their bonds. This advantage grows with maturity and draws them into long municipal bonds despite their smaller exemption benefits. Because the exemption saves these investors less tax, they accept a smaller reduction in yield relative to taxable bonds than high-tax households do. Long municipal yields therefore reflect their smaller tax benefits, explaining why yields exceed the household benchmark. Calibrated to U.S. taxes, balance sheets and holdings, without targeting municipal yields, the model explains 83% of the thirty-year puzzle. Untargeted empirical tests support its prediction that household tax effects weaken with maturity while those of banks and insurers strengthen. The framework applies beyond municipals because it jointly determines relative yields across markets. In the counterfactuals, investor substitution transmits tax reforms to the financing costs of both municipal and taxable issuers.

Household mortgage debt, banks' assets and deposit funding, and insurers' long-term obligations create different net exposures to rate movements. A two-investor illustration shows how these balance-sheet differences interact with exemption benefits to determine ownership across maturities. The high-tax household gains more from exemption, while the lower-tax investor has a stronger

liability hedge and therefore a lower cost of bearing rate risk. The resulting duration benefit grows with a bond's rate exposure and hence with maturity, whereas the difference in exemption benefits stays constant. Beyond a maturity threshold, the duration benefit outweighs the exemption disadvantage. In this illustration, the household therefore holds shorter municipal bonds, while the lower-tax investor holds longer municipal bonds and risky taxable bonds.

This sorting raises municipal yields above the household benchmark through three channels. First, lower-tax investors value the exemption less and therefore accept a smaller yield reduction, keeping the long municipal yields they price closer to taxable yields than the benchmark predicts. Second, the household alone holds shorter risky municipal bonds and bears rate risk at a higher cost than the lower-tax investor holding comparable taxable bonds. The household's higher required compensation raises municipal yields above the benchmark even at these maturities. Third, taxes scale risk as well as returns. Under the household benchmark, municipal yields and their sensitivity to interest rates equal the corresponding taxable quantities multiplied by the fraction of taxable interest the household keeps. Lower-tax investors keep a larger fraction, so the long municipal bonds they price are more sensitive to rates than this benchmark predicts. Holders require compensation for that additional rate risk, and household portfolio substitution carries the effect to shorter municipal bonds. All three channels vanish at the riskless shortest maturity, where the household's tax advantage determines relative yields.

Other motives can also lead lower-tax investors to hold long municipal bonds and generate the same qualitative pricing pattern. Lower trading costs or better default diversification can do so when the advantage grows with maturity and applies equally to taxable and tax-exempt bonds (Section 4.7). I focus on liability hedging because investors' widely differing liability durations provide an economically important and measurable source of differences in the cost of bearing interest-rate risk. Observed balance sheets and maturity holdings discipline these differences in the quantitative model.

The quantitative application asks how much of the puzzle observed taxes and portfolios can explain. It includes high- and lower-tax households in each state, banks, property-and-casualty insurers, life insurers and tax-neutral investors trading in state municipal markets and a common taxable market. I measure tax treatment, bond supplies, and investor balance-sheet sizes directly from data and estimate the rate process from historical short rates. Observed ownership and the maturities investors hold mainly discipline their unobserved exposure to interest rates, while taxable yields help discipline risk aversion. Municipal yields and empirical test coefficients are not targeted in calibration. The model accounts for 97 of the 118 basis points at thirty years, and measured default, liquidity and capital-gains differentials explain almost all of the remainder.

I evaluate the model against the household benchmark using three empirical tests, none targeted in the calibration. The model predicts that variation in top-bracket households' exemption benefits is fully reflected in implied taxes at the shortest maturity and less strongly at longer maturities. Changes in lower-tax investors' exemption benefits have no effect at the shortest maturity and matter more at longer maturities. First, most states exempt residents' interest on their own

municipal bonds from state income tax. Cross-state variation in state tax rates is reflected almost one-for-one in implied state taxes at short maturities and progressively less at longer maturities. Second, bank-qualified designation gives banks an additional tax benefit on a subset of municipal bonds issued by small borrowers. This bank-only benefit is barely reflected in short municipal yields and increasingly reflected at longer maturities. Third, I compare municipal yield changes with changes in the treasury reference yield at the same maturity. This relative response rises with maturity, consistent with lower-tax holders making long municipal yields more sensitive to interest rates.

Two counterfactual reforms show how removing different tax benefits changes financing costs and tax receipts. The first removes a state's exemption for residents' interest on its own municipal bonds, as Illinois, Iowa and Wisconsin generally do. Residents retain the federal benefit of holding municipal bonds but no longer have a tax incentive to favor home-state over other municipal bonds, like those residents in states without an income tax. High-tax households therefore mainly switch between municipal markets, with relatively modest changes in the maturities they hold. The state taxes the interest on these replacement investments, so selling local bonds does not substantially erode its tax base. Meanwhile, long municipal yields also reflect the valuations of institutions and nonresidents whose tax treatment does not change. Long yields therefore respond only partly to the lost household exemption, consistent with the partial pass-through of state taxes at long maturities documented in the first empirical test. Across seven state experiments averaged over 2000–2017, added state receipts exceed local issuers' added financing costs: own-state issuers incur \$0.68 to \$0.88 in added annual financing costs per dollar of added state receipts. Portfolio substitution also raises borrowing costs for other states' municipal issuers and taxable issuers.

The second reform removes the federal exemption as well, eliminating the tax advantage of municipal over taxable interest for every investor. Taxable municipal issuance has a precedent in the 2009–2010 Build America Bonds program, which subsidized issuers' interest payments. High-tax investors can no longer preserve that exemption benefit by switching between state municipal markets. They reduce their municipal holdings and shift toward shorter bonds or cash, which generate less interest than long bonds and limit the revenue collected from repeal. Lower-tax and tax-neutral investors take over more long bonds, but the reform imposes little additional tax on their interest income. They also require higher yields to bear the additional rate risk, raising borrowing costs in both markets. The increase in financing costs consequently exceeds added tax receipts. Averaged over 2000–2017, municipal and taxable issuers together incur \$1.20 in added annual costs per dollar of added federal and state receipts, including \$0.13 borne by taxable issuers. Annual financing costs rise by \$46.30 billion, while annual tax receipts rise by \$38.71 billion. The excess becomes higher aggregate after-tax interest income, distributed unevenly across investors. Lower-tax households, life insurers and tax-neutral investors gain, while high-tax households, banks and property-and-casualty insurers, which previously benefited more from the exemption, lose on average.

The framework is not only about municipal bonds. It explains relative bond pricing across taxable and tax-exempt markets, with investor portfolio substitution transmitting tax changes in one market to yields in the other. I apply it to municipal and taxable bonds because municipal tax rules provide clean variation to test the mechanism and address a longstanding pricing puzzle. Its applications extend to other settings where tax treatment differs across investors and markets, including state-tax differences between treasury and corporate bonds, tax-favored accounts and cross-border withholding. More broadly, heterogeneous taxes, together with balance-sheet differences and limits to arbitrage, shape endogenous market segmentation and are important for understanding bond yields and the effects of tax policy.

Sections 2–3 review related work and present the tax, ownership and yield evidence. Section 4 develops the model, Sections 5–6 evaluate its quantitative and empirical implications, and Section 7 studies tax policy. Formal results, proofs and supporting evidence are in the appendix.

2 Related Literature

The paper contributes to research on how taxes affect portfolio choice and asset prices. Theoretical work studies how tax differences shape portfolios and equilibrium prices (Miller, 1977; Auerbach and King, 1983; Dybvig and Ross, 1986; Dammon and Green, 1987), and Sialm (2009) provides evidence that taxes are reflected in equity valuations. I study how these differences shape the bond yield curve. Under short-sale constraints, investors choose jointly across markets and maturities. Their taxes and costs of bearing rate risk determine who holds each bond and whose tax treatment its yield reflects.

Research on market segmentation links the yield curve to bond supply, investors’ maturity preferences and their capacity to bear risk (Vayanos and Vila, 2021; Greenwood and Vayanos, 2014; Guibaud et al., 2013). Portfolio adjustment also transmits supply shocks across asset classes and currency bond markets (Greenwood et al., 2018, 2023).¹ I explain how taxes generate partial segmentation across markets and maturities. Participation responds to taxes and balance-sheet exposures, so tax changes alter both yield curves through joint portfolio choice.

The model’s explanation of participation connects to research on institutional balance sheets and asset pricing. Intermediaries’ capital and exposure to interest-rate movements affect the risk premia they require (He and Krishnamurthy, 2013; Haddad and Sraer, 2020). Kekre et al. (2024) show how losses on arbitrageurs’ bond portfolios can amplify the effect of monetary tightening on long yields. Insurers’ long-term liabilities also shape their demand for long bonds and their capacity to bear credit risk (Domanski et al., 2017; Li, 2026).² I show how a balance-sheet advantage in bearing rate risk can offset a smaller tax benefit, explaining why some lower-tax investors hold long tax-exempt bonds.

¹Constraints on arbitrage can limit this adjustment (Gromb and Vayanos, 2018; Siriwardane et al., 2025). Related empirical work measures investors’ asset demand and household rebalancing (Kojen and Yogo, 2019; Gabaix et al., 2026).

²Municipal offering spreads reflect banks’ rate exposure and funds’ liquidity needs (Pati, 2023). Liability structures also help explain investors’ responses to credit-market downturns (Coppola, 2025).

The municipal bond puzzle has motivated a longstanding literature on relative bond yields. Early explanations emphasized tax benefits and investor sorting across maturities (Mussa and Kormendi, 1979; Hendershott and Koch, 1980; Kidwell and Koch, 1983; Skelton, 1983). The Tax Reform Act of 1986 reduced banks’ tax advantages and municipal holdings (Hines and Horton, 2024), yet long municipal yields remained high relative to after-tax taxable yields (Green, 1993; Chalmers, 1998). Subsequent research examined the taxation of bond cash flows, tax risk, default risk, liquidity, convenience premia and callability (Ang et al., 2010; Longstaff, 2011; Schwert, 2017; Fleckenstein and Longstaff, 2025; Brown, 2011).³ I assess bond-side adjustments on a common household after-tax scale and find that they leave a substantial long-maturity gap. My model explains most of this remaining gap through endogenous participation by investors with different taxes and costs of bearing rate risk.

Empirical research on municipal tax policy links state tax benefits to local ownership, financing conditions and risk sharing across states (Hendershott and Kidwell, 1978; Pirinsky and Wang, 2011; Babina et al., 2021; Moussawi, 2022, 2024). Changes in banks’ tax benefits likewise affect municipal investment and access to financing (Hines and Horton, 2024; Dagostino, 2025; Benson and Marks, 2024).⁴ I test how household and institutional tax effects vary with maturity. Household state-tax pass-through declines, while the yield reduction associated with bank qualification rises. This opposite pattern tests the participation mechanism against a single household tax rate applied across the curve.

Fiscal analyses of municipal tax exemption examine its revenue cost and benefits to borrowers. Poterba and Ramírez Verdugo (2011) show that estimates of forgone revenue depend on how households would reinvest without the exemption. Garrett et al. (2023) show how auction competition affects the borrowing-cost savings from tax benefits. I evaluate reforms by running counterfactuals that incorporate the equilibrium impact of investor reallocation, capturing changes in tax receipts and financing costs.

3 Taxes, Ownership and the Municipal Puzzle

This section establishes the empirical setting for the model. I describe the data and the tax benefits available to different investors, then show that municipal ownership extends well beyond high-tax households. I measure the municipal puzzle and examine whether bond characteristics or the choice of taxable benchmark explain its maturity pattern.

³Mankiw and Poterba (1996) link the two bond markets through equity holdings. Related pricing, tax-policy and maturity-clientele analyses include Fama (1977); Livingston (1982); Poterba (1986); Brown and Sirmans (2011). Trading-cost evidence informs my liquidity adjustment (Harris and Piwowar, 2006; Green et al., 2007; Bagley et al., 2023). For surveys, see Cestau et al. (2019); Spreen and Gerrish (2022). Appendix C.5 compares the model with Green’s benchmark.

⁴Municipal fund flows and fragility also affect yields and financing (Li et al., 2024; Adelino et al., 2026).

3.1 Data

I combine data on bond issuance, investor holdings and taxes. Mergent provides issuance records, ratings, bond characteristics and call schedules. Insurers’ filings in Schedule D of the National Association of Insurance Commissioners (NAIC) and a pension holdings panel collected through Freedom of Information Act requests allow me to examine institutional portfolios at the level of individual bond holdings. I use the Financial Accounts to measure aggregate holdings, the Survey of Consumer Finances to measure household balance sheets, and IRS Statistics of Income to document household tax-exempt interest income across income brackets.

The main yield comparison covers 2000–2020. Its taxable benchmark is FRED’s A-or-above corporate zero-coupon curve. Federal Reserve treasury curves (Gürkaynak et al., 2007) provide the reference for the interest-rate response test. I measure trading and liquidity using municipal transactions from the Municipal Securities Rulemaking Board (MSRB) and corporate transactions from the Trade Reporting and Compliance Engine (TRACE) Enhanced database. Markit provides municipal and corporate credit default swap quotes, and the Tax Foundation provides state tax schedules. Appendix D.7 details the sources and their coverage.

3.2 Municipal issuance

Table 1: Municipal issuance: descriptive statistics (sample means).

Variable	Full sample	Tax-exempt, non-AMT	
		AA and above (score ≤ 3)	A– and above (score ≤ 7)
Offering yield (%)	3.11	2.82	2.96
Corporate benchmark yield (%)	4.47	4.33	4.41
Maturity at issuance (years)	10.15	10.76	10.86
Offering price (per \$100 par)	102.88	104.68	103.85
Issue rating (1 = AAA, 3 = AA, 7 = A–)	3.80	2.18	3.55
Implied tax rate $1 - y^M/y^C$	0.29	0.35	0.33
Share callable	0.46	0.51	0.51
Share tax-exempt, non-AMT	0.92	1.00	1.00
Sample size (CUSIPs)	3,014,420	984,210	1,990,791

Notes. The table reports bond-level means for issues of 2000–2020 before the call and issuance-premium adjustments, with one observation per municipal issue (CUSIP). The full sample contains all issues. The AA-and-above and A-or-above columns restrict to federally tax-exempt, non-AMT issues with average agency score ≤ 3 and ≤ 7 , the latter matching the A-or-above corporate credit band. The implied-tax row averages the bond-level ratio $1 - y^M/y^C$, winsorized at the first and ninety-ninth percentiles, so it need not equal one minus the ratio of the displayed yield means. Each row uses non-missing observations, so the full-sample rating mean uses the 2,292,092 rated issues. The corporate benchmark is the corporate zero yield for the issue month and maturity.

Table 1 describes municipal bonds issued during 2000–2020, with one observation per bond identifier (CUSIP). The full sample contains more than three million issues, with an average maturity of about ten years. Most pay federally exempt interest: taxable municipal bonds account for 5.8% of issues, bonds subject to the alternative minimum tax (AMT) account for 1.9%, and the remaining issues are tax-exempt, non-AMT bonds. Call provisions allow issuers to redeem bonds

before maturity and are common in the data. Callable bonds account for 46% of all issues and 51% of the high-grade, tax-exempt samples, making callability important for comparing yields.

The table also reports two high-grade samples of tax-exempt, non-AMT bonds. The broader sample includes A– and above issues (average agency score ≤ 7 on the twenty-two-notch scale) and matches the credit band of the corporate benchmark used for the main yield curves. The main empirical specifications use the AA-and-above sample (≤ 3) to minimize the impact of default risk.

3.3 Bond taxation and the representative-tax benchmark

For a taxable household, interest on taxable bonds is ordinary income, while interest on tax-exempt municipal bonds is federally exempt. Interest includes both coupon payments and annual original-issue-discount accrual. Institutions generally do not receive the full exemption benefit from holding municipal bonds. For example, property-and-casualty (P&C) insurers pay an effective federal tax of 5.25% on municipal interest. Let τ_i^C and τ_i^M denote investor i 's effective tax rates on taxable and municipal interest. I define the relative after-tax share as the fraction of taxable interest retained divided by the fraction of municipal interest retained:

$$\xi_i \equiv \frac{1 - \tau_i^C}{1 - \tau_i^M}. \quad (1)$$

Investor i 's relative municipal tax benefit, the exemption benefit, is $1 - \xi_i = (\tau_i^C - \tau_i^M)/(1 - \tau_i^M)$. A smaller ξ_i means a larger benefit. When municipal interest is fully exempt, $\xi_i = 1 - \tau_i^C$, so the benefit equals the tax rate on taxable interest. Appendix D explains the underlying tax rules.

The representative-tax benchmark uses the top-bracket household's tax treatment at every maturity. Write y_T^M and y_T^C for municipal and taxable yields at maturity T , and H for this household. Under the benchmark, the two bonds offer the household the same after-tax yield, $y_T^M = \xi_H y_T^C$. I call $\xi_H y_T^C$ the household benchmark and $1 - \xi_H$ the household wedge. The benchmark's yield-implied tax is therefore $1 - y_T^M/y_T^C = 1 - \xi_H$. With fully exempt municipal interest, it equals τ_H^C . Investors with different ξ_i cannot all be indifferent at this yield ratio. Their participation in the municipal bond market is the central question of the model.

3.4 Investor taxes and ownership

The tax benefit of municipal interest varies substantially across investors in the roughly \$4 trillion municipal market. Table 2 summarizes these differences before the 2017 Act. For households and insurers, the net exemption benefit is $1 - \xi_i$ in (1). Households in higher tax brackets receive a larger benefit because they pay more tax on taxable interest. Exemption from home-state income tax adds to the benefit on in-state municipal bonds. Appendix D gives the detailed tax treatment of each investor group and the changes after the 2017 Act.

Insurers receive smaller benefits than high-tax households because exempt interest reduces their deductions. The reduction is less severe for P&C insurers than for life insurers, giving P&C insurers

Table 2: Investor heterogeneity in the municipal market before the 2017 Tax Cuts and Jobs Act

Investor type	Federal treatment	State treatment	Net exemption benefit
High-tax household	Interest exempt	Exempt on in-state bonds only	0.35–0.43
Middle-tax household	Interest exempt	Exempt on in-state bonds only	0.25–0.33
Low-tax household	Interest exempt	Exempt on in-state bonds only	0.10–0.15
P&C insurer	Exempt, less proration	Premium taxes, no distinction	0.31
Life insurer	Exempt, less proration	Premium taxes, no distinction	0.21
Bank	Exempt, funding cost disallowed	Varies by state	0.15–0.33
Other (pensions, etc.)	Varies	Varies	< 0.05

Notes. For households and insurers, the net exemption benefit is $1 - \xi_i$, from statutory rates and proration over 2000–2017. For banks it also deducts the financing-tax cost described below. Appendix Table 49 gives the values after the 2017 Act. Household values are federal marginal rates in three groups (35% and above, 25% to 33%, below 25%), and the top value includes the 3.8% net investment income tax from 2013. The quantitative model pools middle- and low-tax households into one lower-tax group, whose benefit with home-state taxes is 0.21 to 0.27 across states in 2015–2017, against 0.43 to 0.50 for the high-tax group (Table 3). Proration cuts a P&C insurer’s losses-incurred deduction by 15% of its exempt interest and a life insurer’s reserve deduction by the policyholders’ share. Section 265(b) disallows the share of a bank’s interest expense allocable to exempt holdings, of which 80% stays deductible for bank-qualified bonds. The bank value, net of that disallowance, is the break-even implied tax, equal to the bank’s tax rate minus its annual financing-tax cost per dollar of municipal holdings divided by the taxable yield. That cost is 1.02% in 2000–2002 and 0.11% in 2015–2017, and the bank range covers the six three-year periods of 2000–2017 and maturities of two to thirty years. States generally tax insurers through premium taxes, which do not distinguish municipal from taxable interest, and banks through franchise or income taxes.

the larger benefit. Banks generally lose the interest deduction allocated to municipal holdings, with partial relief for bank-qualified bonds, so the table reports their benefit net of this financing-tax cost.⁵ Tax-exempt institutions such as qualified pension funds and endowment funds already receive taxable interest free of current income tax and gain no additional benefit from the municipal exemption. They hold almost no municipal bonds (Appendix C.9).

Figure 1 shows that municipal ownership spans these tax groups. Panel (a) reports aggregate holdings: households and municipal funds hold the largest amounts, while banks and P&C insurers also hold substantial portfolios despite smaller exemption benefits. P&C insurers allocate about 40% of their bond portfolios to municipals and maintain a larger municipal portfolio share than life insurers throughout the sample (Appendix C.9). Panel (b) documents household tax-exempt interest income across income brackets. Assigning municipal-fund holdings to households and allocating household holdings using tax-exempt interest shares puts roughly half the market with high-tax households.⁶ The remaining holdings by investors with smaller exemption benefits motivate a model in which investors with different exemption benefits choose municipal and taxable bonds jointly. The analytical benchmark has two investors. The quantitative model includes household groups by state of residence, insurers, banks and a residual group (Section 5.1).

⁵The quantitative model charges the bank financing-tax cost separately. Table 3 therefore reports the bank’s exemption benefit $1 - \xi_i$ before that cost.

⁶I use IRS Statistics of Income shares of tax-exempt interest on taxable returns, including fund distributions, to approximate household holdings by income group in 2010–2017. Adjusted gross income of at least \$200,000 proxies for the high-tax group.

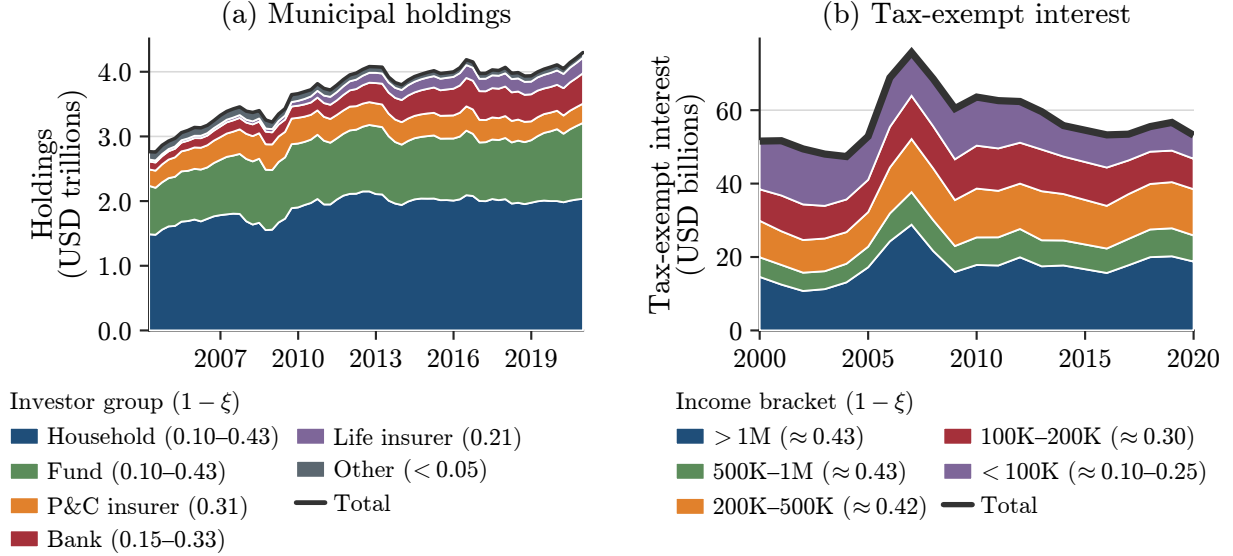


Figure 1: Municipal holdings and tax-exempt interest by investor group. Panel (a): direct holdings of taxable and tax-exempt municipal securities, 2005–2021 (USD trillions, Financial Accounts, with households including nonprofits). Panel (b): tax-exempt interest, an income flow, on taxable individual returns by adjusted-gross-income bracket, 2000–2020 (USD billions, IRS Statistics of Income). 2015 lacks comparable bracket detail and is omitted. Legend values in parentheses are net exemption benefits: in panel (a), the values before the 2017 Act from Table 2, with funds assigned their household shareholders’ range, and in panel (b), approximate values by bracket.

3.5 Limits to arbitrage

Tax rules and trading frictions limit the trades that could exploit yield differences or equalize investors’ costs of bearing rate risk. Interest and short-sale payments incurred to carry exempt bonds generally lose their deduction. Consider a household that shorts a long-maturity corporate bond to hedge its municipal holdings. With the deduction denied, an equal-principal position earns coupon carry equal to the pretax municipal-minus-corporate yield difference. This carry is negative along the sample-average par curves and about -95 basis points a year at thirty years. Municipal shorts are scarce and costly because substitute payments to the lender do not retain the interest exemption. Off-exchange swaps face eligibility restrictions, while futures losses can face deduction limits and deferral. The model summarizes these tax asymmetries and trading frictions through nonnegative bond positions and cash balances, ruling out short sales and borrowing. Appendix D.6 examines the tax treatment, carry and residual risk of these trades.

3.6 The municipal puzzle

The municipal puzzle is a maturity pattern: short municipal yields are close to the household benchmark, while long municipal yields substantially exceed it. To measure this pattern, I use the seasoned, option-free A-or-above corporate zero-coupon curve as the taxable comparison. High-grade corporate bonds are closer to municipals in credit and liquidity than treasuries. Treasury yields also contain a premium for safety and liquidity (Krishnamurthy and Vissing-Jorgensen, 2012),

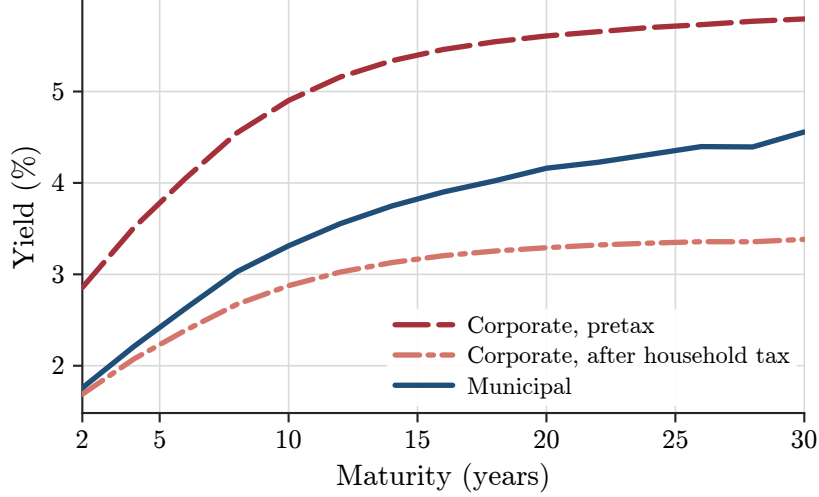


Figure 2: The municipal puzzle: A-or-above zero-coupon yields, 2000–2020. Municipal yields remove callability and the issuance premium. Averaging follows Appendix C.1.1. Appendix C.12.1 gives the AA-and-above counterpart.

which would enter the measured gap if treasuries supplied the benchmark. Studies using treasury yields (Chalmers, 1998) or swap rates (Longstaff, 2011) as the taxable benchmark also find the municipal puzzle.

I construct municipal zero-coupon yields in the same credit band from offering prices and contractual cash flows. Issuance data provide broad coverage and a common measurement date. Offering yields also avoid the direct tax effects of capital gains and market discounts arising from prior secondary-market trades. I remove callability and the issuance premium to make municipal yields comparable with the seasoned, option-free corporate curve.⁷ Let $y_t^C(T)$ denote the taxable yield and $y_{s,t}^M(T)$ the municipal yield after these two adjustments, where T is maturity, s is the issuer’s state and t is the issue date. The municipal puzzle is the gap between $y_{s,t}^M(T)$ and the household benchmark $\xi_H y_t^C(T)$. The yield-implied tax expresses the same comparison in tax-rate units:

$$ImpliedTax_{s,t}(T) \equiv 1 - \frac{y_{s,t}^M(T)}{y_t^C(T)}. \quad (2)$$

For a given household relative after-tax share ξ_H , the two measures satisfy

$$\underbrace{y_{s,t}^M(T) - \xi_H y_t^C(T)}_{\text{municipal yield gap}} = \underbrace{[(1 - \xi_H) - ImpliedTax_{s,t}(T)]}_{\text{shortfall from the household wedge}} y_t^C(T). \quad (3)$$

For positive taxable yields, a positive gap means an implied tax below the household wedge $1 - \xi_H$. In Figure 2, the municipal yield is near the household benchmark at the short end but about 118 basis points above it at thirty years. The implied tax falls from about 38% to 21% while the household wedge remains near 41% (Figure 7(d)).

⁷The market-discount tax applies to later secondary purchasers. Its expected effect on resale values can also affect offering prices. The capital-gains adjustment includes this effect as well as taxes on realized gains.

The declining implied-tax pattern also appears against treasuries (Chalmers, 1998), within issuer-matched tax-exempt and taxable municipals (Section 3.8), and in the AA-and-above and unlimited-tax general-obligation samples (Appendices C.12.1 and C.4).

3.7 Bond-side adjustments

Studies such as Fleckenstein and Longstaff (2025) use secondary-market prices of option-free bonds to examine the municipal puzzle. Consistent with that comparison, I remove callability and the issuance premium before defining the municipal puzzle as the gap over the household benchmark. I then assess default risk, liquidity and capital-gains taxation as explanations of this gap. Each adjustment measures the difference in compensation between municipal and corporate bonds, expressed in the household’s after-tax yield units. Compensation common to both markets therefore cancels. The household wedge, one minus the ratio of averaged after-tax to pretax corporate yields, averages 41.3% and is nearly flat across maturities. Appendix C explains the adjustments, and Appendix Table 25 reports their contributions by maturity. Appendix C.5 discusses the tax timing of taxable bonds in Green (1993).

At thirty years, the call and issuance-premium adjustments reduce the 160-basis-point spread over the household benchmark to the 118-basis-point municipal puzzle. The default, liquidity and capital-gains differentials explain 51% of the puzzle at ten years but only 16% at thirty. Subtracting them leaves 99 basis points at thirty years for comparison with the model in Section 5.4. At two years, their sum exceeds the small puzzle, leaving a residual of -9 basis points, consistent with the small convenience premium discussed in Fleckenstein and Longstaff (2025). The adjustments leave the central pattern intact: the residual widens with maturity.

3.8 Issuer-matched evidence

Taxable and tax-exempt bonds issued by the same municipal borrower provide a further comparison that does not rely on corporate bonds. During 2009–2010, the Build America Bonds (BAB) program allowed municipal issuers to sell taxable bonds with a federal subsidy equal to 35% of their interest payments. The top federal household income-tax rate was also 35%. I leave tax-exempt offering yields unchanged and multiply BAB yields by 0.65 to put both on the household’s after-tax scale. I then compare bonds from the same issuer and estimate how their yield difference varies with maturity:

$$y_b = \alpha_1 EXMP_b + \beta (EXMP_b \times n_b) + \gamma_C y_{n_b, t_b}^C + \lambda_{j(b) \times n_b \times q_b} + \varepsilon_b, \quad (4)$$

where y_b is the offering yield for a tax-exempt bond and 0.65 times the offering yield for a BAB. The indicator $EXMP_b$ equals one for a tax-exempt bond and zero for a BAB, and n_b is maturity in years. The fixed effect $\lambda_{j(b) \times n_b \times q_b}$ compares bonds from issuer $j(b)$ at the same maturity and in the same issue quarter q_b . The call-matched specification also requires the same call status. The corporate yield y_{n_b, t_b}^C , measured at the bond’s maturity and issue month t_b , controls for market

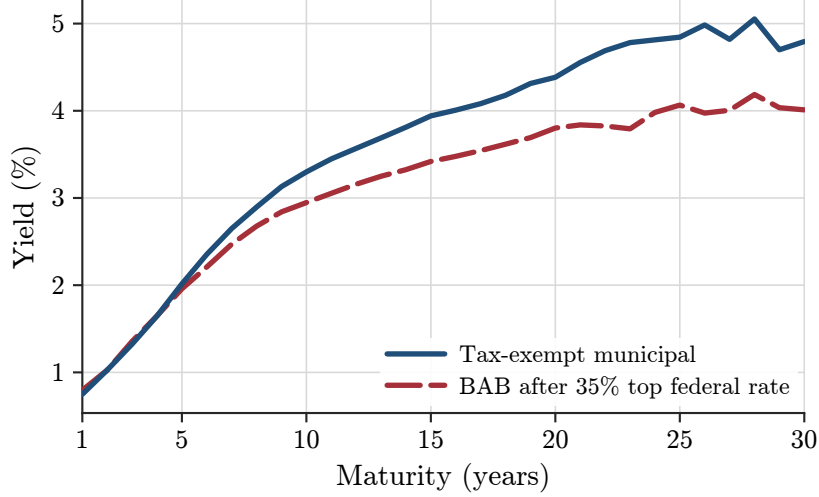


Figure 3: Tax-exempt and Build America Bond offering yields by maturity, AA-and-above issues, April 2009–December 2010. BAB yields are multiplied by 0.65, one minus the 35% top federal rate in 2009–2010. Appendix Table 28 reports regressions with and without call-status matching. Appendix C.3 describes the sample.

yield movements within that quarter. The coefficient β therefore measures how the tax-exempt yield minus the household after-tax BAB yield changes with maturity. Standard errors are clustered by state. Comparing bonds from the same issuer at the same maturity and with matched call status holds borrower identity and these bond characteristics fixed.

The estimated gradient β is 3.4 basis points per year in the AA-and-above sample and 3.3 in the A-or-above sample. Matching call status gives gradients of 3.1 and 3.0 basis points. The binned estimates also rise monotonically: in the AA-and-above sample with call matching, the [25, 30]-year coefficient is 63.9 basis points above the one-to-five-year bin. Figure 3 shows mean yields by maturity before issuer matching. Appendix Table 28 reports the issuer-matched estimates. Both show a widening gap, and the regressions establish that it persists within the same issuer, using offering yields for both bond types and matching call status.

4 Model

I develop a term-structure model in which investors differ in tax treatment and balance-sheet exposure to interest rates. They choose bonds jointly across markets and maturities, subject to short-sale constraints. The model answers two questions: why investors with smaller exemption benefits hold tax-exempt bonds, and how their participation changes yields relative to the household benchmark.

I first present the portfolio-pricing framework. A two-investor, two-market affine benchmark then isolates the mechanism. The two propositions below summarize the return-gap decomposition and rate-risk transmission, with formal conditions and proofs in Appendix A. Section 5 extends the framework for calibration.

4.1 Bonds, investors and portfolio choice

Time is discrete. Each investor solves an after-tax mean–variance portfolio choice problem across all assets and maturities, subject to short-sale constraints.

Interest rates. A single rate state z_t drives the short rate r , the pretax return on cash, which is subject to a zero lower bound:

$$z_{t+1} = \rho z_t + \sigma_r \varepsilon_{t+1}, \quad r(z_t) = \max\{\bar{r} + z_t, 0\}, \quad 0 < \rho < 1. \quad (5)$$

Here ε_{t+1} is standard normal, ρ measures persistence, σ_r the size of rate shocks and $\bar{r}$ the mean rate before the floor. Taxable cash earns this known log gross return $r(z_t)$ over the next period.

Bonds and returns. Bonds are zero-coupon and grouped into markets a . Taxes are investor-market specific. For a given investor, all maturities in a market receive the same tax treatment. A bond of market a with T periods to maturity has price $P_T^a(z)$, pays one unit at maturity and becomes a $(T - 1)$ -period bond next period. Cash is the one-period taxable asset. Dollar supplies Q_T^a are fixed.

An investor who buys a T -period bond today and sells it next period earns

$$\text{gross holding return} = 1 + R_{T,t+1}^a = \frac{P_{T-1}^a(z_{t+1})}{P_T^a(z_t)}, \quad P_0^a = 1.$$

The price ratio is the resale value per dollar invested. It carries rate risk because next period's price depends on next period's rate state.

With one period to maturity, the payment is known, so the bond is fully riskless. A zero-coupon bond's yield is $y_T^a = -\log P_T^a/T$. The holding return above covers one period, from purchase to resale next period. Investors optimize over this one-period horizon.

Consistent with the discrete-time return approximations used in the literature, I approximate the net holding return for portfolio choice as⁸

$$R_{T,t+1}^a \approx \mathcal{R}_{T,t+1}^a = \mu_T^a(z_t) + u_{T,t+1}^a. \quad (6)$$

Here $\mu_T^a(z_t) = \log \mathbb{E}_t[1 + R_{T,t+1}^a]$ is the mean of the approximated return used in the portfolio problem. The shock $u_{T,t+1}^a$ is the unexpected component of the log holding return, with conditional mean zero. Thus μ determines expected portfolio income, while u determines risk. The approximation becomes exact in the continuous-time limit.

Investors. Investor i has funding W_i that can be invested in any bonds or cash, risk aversion γ_i and outside rate exposure ℓ_i from assets and liabilities beyond its bond portfolio. A positive ℓ_i raises net wealth when rates rise, as does a long-term liability whose value falls. A negative ℓ_i adds

⁸I follow the discrete-time return approximation in [Greenwood et al. \(2024\)](#), Section 2.1. See also [Greenwood et al. \(2023\)](#), Section III.A, and [Hanson and Stein \(2015\)](#), footnote 18. Appendix [A.1](#) gives the precise return convention and wealth equation.

to bond losses. Investor i keeps the after-tax share $\xi_i^a = 1 - \tau_i^a$ of returns in market a , where τ_i^a is its tax rate on bonds in market a . Cash is taxed like taxable bonds.

The investor chooses dollar positions $x_{i,T}^a \geq 0$ across all markets a and all maturities T , keeps the rest of its funding in cash, $x_i^0 = W_i - \sum_{a,T} x_{i,T}^a \geq 0$, and maximizes expected next-period wealth less a penalty for its variance. With pretax returns measured by $\mathcal{R}$, next-period after-tax wealth is $\widehat{W}_i'$, and the problem is

$$\begin{aligned} \max_{x_i \geq 0, x_i^0 \geq 0} \quad & \mathbb{E}_t[\widehat{W}_i'] - \frac{\gamma_i}{2} \text{Var}_t(\widehat{W}_i'), \\ \text{subject to} \quad & \widehat{W}_i' = W_i + \xi_i^C r(z_t) x_i^0 + \sum_{a,T} \xi_i^a x_{i,T}^a \mathcal{R}_{T,t+1}^a + \ell_i \sigma_r \varepsilon_{t+1}. \end{aligned} \quad (7)$$

Next-period wealth adds after-tax cash income, after-tax bond returns and the change in outside wealth to current funding. Investors choose portfolios each period, taking prices as given and conditioning on today's rate state. Nonnegative positions rule out short sales and borrowing (Section 3.5).

Required returns. The portfolio problem gives the Karush–Kuhn–Tucker conditions

$$\underbrace{\xi_i^a \mu_T^a - \xi_i^C r}_{\text{after-tax return over cash}} \leq \underbrace{\gamma_i \xi_i^a \text{Cov}_z(u_{T,t+1}^a, \widehat{W}_i')}_{\text{cost of the risk the bond adds}} + \underbrace{\kappa_i}_{\text{shorting constraint on cash}}, \quad (8)$$

with equality whenever $x_{i,T}^a > 0$. The covariance measures the risk the bond adds to total wealth, conditional on today's state. The multiplier $\kappa_i \geq 0$ on the shorting constraint on cash is zero when the investor holds cash and can invest another dollar without borrowing. A held bond must compensate for both its added risk and, when this constraint binds, the dollar it displaces from other bonds. If it offers less than an investor's required return in equilibrium, that investor holds none. Differences in taxes and balance sheets therefore allow the same bond to suit one investor but not another in equilibrium.

Equilibrium. Prices clear every market at every rate state:

$$\sum_i m_i x_{i,T}^a(z) = Q_T^a, \quad (9)$$

where m_i is the mass of type i , converting its positions and funding into aggregate holdings $m_i x_{i,T}^a$ and funding $m_i W_i$. In the calibration, one unit of mass represents one dollar of baseline portfolio assets (debt securities and cash). Thus m_i measures a type's portfolio size, and W_i is the share available for the bonds and cash included in the model. The rate process, supplies, taxes, funding, risk aversion and outside exposures are exogenous inputs. Equilibrium prices imply holding returns for which investors satisfy (8) and all markets clear. Who holds each bond is determined by these conditions. Appendix Table 6 collects the notation.

4.2 Taxes and the cost of bearing rate risk

To isolate the mechanism, I use a special case with two investors and two markets. Consider a high-tax household H and a lower-tax investor I choosing municipal bonds M , taxable bonds C and cash. Municipal maturities begin at $T = 1$, while risky taxable bonds begin at $T = 2$. The two determinants of required returns are the investor's tax treatment and the cost of adding risk to its portfolio.

The relative after-tax share of Section 3.3 summarizes the tax comparison:

$$\xi_i = \frac{\xi_i^C}{\xi_i^M} = \frac{1 - \tau_i^C}{1 - \tau_i^M}, \quad \xi_H < \xi_I.$$

The household's smaller ξ_H means a larger exemption benefit, $1 - \xi_H$. I measure each investor's payoffs in units of its after-tax municipal income. Municipal returns are then kept in full, and taxable returns and cash income are multiplied by ξ_i . This change of units preserves portfolio choices.⁹

In the full model, prices are nonlinear in the rate state for three reasons: the short rate cannot fall below zero, the shorting constraint on cash can bind, and which investors hold which bonds can change as rates move. A bond's price today depends on its resale price next period, so these future nonlinearities also affect today's prices. For tractability, the illustration fixes participation: prices are computed as if each bond were held by the same investors at every future rate state, for example the household alone up to some maturity and both investors beyond it. I call such a fixed holding pattern a participation branch. The illustration also removes the rate floor, setting $r = \bar{r} + z$, and the shorting constraint on cash, so $\kappa_i = 0$. Each log price is then a constant plus a constant multiple of the rate state, $\log P_T^a(z) = -A_T^a z - C_T^a$, where A_T^a is the bond's rate loading defined below. The quantitative model of Section 5 imposes none of these simplifications: participation changes with rates, and the rate floor and the shorting constraint on cash are kept. Fixing participation abstracts from the possibility that rate movements over a bond's remaining life change who holds it, and hence how strongly its price responds to rates (Appendix equation (48)).

Define a bond's rate loading as $A_T^a \equiv -\partial \log P_T^a / \partial z$, which measures its proportional price decline when the rate state rises. It is determined by equilibrium prices, as Section 4.5 shows. Rate shocks affect the holding return through next period's resale price, when the bond has $T - 1$ periods remaining. Define the investor's net rate exposure as

$$\widehat{D}_i \equiv -\frac{\partial \widehat{W}_i}{\partial z_{t+1}} = \sum_T x_{i,T}^M A_{T-1}^M + \xi_i \sum_T x_{i,T}^C A_{T-1}^C - \ell_i. \quad (10)$$

The derivative measures how much next-period wealth falls when the rate state rises, holding today's portfolio fixed. Liabilities whose value also falls offset losses on bonds, so a larger liability hedge ℓ_i lowers this net rate exposure.

⁹With payoffs divided by ξ_i^M , risk aversion $\gamma_i \xi_i^M$ and outside exposure ℓ_i / ξ_i^M , the objective is $1 / \xi_i^M$ times the original, up to a portfolio-independent constant. Below, γ_i and ℓ_i denote these normalized values.

The investor's cost per unit of additional rate exposure is $\gamma_i \sigma_r^2 \widehat{D}_i$, which falls with the liability hedge at given bond holdings. The portfolio first-order conditions are

$$\begin{aligned}\mu_T^M - \xi_i r &\leq \gamma_i \sigma_r^2 \widehat{D}_i A_{T-1}^M, \\ \mu_T^C - r &\leq \gamma_i \sigma_r^2 \widehat{D}_i A_{T-1}^C,\end{aligned}\tag{11}$$

with equality for a positive position. The municipal condition compares μ_T^M with after-tax cash, $\xi_i r$. The taxable condition, divided by ξ_i , is on a pretax scale and compares μ_T^C with r . In both lines the right side is the return the investor requires for the rate risk added by the bond. The quantitative model uses common risk aversion in these normalized payoff units, so its investors differ in tax treatment ξ_i and net rate exposure $\widehat{D}_i$, which depends on portfolio choice and outside exposures. The analytical benchmark allows γ_i to differ.

These conditions separate the two forces behind participation. The household's larger exemption benefit lowers its required municipal return. A lower unit cost of rate risk also lowers the required return, increasingly so for a bond with a larger loading.

4.3 Participation at longer maturities

At the riskless maturity, only the cash alternative matters. With $A_0^M = 0$, a municipal holder requires $\mu_1^M = \xi_i r$. The household accepts the lowest return because it keeps the smallest share of taxable interest. With a positive short rate and a slack shorting constraint on cash, it holds the municipal supply and

$$y_1^M = \mu_1^M = \xi_H r, \quad y_1^C = r.\tag{12}$$

At the riskless short end the municipal yield equals the household benchmark exactly.

At a risky maturity, compare the returns the two investors require on the same municipal bond. The lower-tax investor requires less when its duration benefit exceeds the difference in exemption benefits:

$$\underbrace{\sigma_r^2 A_{T-1}^M (\gamma_H \widehat{D}_H - \gamma_I \widehat{D}_I)}_{\text{duration benefit}} > \underbrace{(\xi_I - \xi_H) r}_{\text{difference in exemption benefits}}.\tag{13}$$

All quantities are evaluated at equilibrium. A positive $\gamma_H \widehat{D}_H - \gamma_I \widehat{D}_I$ means that the lower-tax investor requires less return per unit of rate exposure. With the same γ_i , this happens when its outside liability exposure ℓ_I is large enough to give it a lower equilibrium $\widehat{D}_I$. Where the duration benefit is smaller than the exemption difference, the household requires less return and holds the bond. Where it is larger, the lower-tax investor requires less and holds the bond. Both can hold a maturity at which the two terms are equal. Thus participation is determined by the comparison at the equilibrium allocation, including the risk each investor bears.

In the illustration, the municipal rate loading A_{T-1}^M is endogenous and increases with maturity (Section 4.5). At a given equilibrium, the difference in unit risk costs is common across maturities. The duration benefit therefore grows with maturity, while $(\xi_I - \xi_H) r$ stays fixed. Lower-tax par-

ticipation can begin at the long end. For taxable bonds, bond returns and cash income receive the same tax treatment, so there is no difference in exemption benefits to overcome. The investor with lower $\gamma_i \hat{D}_i$ requires less compensation for rate risk at every risky maturity.

Figure 4 shows the resulting allocation. The household holds municipal maturities 1 to 22, both investors hold maturity $T^* = 23$, and the lower-tax investor holds maturities 24 to 30 and all risky taxable bonds.

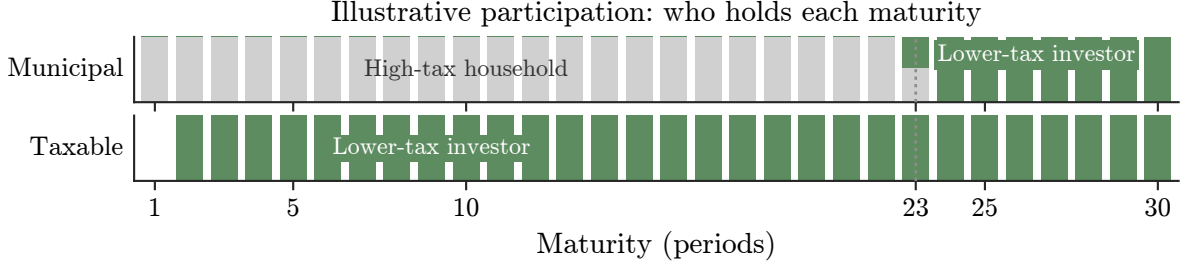


Figure 4: Participation in the two-investor illustration. Bars show supply shares at the mean rate state. Both investors hold municipal maturity $T^* = 23$ (dotted line). Households hold shorter municipals and the lower-tax investor holds longer municipals and risky taxable bonds. The one-period taxable claim is cash. Appendix A.9 gives the parameters and checks.

4.4 Three channels in required returns

Given this allocation, I compare municipal and taxable bonds of the same maturity, the relevant comparison for the municipal puzzle. The municipal return above the household's after-tax taxable return is

$$g_T = \mu_T^M - \xi_H \mu_T^C. \quad (14)$$

This compares municipal holding returns with the household-scaled taxable holding return. The portfolio conditions separate the gap into three channels. Section 4.5 then translates holding returns into yields, including the convexity adjustment.

Proposition 1 (Three channels of the one-period return gap). In the two-investor affine allocation just described, the lower-tax investor holds risky taxable bonds. For municipal bonds, use the household's condition ($i = H$) below T^* and the lower-tax investor's condition ($i = I$) from T^* onward. Subtracting ξ_H times the taxable holder's condition from the municipal holder's condition in (11) gives

$$g_T = \underbrace{(\xi_i - \xi_H)r}_{\text{exemption benefit}} + \underbrace{\sigma_r^2 A_{T-1}^M (\gamma_i \hat{D}_i - \gamma_I \hat{D}_I)}_{\text{duration cost}} + \underbrace{\gamma_I \sigma_r^2 \hat{D}_I (A_{T-1}^M - \xi_H A_{T-1}^C)}_{\text{rate loading}}. \quad (15)$$

The three channels compare exemption benefits, costs per unit of rate exposure, and the exposures themselves. The exemption-benefit term is zero below the threshold, the duration-cost term is zero from the threshold onward, and the loading term applies to both ranges. A zero duration-

cost term does not mean that the lower-tax investor bears no rate risk. It means that the same investor's unit risk cost applies to both bonds in this comparison. The appendix states and proves the decomposition.

Exemption benefit. At long maturities, the lower-tax investor holds both municipal and taxable bonds. Because it pays less tax on taxable interest, tax exemption saves it less than it saves the high-tax household. It therefore requires a higher municipal return than the household benchmark predicts. This difference in exemption benefits, $(\xi_I - \xi_H)r$, contributes to this gap.

Duration cost. At shorter risky maturities, the household alone holds municipal bonds, while the lower-tax investor holds taxable bonds. The household needs more return to bear the same rate exposure because rate risk is more costly for it. This difference contributes $\sigma_r^2 A_{T-1}^M (\gamma_H \hat{D}_H - \gamma_I \hat{D}_I)$ to municipal returns above the household benchmark, even where the lower-tax investor holds no municipal bonds.

Rate loading. When taxes scale returns, they also scale risk. The household benchmark scales taxable yields and rate exposure by ξ_H , predicting $A_{T-1}^M = \xi_H A_{T-1}^C$. In the illustration, lower-tax participation raises municipal rate loadings to $A_{T-1}^M \geq \xi_H A_{T-1}^C$. The extra exposure contributes positively to the gap when the lower-tax investor's net rate exposure is positive, as in the illustration. The contribution equals the extra exposure times $\gamma_I \sigma_r^2 \hat{D}_I$, the lower-tax investor's unit risk cost. Section 4.5 derives the extra rate loading $A_{T-1}^M - \xi_H A_{T-1}^C$.

All three channels vanish at the riskless maturity, and $g_1 = 0$. Since $A_0^M = A_0^C = 0$, the duration-cost and rate-loading terms vanish. The exemption-benefit term is also zero because the household holds the riskless municipal bond. Figure 5 shows how the channels combine across maturities. At a jointly held maturity, either holder's condition gives the same total gap. The choice changes how the gap is divided between the exemption-benefit and duration-cost terms.

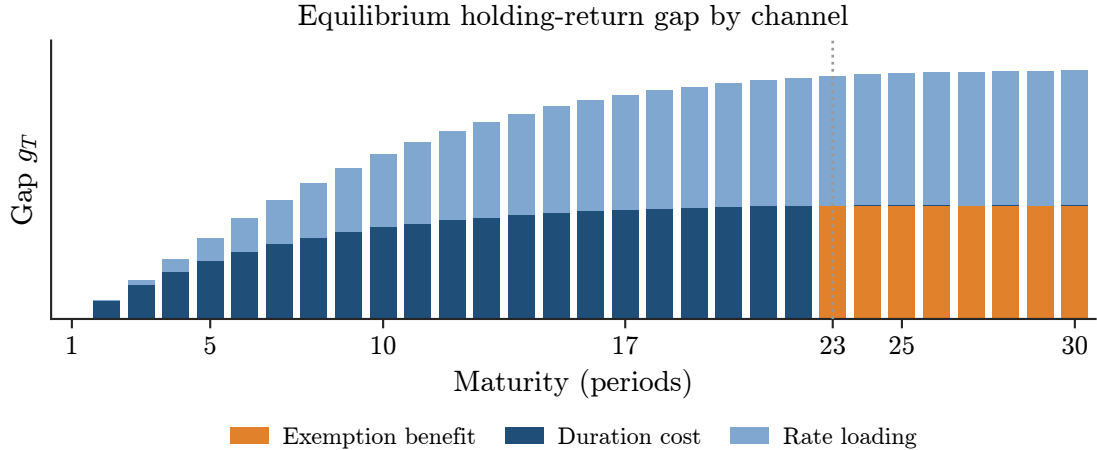


Figure 5: Required-return gap by channel. The allocation is that of Figure 4. Bars decompose $g_T = \mu_T^M - \xi_H \mu_T^C$ into the exemption-benefit, duration-cost and total rate-loading terms of (15). The vertical scale is omitted because levels depend on the illustrative parameters.

4.5 Rate loadings and the yield gap

A bond's rate loading depends on how its expected holding return responds to rates. Let $s_T^a = \partial\mu_T^a/\partial z$ denote this response. The cash terms in the household's portfolio conditions contribute responses ξ_H for municipal bonds and 1 for taxable bonds. Changes in the holder's portfolio exposure also change the risk compensation it requires. Both the holder's tax treatment and its portfolio response therefore enter the bond's return response.

As rates rise, the difference in exemption benefits, $(\xi_I - \xi_H)r$, grows, making municipal bonds relatively less attractive to the lower-tax investor. Trading at the shared maturity transfers rate exposure to the household. The household then requires more compensation on all its risky municipal bonds, including those it holds alone. The lower-tax investor sheds exposure and requires less risk compensation on its municipal and taxable bonds. Portfolio substitution thus transmits the effect beyond the bond that is traded.

Proposition 2 (Participation and rate responses). On a feasible affine participation branch, let $\bar{\xi}_T$ denote the municipal holders' tax component. It equals ξ_H when the household holds the bond alone, ξ_I when the lower-tax investor holds it alone, and $\xi^* = \omega_H \xi_H + \omega_I \xi_I$ under joint holdings. Here ω_i is each investor's share of total risk tolerance.¹⁰ Differentiating (11) gives

$$s_T^M = \begin{cases} \xi_H + \gamma_H \sigma_r^2 A_{T-1}^M \frac{\partial \hat{D}_H}{\partial z}, & \text{household alone,} \\ \xi^*, & \text{both investors,} \\ \xi_I + \gamma_I \sigma_r^2 A_{T-1}^M \frac{\partial \hat{D}_I}{\partial z}, & \text{lower-tax investor alone,} \end{cases} \quad (16)$$

$$s_T^C = 1 + \gamma_I \sigma_r^2 A_{T-1}^C \frac{\partial \hat{D}_I}{\partial z}.$$

Here $\partial \hat{D}_i / \partial z$ measures how the investor's net rate exposure changes as rates rise. On the illustrated allocation, $\partial \hat{D}_H / \partial z > 0 > \partial \hat{D}_I / \partial z$. The municipal portfolio response is therefore positive at household-only maturities, zero at the shared maturity and negative where the lower-tax investor holds municipals alone. A negative portfolio response means that risk compensation falls as the investor sheds exposure, not that the compensation itself is negative. At the longest illustrated maturity, the total municipal response is 0.690, equal to the lower-tax share 0.780 less a portfolio offset of 0.090. It remains above the household benchmark, 0.550.

Combining the tax component and municipal portfolio response separates the net effects in the two markets:

$$s_T^M - \xi_H s_T^C = \underbrace{s_T^M - \xi_H}_{\text{municipal side}} + \underbrace{\xi_H(1 - s_T^C)}_{\text{taxable side}}.$$

The municipal term measures how much more the municipal return responds than the household's after-tax cash alternative. The taxable term reflects the lower-tax investor's falling risk compensation in the taxable market. Both are positive at risky maturities in the illustration. The appendix

¹⁰The weights are $\omega_i = (m_i/\gamma_i)/(m_H/\gamma_H + m_I/\gamma_I)$. They sum to one.

proves the response formulas and the increasing relative sensitivity on this branch. Figure 6 shows the net effect in each market, combining the municipal tax and portfolio terms.

Because $\mu_T^a(z)$ is the log of the expected gross holding return, it links expected holding returns to prices. Rearranging its definition gives the log-price recursion

$$\log P_T^a(z) = -\mu_T^a(z) + \log \mathbb{E}_z[P_{T-1}^a(z')].$$

Today's log price equals the log expected resale value less the required holding return. With affine log prices and $\mathbb{E}_z[z'] = \rho z$, the log expected resale value has loading ρA_{T-1}^a . Taking the negative state derivative gives

$$A_T^a = \rho A_{T-1}^a + s_T^a, \quad A_0^a = 0. \quad (17)$$

The first term carries forward the shorter bond's loading, weighted by rate persistence, and the second adds the current return response. Starting from the riskless holding return, this relation carries the response differences into $A_T^M > \xi_H A_T^C$ for $T \geq 2$. Municipal bonds consequently carry more rate exposure than the household benchmark implies. Figure 6(b) accumulates the municipal and taxable contributions into this loading gap.

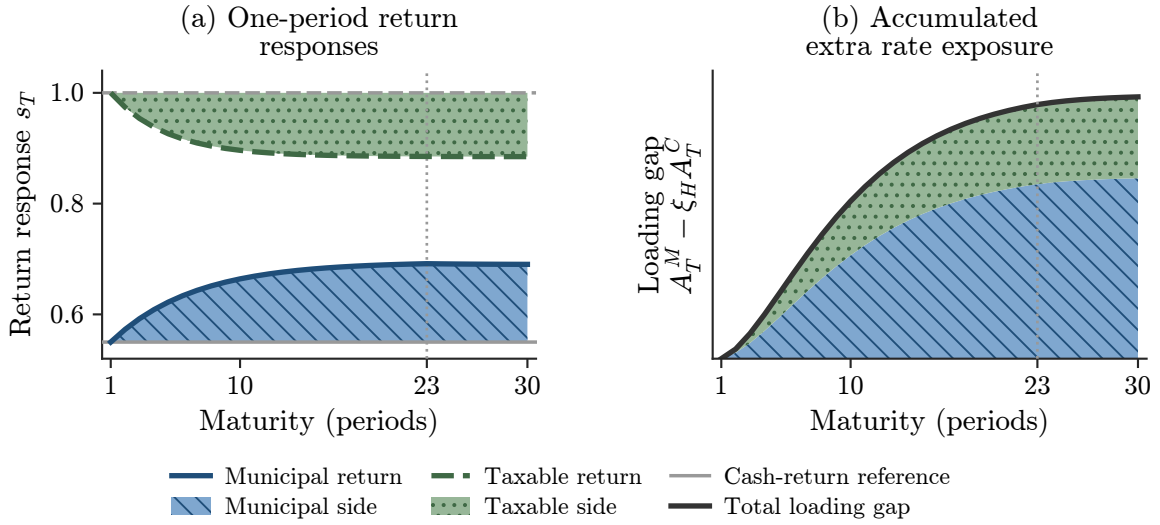


Figure 6: Rate responses in the illustration of Figure 4. Panel (a) shows municipal and taxable holding-return responses. Gray lines show the household's after-tax cash response, ξ_H , and the pretax cash response, 1. Blue shading combines the municipal tax component and portfolio response into $s_T^M - \xi_H$. Green shading shows $1 - s_T^C$. Panel (b) accumulates $s_T^M - \xi_H$ and $\xi_H(1 - s_T^C)$ into the loading gap. The dotted line marks the shared municipal maturity $T^* = 23$.

From return gaps to yield gaps. A long bond earns a sequence of holding returns before it matures. Its yield averages the expected returns along this path, less convexity (Appendix equa-

tion (43)). Applying this relation to both markets gives

$$y_T^M - \xi_H y_T^C = \frac{1}{T} \sum_{j=0}^{T-1} \mathbb{E}_t[g_{T-j}(z_{t+j})] - \text{convexity difference.} \quad (18)$$

The yield gap accumulates the three return-gap channels over the bond's remaining life. In the illustration, the equilibrium municipal yield exceeds the household benchmark at every maturity beyond one period.

4.6 Empirical predictions

The representative-tax benchmark predicts full pass-through of household tax changes at every maturity and no response to tax benefits specific to lower-tax institutions. In the model, household taxes determine the riskless short municipal yield, while longer yields also reflect lower-tax participants' tax treatment and portfolio responses. I first use tax variation specific to households to test whether household tax pass-through declines with maturity (P1). I then use tax variation specific to banks and insurers to test whether lower-tax institutions' benefits affect longer yields while leaving the riskless short yield unchanged (P2). Together, these tests distinguish the participation mechanism from a single household tax rate applied across the yield curve. P3 tests the associated prediction that municipal rate sensitivity rises relative to the taxable benchmark with maturity.

The outcome of interest is the yield-implied tax $\tau_T^{Y,\text{impl}} = 1 - y_T^M / y_T^C$ from (2). Investor i 's pass-through measures how much this equilibrium implied tax changes when investor i 's exemption benefit changes:

$$\mathcal{P}_i^Y(T) = \frac{d\tau_T^{Y,\text{impl}}}{d(1 - \xi_i)}, \quad i \in \{H, I\}. \quad (19)$$

Here, pass-through of one means full transmission and zero means no effect. The representative-tax benchmark predicts pass-through of one for household taxes by definition.

In the affine benchmark, the household alone holds the riskless short municipal bond, whose yield equals its after-tax return on taxable cash. Since $y_1^C = r$, this gives $y_1^M = \xi_H r$, so the implied tax at the riskless shortest maturity is $1 - \xi_H$. Therefore

$$y_1^M = \xi_H y_1^C \implies \mathcal{P}_H^Y(1) = 1, \quad \mathcal{P}_I^Y(1) = 0. \quad (20)$$

Only the household's benefit enters this pricing relation.

At longer maturities, lower-tax participation changes whose exemption benefit affects returns. A larger benefit makes municipal bonds more attractive to its recipient and changes portfolios in both markets. The household's benefit therefore no longer has the exclusive role it has at the riskless short end. An institution's tax treatment affects returns where it holds bonds, while portfolio substitution transmits the effect to other risky maturities. A long bond's yield also reflects the

returns earned as its remaining maturity shortens. Its yield can therefore respond to household taxes even when an institution holds it today.

I compute tax pass-through by changing one investor's after-tax share and re-solving loadings and portfolios, keeping non-tax primitives fixed. In the two-investor illustration, household pass-through declines from 1 at the shortest maturity to 0.242 at maturity 30, while institutional pass-through rises from 0 to 0.721. These are full yield-ratio derivatives, including the response of the taxable curve. The long-end continuation of the illustrated price recursions has the same ordering:

$$\mathcal{P}_H^Y(\infty) < 1, \quad \mathcal{P}_I^Y(\infty) > 0. \quad (21)$$

Appendix A.9 derives the loading result and reports the re-solved tax derivatives. Appendix Proposition A.9 gives sufficient conditions for the yield-pass-through inequalities in the extension with joint holdings across long maturities. The empirical analysis tests the following three predictions, P1 to P3.

Household taxes (P1). A home-state exemption gives resident households an additional benefit on their state's bonds. That benefit is fully reflected in implied state taxes at the riskless short end and less strongly at longer maturities in the solved model. The representative-tax benchmark predicts full pass-through at every maturity.

Institutional taxes (P2). A benefit specific to a lower-tax institution has no effect at the riskless short end and matters more at longer maturities. It affects returns where the institution holds bonds and spreads to other risky maturities through portfolio substitution. Bank qualification and the 2017 corporate tax change for insurers test this prediction. For banks, the measure is the share of the bank-only benefit reflected in yields (Section 6.2). The representative-tax benchmark gives these benefits no role.

Interest-rate responses (P3). Section 4.5 shows why municipal bonds carry more rate exposure than the household benchmark predicts. Since $y_T^a = -\log P_T^a/T$, yield sensitivity is A_T^a/T . Comparing bonds at the same maturity cancels this division by T . For the illustration's affine branch, this gives

$$\frac{\partial y_T^M / \partial z}{\partial y_T^C / \partial z} = \frac{A_T^M}{A_T^C}, \quad \frac{A_1^M}{A_1^C} = \xi_H, \quad \lim_{T \rightarrow \infty} \frac{A_T^M}{A_T^C} = \xi_I > \xi_H.$$

The lower-tax investor prices both markets at long maturities. Its changing rate exposure affects both loading recursions through the same risk-compensation coefficient in (16). This common adjustment cancels in the limiting ratio, leaving ξ_I (Appendix A.9). It still affects each bond's loading and return response. On the illustrated branch, the ratio rises with maturity from ξ_H toward ξ_I . P3 tests whether municipal yield sensitivity relative to same-maturity taxable bonds increases with maturity. The representative-tax benchmark predicts the constant ratio ξ_H .

4.7 Liability hedging and other motives

Liability hedging is one source of the participation advantage. Lower trading costs or better default diversification can also make municipal bonds less costly for a lower-tax investor to hold. An investor may trade less often or at lower cost than households, or diversify default risk across markets more effectively than a household concentrated in its home state's bonds. When these advantages grow with maturity, they can outweigh the maturity-invariant difference in exemption benefits and sustain participation in long municipal bonds.

The pricing mechanism extends to these motives when the relevant advantage is neutral across taxable and tax-exempt bonds. Neutrality means that the advantage itself creates no return gap relative to the household benchmark when the household holds both bonds. Lower-tax participation then brings the participant's smaller exemption benefit into the return comparison. Appendix A.11 states the result, and Appendix A.11.4 gives solved trading-cost and default-diversification examples, with the default comparison net of default compensation. A preference for municipal bonds as such, including warm glow from financing local projects, lowers the return an investor requires on them. Such a preference alone does not explain yields above the household benchmark.

I use liability hedging in the quantitative application because rate risk is common to both markets and balance sheets provide a measurable source of differences in investors' risk-bearing costs. Observed differences in maturity holdings help discipline the outside exposures. The trading-cost and diversification examples show that liability hedging is not the only motive that can sustain lower-tax holdings and bring a smaller exemption benefit into relative returns.

5 Quantitative Results

5.1 The quantitative model

The quantitative model asks how much of the municipal puzzle measured taxes, balance sheets and ownership can explain. It extends Section 4.1 to municipal markets across the 50 U.S. states and the District of Columbia, a taxable market and six investor groups. Tax rules, quantities and household balance sheets are measured directly. Seventeen common coefficients governing risk aversion, bond-specific risk, outside exposures and within-group asset shares are estimated from twenty-two moment conditions on ownership, maturities, portfolios and the taxable yield curve. Municipal yields are untargeted.

States with a personal income tax have separate municipal markets. The seven states with no state income tax are combined into one market because investors treat their bonds alike in the model. The model assigns the home-state exemption to resident households, while institutions have no modeled state-income-tax wedge. A model period is two years, giving fifteen maturities from two to thirty years in each of the 45 municipal markets. The taxable market pools treasury and corporate supply. Cash and the two-year municipal claim are riskless. Prices in both markets are determined in equilibrium.

For calibration, I add a small independent bond-specific return shock, which makes bonds imperfect substitutes and permits several investor types to share a maturity. It vanishes at the riskless claim and has a small effect on the yield gap (Appendix B.8). Investors solve (7) with market-specific after-tax shares, and prices satisfy the covariance-based portfolio conditions (8) and clearing (9) in every market and rate state. Banks’ measured financing-tax cost is subtracted from municipal returns according to their tax treatment (Appendix B.7). The model retains the rate floor in (5) and allows the shorting constraint on cash to bind and participation to change across rate states. I solve $P_T^a(z)$ numerically through the general resale relation in Appendix equation (39), integrating over next-period states. Returns, portfolio risks, holdings and both yield curves are determined jointly. Prices can be nonlinear in the rate state because future participation, binding shorting constraints on cash and the rate floor affect resale values.

The measured inputs cover six three-year blocks, from 2000–2002 through 2015–2017, centered on Survey of Consumer Finances (SCF) waves. Financial Accounts stocks give each sector’s portfolio assets and bond quantities. Mergent gives maturity profiles and municipal state shares. The SCF supplies household balance sheets and the income and filing status needed to apply statutory taxes. Historical short rates determine the rate process. Appendix B.7 details the construction, and Table 14 lists the sources.

Table 3 lists the groups. For type i , m_i is its baseline portfolio assets (debt securities and cash), and W_i is the share available for the asset categories traded in the model. Its investment budget is therefore $m_i W_i$. Positions $x_{i,T}^a$ are measured per dollar of portfolio assets, so aggregate bond holdings are $m_i x_{i,T}^a$. Households differ by federal tax bracket and outside assets and debt. Outside assets add rate risk and liabilities offset it. Their mapping into exposures is estimated. Insurers differ in exposures and shares of sector assets, banks face municipal financing-tax costs, and the residual group is tax-neutral. Section 3.4 and Appendix D.1 distinguish gross exemption benefits from these financing costs.

5.2 Estimation and fit

I measure directly the inputs available in the data: investor tax treatment, each sector’s portfolio assets, bond quantities and maturity profiles, household assets and debt, and the short-rate process. These inputs vary across the six blocks of 2000–2017. The remaining seventeen coefficients govern risk aversion, bond-specific risk, outside rate exposures and within-group asset shares. I estimate one common set of coefficients across all six time blocks.

Observed ownership provides the main discipline on outside rate exposure, which balance-sheet amounts alone cannot measure. Given each group’s tax treatment and portfolio assets, its municipal and taxable holdings reveal how much rate risk it is willing to bear. For households, three coefficients translate measured outside assets and debt into rate exposure. For institutions, the estimated exposures and within-sector asset shares account for both the amounts held and their maturity composition. The taxable yield curve and long-bond allocations discipline risk aversion

Table 3: Investor groups in the quantitative model, 2015–2017

Group	Types	Types differ by	Gross exemption benefit $1 - \xi_i$	Portfolio assets m_i
High-tax households	$7 \times 45 = 315$	Rate exposure, home state	0.43–0.50	6.97
Lower-tax households	$7 \times 45 = 315$	Rate exposure, home state	0.21–0.27	8.09
P&C insurers	2	Rate exposure	0.31	1.11
Life insurers	3	Rate exposure	0.21	3.05
Banks	1	None	0.35	4.83
Residual group	7	Rate exposure	0	26.47
Total	643			50.52

Notes. Each household group has seven rate-exposure types in each of the 45 markets and in the one-market economy used to estimate the coefficients (Section 5.2). Estimated coefficients map each household type’s measured outside assets and debt into exposure. Institutional exposures are estimated directly (Appendix Table 9). The gross exemption benefit $1 - \xi_i$, with $\xi_i = (1 - \tau_i^C)/(1 - \tau_i^M)$, follows from tax law. Households multiply the federal and home-state after-tax shares of each SCF tax unit, with the range running across states, and institutions face the 35% corporate rate, with proration for insurers. Banks also pay a financing-tax cost of 0.11% a year on municipal holdings, which lowers their net benefit below this gross value (Table 2). Portfolio assets m_i , in trillions of dollars, are all of a group’s debt securities and cash in the Financial Accounts, and they scale the group’s risk-bearing capacity.

and bond-specific risk. The coefficients are estimated jointly because portfolio choices and yields adjust together in equilibrium.

The estimation uses twenty-two targets. Nine are ownership stocks: municipal holdings of the two household groups, P&C insurers, life insurers and banks, and taxable holdings of households and the same three institutional sectors. Four are the mean maturities of insurers’ municipal and taxable portfolios. Six describe the distribution of insurer holdings: two measure whether municipal holdings are spread across insurers or concentrated in portfolios with high municipal shares, giving larger portfolios more weight, and four measure the fractions of each sector’s municipal and taxable portfolios maturing beyond twenty years. The remaining three are taxable yields at ten, twenty and thirty years. I evaluate these conditions at observed rate states, average them across the six blocks and minimize their weighted squared deviations from the data.¹¹

For computational tractability, I estimate the coefficients in an economy with one national municipal market, using municipal supply weights to average each household type’s home-state after-tax shares. I then restore the 45 municipal markets and their measured tax differences, holding every coefficient fixed. No estimated coefficient is state-specific.¹²

The one-market and full economies give similar relative municipal yields. Their thirty-year gaps average 102.6 and 102.4 basis points across the six blocks, differing by at most 2.5 basis points in any block. Thirty-year taxable yields differ by at most 0.5 basis points. Changing each coefficient group in turn (Appendix Table 12) moves the one-market gap by at most about 4 basis points,

¹¹Each averaged deviation is scaled by 10% of its observed normalizer. Maturity and long-portfolio conditions include model-portfolio weights. Appendix equation (111) gives the criterion, and Tables 9 and 10 report the coefficients and target definitions.

¹²Fourteen of the twenty-two absolute scaled deviations are below one. Appendix B.6 reports the estimation procedure and fit.

except for the common risk aversion.¹³ Section 5.4 uses the 2000–2020 bond sample’s weights and expresses the model’s relative municipal yields on the observed corporate curve for the headline comparison.

Municipal yields, the municipal puzzle and empirical tax-pass-through coefficients do not enter the estimation criterion. Ownership disciplines outside rate exposure, which combines with observed taxes to determine municipal yields in equilibrium. The relative municipal yield curve and cross-state pass-through therefore test the mechanism without being targeted (Sections 5.3–5.4 and 6.1).¹⁴

5.3 Model and data yield curves

Figure 7 sets the calibrated model curve (top) against the data (bottom). To cover the full data sample, I keep the estimated coefficients fixed and add a seventh economy for 2018–2020 after the 2017 Tax Cuts and Jobs Act (TCJA), built from that period’s taxes, rates and quantities. Both rows show a municipal yield near the household benchmark at the short end and a spread over it that widens with maturity. In nominal semiannual yields, the model’s implied tax $1 - y_T^M / y_T^C$ is 41.0% at two years, 34.7% at ten and 24.8% at thirty, a decline of 16.2 percentage points. The model’s thirty-year municipal and taxable yields are 4.651% and 6.185%, against 4.558% and 5.794% in the adjusted data.

5.4 Accounting for the puzzle

The accounting asks how much of the municipal puzzle the model explains, alongside the measured capital-gains, liquidity and default differentials. To compare these contributions, I express them on the same observed taxable yield curve. Otherwise, a higher taxable yield in the model would mechanically enlarge its gap in basis points, even with the same implied tax.

At each maturity, I keep the model’s municipal-to-taxable yield ratio and multiply it by the observed taxable yield. I then subtract the household after-tax benchmark. The resulting gap, denoted $\tilde{\mathcal{G}}_T$, measures the model’s contribution on the data’s taxable curve:

$$\begin{aligned} \text{Normalized municipal yield} &= \frac{\text{Model municipal yield}}{\text{Model taxable yield}} \times \text{Observed taxable yield}, \\ \tilde{\mathcal{G}}_T &= \text{Normalized municipal yield} - \text{Household benchmark}. \end{aligned} \tag{22}$$

I call this the corporate normalization. It preserves the model’s implied tax and changes only the scale of this comparison. The solved equilibrium and the curves in Figure 7 are unchanged. Appendix C.1.7 gives the averaging and normalization details.

¹³Changing risk aversion alone scales both curves and the spread, moving the taxable yield level away from its calibration targets.

¹⁴Appendix B describes the solution method, and Appendix B.3 reports accuracy.

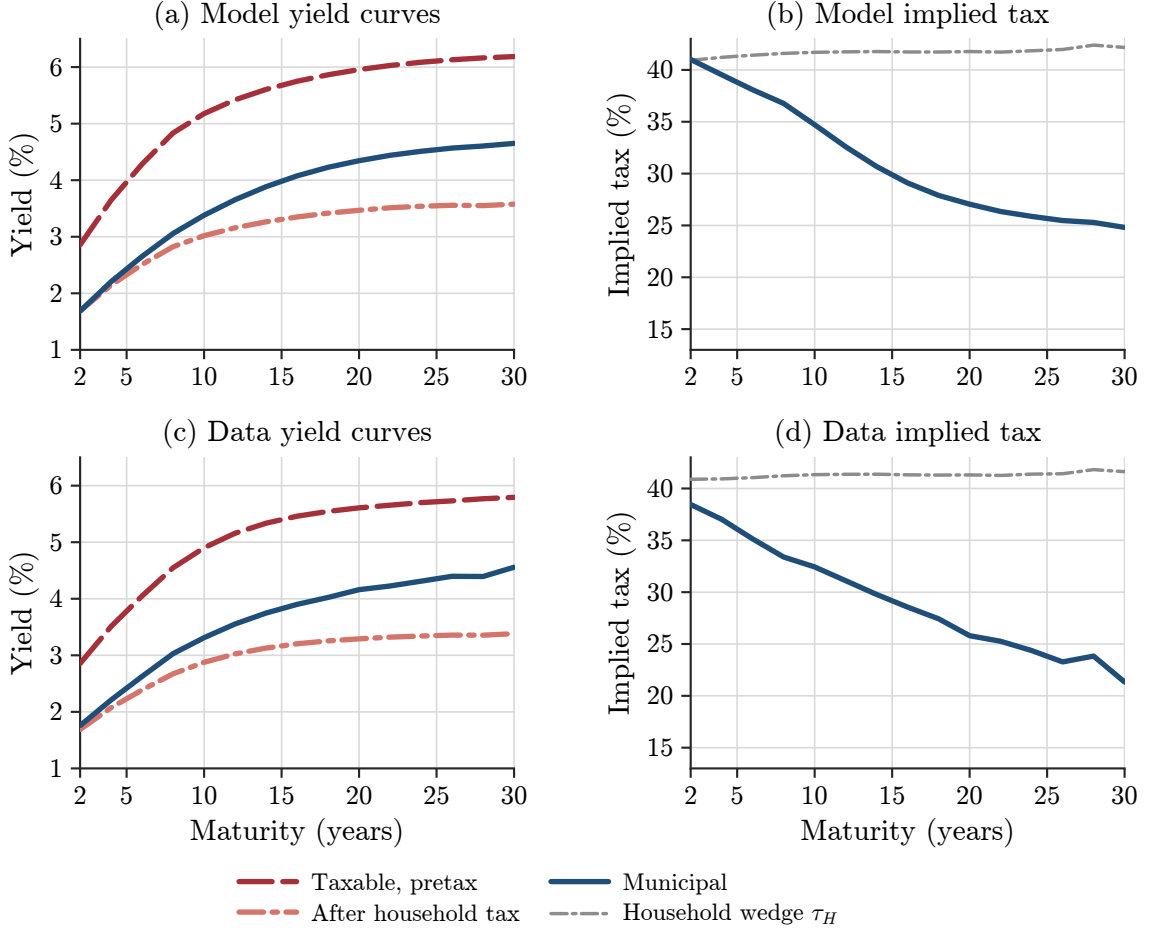


Figure 7: Model (top) and adjusted data (bottom), 2000–2020. Panels (a) and (c) plot the pretax taxable yield, the taxable yield after household tax (the household benchmark) and the municipal yield. Panels (b) and (d) plot the implied tax $1 - y_T^M/y_T^C$ and the household wedge. Corresponding panels share axis scales. The taxable curve is the model’s pooled taxable market (top) and the observed corporate benchmark (bottom). The municipal data curve removes callability (Appendix C.1.2) and the issuance premium, and panel (d) uses this same adjusted curve. The model evaluates the estimated coefficients at observed issue-quarter rate states in the six calibration blocks and a 2018–2020 economy. Both rows use the empirical bond, rating and quarter weights and nominal semiannual yields. The household benchmark multiplies each observation’s taxable yield by the high-tax household’s relative after-tax share for that bond’s state and date before averaging. Implied taxes and the household wedge are ratios of averaged yields.

The accounting has three steps at each maturity T :

$$\begin{aligned}
 \text{Spread}(T) - \underbrace{[\text{Call}(T) + \text{Prem}(T)]}_{\text{comparability adjustments}} &= \underbrace{\text{Puzzle}(T)}_{\text{municipal puzzle}}, \\
 \text{Puzzle}(T) - \underbrace{[\text{CG}(T) + \text{Liq}(T) + \text{Def}(T)]}_{\text{bond-side adjustments}} &= \underbrace{\text{Puzzle}^{\text{adj}}(T)}_{\text{residual before the model}}, \\
 \text{Puzzle}^{\text{adj}}(T) - \underbrace{\tilde{\mathcal{G}}_T}_{\text{model}} &= \underbrace{\text{Residual}(T)}_{\text{unexplained}}.
 \end{aligned} \tag{23}$$

The spread, call adjustment and issuance premium follow Section 3.7. The terms CG, Liq and Def denote capital-gains, liquidity and default differentials. Since the model excludes these features, its gap is compared with $\text{Puzzle}^{\text{adj}}(T)$. The final residual equals the observed corporate yield times the implied-tax difference (model minus data), with the data reflecting all five adjustments.

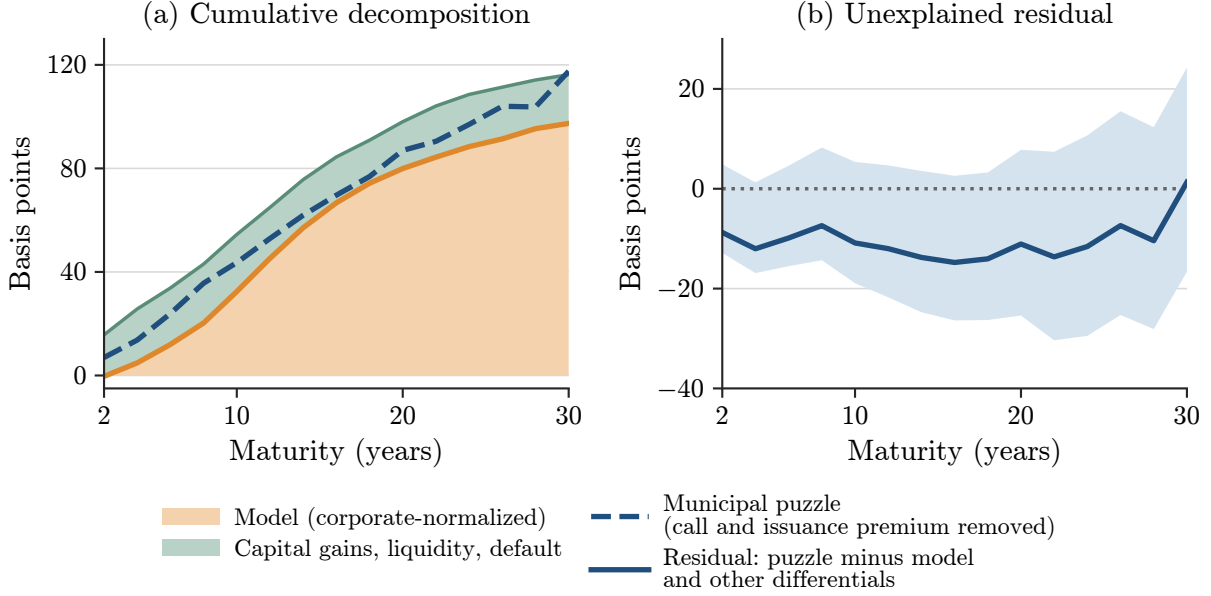


Figure 8: Cumulative decomposition of the municipal puzzle, 2000–2020, as in (23). The puzzle is the observed spread after removing callability and the issuance premium. Panel (a) stacks the model gap $\tilde{G}_T$ of (22) and the capital-gains, liquidity and default differentials against that puzzle. Panel (b) shows $\text{Residual}(T)$, the puzzle less their sum. Shading spans the ranges of the five empirical adjustments (Appendix Table 26) and excludes model uncertainty. Weights and yield conventions match Figure 7.

At thirty years (Figure 8 and Table 4), the model accounts for 97 basis points of the 118-basis-point puzzle, or 83%, and the bond-side adjustments account for 19, or 16%, leaving about 1 basis point unexplained. The model’s contribution grows with maturity in basis points. Its share of the puzzle is 74% at ten years, 97% at eighteen and 83% at thirty. The unexplained residual is negative at every maturity except thirty years. The model gap exceeds $\text{Puzzle}^{\text{adj}}(T)$ by 11 to 15 basis points from twelve to twenty-four years, and the largest absolute residual is 15 basis points at sixteen years. This negative residual is consistent with the convenience premium interpretation in Fleckenstein and Longstaff (2025).

Table 4: Accounting for the municipal puzzle. Removing callability and the issuance premium from the raw spread gives the puzzle. The model contribution applies its implied tax to observed corporate yields, as in equation (22).

Maturity (years)	Raw spread	Comparability		Puzzle	Accounting for the puzzle			
		Call	Issuance premium		Model	Model share	Bond-side adjustments	Unexplained
2	9.9	0.1	2.9	6.9	−0.3		16.0	−8.8
6	27.9	0.7	3.3	23.9	12.0	50%	21.8	−9.9
10	52.2	4.6	4.0	43.6	32.4	74%	22.0	−10.9
14	87.0	17.9	7.2	61.9	57.1	92%	18.6	−13.8
18	110.3	23.9	9.6	76.9	74.3	97%	16.6	−14.0
22	128.8	27.0	11.4	90.5	84.3	93%	19.8	−13.6
26	145.0	28.5	12.5	104.0	91.4	88%	20.0	−7.4
30	159.8	29.5	12.8	117.5	97.4	83%	18.8	1.3

Notes. Basis points, except maturity and shares. The accounting is puzzle = model + bond-side + unexplained, as in (23). The model column uses the corporate normalization, and the bond-side adjustments sum the capital-gains, liquidity and default differentials. The model share divides its contribution by the puzzle, using unrounded values, and is omitted where the puzzle is below 20 bp. Appendix C.1.7 reports the unnormalized gaps. Model and data both use the 2000–2020 sample, the averaging of Appendix C.1.1 and the nominal semiannual convention. The separately measured adjustments are not jointly estimated. Rows show every second model maturity. Figure 8 shows all fifteen.

6 Empirical Tests

I test the model’s participation mechanism using variation in household taxes, institutional tax benefits and interest rates (Section 4.6). None of the estimates is a calibration target.

P1 uses differences in residents’ state income-tax rates. P2 uses bank qualification and the 2017 corporate tax cut for insurers. These designs compare household and institutional tax effects across maturities.

P3 compares changes in municipal and treasury yields at the same maturity. For traded coupon bonds, I account for the timing of their payments when constructing the benchmark response.

The P1, TCJA and P3 tests use the AA-and-above municipal sample (agency score three or better) to limit credit differences that could affect both yields and trading. The curve accounting of Sections 3 and 5 uses the wider A-or-above sample (score seven or better), which matches the corporate benchmark’s credit band, and Appendix C.12.2 repeats each test on it with the same maturity ordering. The bank-qualified design uses all new issues with rating-by-year fixed effects. Throughout, n_b denotes a bond’s remaining maturity in years.

6.1 P1. State-tax pass-through

P1 asks whether resident households’ state-tax benefit is reflected less in municipal yields at longer maturities. Most states exempt residents’ interest on in-state municipal bonds and tax interest on out-of-state bonds. A household in state s has a state after-tax share of one on in-state municipal interest and $1 - \tau_s^{\text{st}}$ on out-of-state municipal and taxable interest.

State rates move resident households' exemption benefits, a wedge institutions largely do not face. Pensions and tax-deferred accounts owe no current state tax on interest, and insurer premium taxes generally do not distinguish municipal from taxable interest. Funds pass the benefit through according to shareholder residence and portfolio eligibility: single-state funds preserve the home-state exemption, while national funds pass through only the eligible home-state share (Appendix D.3). Under the conditions in Section 4.6, Appendix Proposition A.9 gives household pass-through of one at the riskless boundary and below one in the long-end limit, and Appendix Proposition A.16 gives the corresponding responses in state markets.

The yield ratio y^M/y^C contains both federal and state taxes. Under the pre-2018 federal deduction convention, the after-tax share of taxable interest is $\xi_s^{\text{st}}\xi_t^{\text{fed}}$, where $\xi_s^{\text{st}} = 1 - \tau_s^{\text{st}}$ is the state share and ξ_t^{fed} the federal share, while home-state municipal interest is exempt at both levels. Household indifference under the representative-tax benchmark gives $y_{s,t}^M = \xi_t^{\text{fed}}\xi_s^{\text{st}}y_t^C$. Dividing out the federal share defines the implied state tax

$$\text{ImpliedStateTax}_{s,t}(T) \equiv 1 - \frac{y_{s,t}^M(T)/y_t^C(T)}{\xi_t^{\text{fed}}}. \quad (24)$$

It equals the statutory rate τ_s^{st} under the benchmark, giving pass-through of one. I set ξ_t^{fed} to one minus the dated top federal ordinary-income rate, excluding the net investment income tax: 0.604 in 2000, 0.609 in 2001, 0.614 in 2002, 0.65 in 2003–2012 and 0.604 in 2013–2017. The sample ends before the Tax Cuts and Jobs Act capped state and local tax deductions at \$10,000 from 2018, and Appendix C.7 reports the later period under its tax mapping.

I estimate pass-through separately in five maturity bins $([0, 5), [5, 10), [10, 15), [15, 20)$ and $[20, 30]$ years):

$$\begin{aligned} \text{ImpliedStateTax}_{b,t} = & \sum_{g=1}^5 \beta_g (\text{StateTax}_{s(b),t} \times \mathbf{1}\{n_b \in g\}) + \sum_{g=1}^5 \delta'_g (Z_{b,t} \times \mathbf{1}\{n_b \in g\}) \\ & + \lambda' V_b + \alpha_{t \times n \times callable} + \varepsilon_{b,t}, \end{aligned} \quad (25)$$

where $\text{StateTax}_{s,t}$ is the state's top statutory rate and the outcome (24) uses call-adjusted zero-coupon yields. The coefficient β_g is the change in the implied state tax per percentage point of statutory rate in maturity bin g , and the benchmark gives $\beta_g = 1$ in every bin. The controls $Z_{b,t}$, also interacted with the bins, are the rounded average rating and six lagged state variables: par-weighted outstanding maturity, the maturity of the preceding year's issuance, outstanding par relative to personal income, log income, log preceding-year issuance and the outstanding share in the bond's maturity bin. The stock and issuance variables cover federally exempt bonds, and the stock and income variables are measured at the preceding quarter-end. V_b contains log offering par and security-type indicators, and $\alpha_{t \times n \times callable}$ denotes issue-year-month-maturity-callability fixed effects.

The 2000–2017 sample has 681,838 observations with complete controls. Standard errors are clustered by state, allowing dependence among a state’s bonds across years and maturities. As an IV comparison, I instrument the current statutory rate with its 1990 historical counterpart, following the policy-history approach of Babina et al. (2021). The results are robust to the choice of historical year, and Appendix C.7 reports other vintages with the same controls and sample. Cross-sectional variation in state tax rates identifies pass-through into yields at different maturities. Historical rates address feedback from current municipal yields to current tax policy. The IV interpretation requires that, conditional on the controls, historical rates affect current yields through current tax rates.

Figure 9 compares the OLS estimates of (25) with the full-state model’s cross-state slope. For the model, I compute (24) using its high-tax household’s federal share ξ^{fed} and nominal semiannual yields. At each maturity, I regress this implied tax on the state’s top rate averaged within the three-year block, with an intercept and equal weights across the 43 income-tax states and the District of Columbia. All states and economies are evaluated at the same fixed rate state, and the slopes are averaged equally across the six pre-2018 blocks.

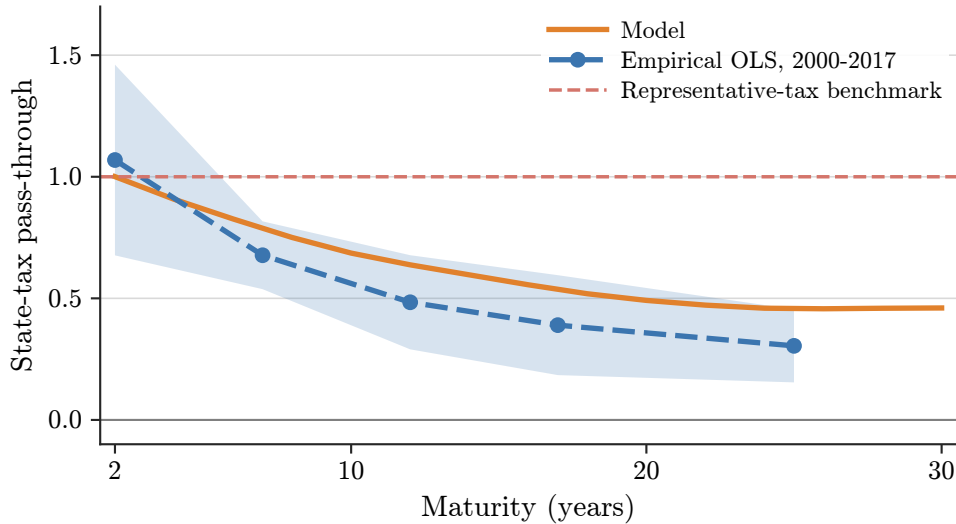


Figure 9: State-tax pass-through by maturity. Blue: OLS estimates for 2000–2017 call-adjusted zeros, with state-clustered 95% intervals using $t(46)$. The regression is unweighted and includes the controls in (25). Orange: the model’s cross-state slope, using its own federal after-tax share and nominal semiannual yields at a common fixed rate state and averaged equally across the six pre-2018 economies. Both use (24). The red dashed line is the representative-tax benchmark.

Estimated pass-through declines across all five bins, from 1.069 below five years to 0.305 beyond twenty (Appendix Table 29). A state top rate ten percentage points higher raises the implied tax by about 10.7 points at the short end and 3.0 points at the long end. The short-to-long decline is 0.764 (s.e. 0.235, $p = 0.002$). Unit pass-through cannot be rejected in the shortest bin and is strongly rejected in the longest, and every bin coefficient is positive and significant at one percent. The IV estimates also decline, from 1.088 to 0.229, with an endpoint-equality p -value below 0.001.

The model's untargeted profile declines from 1 at two years to 0.69 at ten and 0.46 at thirty. The empirical decline survives alternative yield constructions, credit controls, weights, instruments, state fixed effects and single-state omissions (Appendices C.6 and C.7). The appendix also reports the post-2018 period separately under its changed tax mapping.

6.2 P2. Bank qualification

P2 uses a tax benefit that only banks receive. Banks ordinarily cannot deduct interest expense allocated to tax-exempt holdings (Section 265(b) of the Internal Revenue Code). For bank-qualified (BQ) bonds, Sections 265(b)(3) and 291 restore 80% of that deduction. An issuer can designate its bonds as bank-qualified if its reasonably anticipated tax-exempt issuance for the year does not exceed \$10 million (\$30 million in 2009–10 under the American Recovery and Reinvestment Act). Designation changes neither the bonds' payments nor any other investor's taxes. [Dagostino \(2025\)](#) studies how the expanded cap affected issuance and employment.

A bank allocates interest expense across assets in proportion to tax basis, so each dollar of municipal basis carries the same annual funding cost c at every maturity. Designation saves $X = 0.8 \tau_{BC}$ a year, where τ_B is the federal bank tax rate. Measured from Call Report interest expense, X is about 65 basis points in 2005–08 and 9 to 14 from 2011 to 2019 (Appendix C.8.3). Because a bond's benefit accrues over its life, I also scale premia by $\bar{X}_b$, the expected annual benefit over bond b 's life (Appendix C.8.3). Its median is 45 basis points.

Three hypotheses differ in how much of X is reflected in the designation premium $\varpi_k = y_k^N - y_k^Q$, the non-designated minus the designated yield at maturity k . Under the representative-tax benchmark, households hold both bonds at the same yield, $y_k^N = y_k^Q = (1 - \tau_H) y_k^C$. A bank then prefers a taxable bond, which pays $(1 - \tau_B) y_k^C$ after tax, to a designated bond, which pays the household benchmark yield less the tax value of the disallowed fifth of its funding cost, whenever

$$(1 - \tau_B) y_k^C > (1 - \tau_H) y_k^C - 0.2 \tau_{BC} \iff \tau_H y_k^C > \tau_B y_k^C - 0.2 \tau_{BC},$$

that is, whenever the household's exemption benefit exceeds the bank's net exemption benefit on a designated bond, as it does when household rates are at least as high as banks'. Banks then hold neither municipal bond, and $\varpi_k = 0$. Under a bank-only benchmark, banks' indifference to taxable bonds gives full capitalization, $\varpi_k = X$, flat in maturity. Holding outside investment opportunities and other bond characteristics fixed, the premium is the benefit times the bank's capitalization share θ_k , the average bank share of the total demand response to the bond's yield (Appendix C.8):

$$\varpi_k = \theta_k X, \quad \underbrace{\theta_k = 0}_{\text{representative-tax}}, \quad \underbrace{\theta_k = 1}_{\text{bank-only}}, \quad \underbrace{\theta_{\text{short}} \simeq 0 < \theta_{\text{long}}}_{\text{model}}. \quad (26)$$

The capitalization share is the bank counterpart of $\mathcal{P}_I^Y(T)$ in Section 4.6. Household after-tax cash returns anchor short yields and banks hold few short bonds, so θ_k is near zero at short maturities.

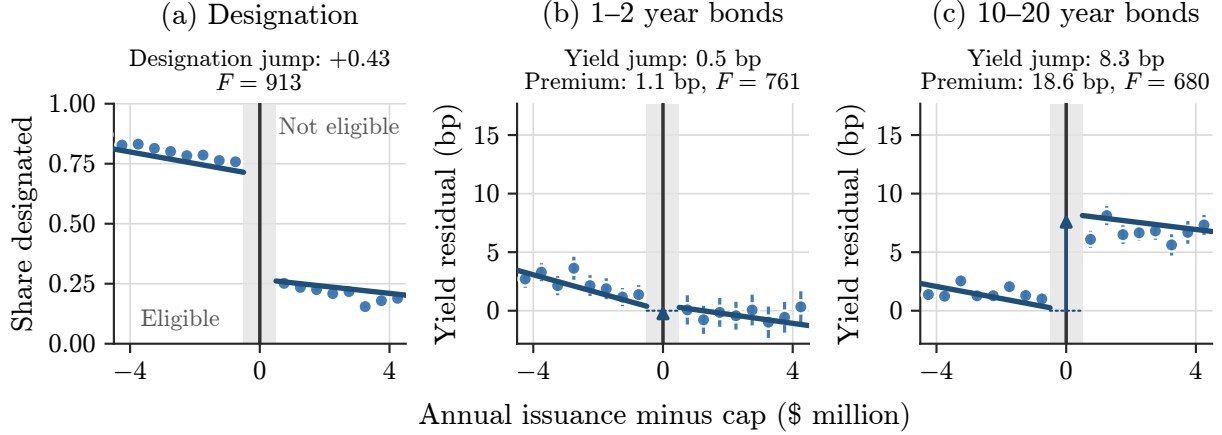


Figure 10: The bank-qualified discontinuity. The horizontal axis is an issuer’s measured annual tax-exempt issuance minus that year’s cap. Designation is more frequent to the left of zero. Dots are means in \$0.5 million bins with 95% intervals: the share of bonds designated in panel (a), and offering-yield residuals after the fixed effects and log par of (27) for 1–2 and 10–20 year bonds in panels (b) and (c). Lines are the estimated local linear fits, which exclude the shaded donut. Yield fits are shifted to zero on the below-cap side at the cap, and the arrow marks the yield jump. Each premium divides the yield jump by its bin’s designation jump. F is the first-stage Wald statistic with issuer-clustered standard errors. Appendix Figure 24 shows the other maturities.

In the calibrated model, banks hold intermediate maturities, where their duration benefit offsets their smaller exemption benefit (Section 4.3), and yields there reflect more of the bank benefit.

The cap provides a comparison among issuers of similar size. Designation averages 0.79 below the cap and 0.21 above it across the estimation window. Measured annual issuance need not equal the issuance anticipated at each sale or counted under the statute, so some designated bonds appear above the measured cap. Below-cap status changes the probability of designation rather than determining it exactly, giving a fuzzy discontinuity. It instruments designation, and the premium is the local yield jump divided by the local designation jump.

The variable R_{jt} is issuer j ’s annual tax-exempt issuance minus that year’s cap, in millions of dollars. The indicator $E_{jt} = \mathbf{1}\{R_{jt} \leq 0\}$ marks below-cap issuer-years, and B_b indicates whether bond b is bank-qualified. The Mergent sample covers tax-exempt new issues from 2000 to early 2021 with $0.5 \leq |R_{jt}| \leq 4$. The half-million exclusion around the cap, the donut, removes managed bunching and issuer-years exactly at the cap. With bins g of maturity at issuance (1–2, 2–5, 5–10, 10–20 and more than 20 years), I estimate

$$y_b = \sum_g \beta_g B_b \mathbf{1}\{n_b \in g\} + W_b' \delta + u_b, \quad (27)$$

by two-stage least squares, instrumenting each designation interaction $B_b \mathbf{1}\{n_b \in g\}$ with the eligibility interaction $E_{jt} \mathbf{1}\{n_b \in g\}$ of the same bin. The first stage is the local designation jump conditional on controls, 0.42 to 0.45 across bins (Appendix Table 35). Here W_b contains local linear controls in R_{jt} on each side of the cap, log tranche par, and fixed effects for maturity, state, offering month, rating, security type, use of proceeds, coupon and callability, with the interactions listed in

Appendix C.8.2. The outcome is the offering yield in basis points, or that yield divided by $\bar{X}_b$, so $-\beta_g$ is the premium in basis points or in expected-benefit units. Standard errors are clustered by issuer.

Identification requires that, given these controls, crossing the cap moves offering yields only through designation. Callability and the number of maturities per issue are balanced at the cap. Placebo cutoffs at \$5 million and \$15 million show essentially no designation effects, and their 10–20 year yield jumps exceed the 1–2 year jumps by only 0.4 and 2.5 basis points, against 7.8 at the cap (Appendix C.8).

Figure 10 illustrates the discontinuity with separate local fits for each maturity bin. The 1–2 year yield barely moves at the cap, by 0.5 basis points, while the 10–20 year yield jumps by 8.3, implying bank-qualification premia of 1.1 and 18.6 basis points after dividing by the designation jumps. The main estimates instead come from the joint 2SLS regression (27), which estimates all five bins together. Figure 11 reports these estimates (Appendix Table 35). The premium is 1.8 basis points at 1–2 years, indistinguishable from zero, and 8.7, 13.8, 18.4 and 18.2 at 2–5, 5–10, 10–20 and more than twenty years. The 10–20 year premium exceeds the 1–2 year premium by 16.5 basis points (s.e. 2.7). In expected-benefit units, the premium rises from 0.17 at 1–2 years to 0.52 at 10–20 years: the bank-only benefit is barely priced in short municipal yields, and about half of it is priced at 10–20 years.

The positive premia at 5–20 years reject the representative-tax benchmark, with t statistics of 6 to 8, and the rise in expected-benefit units, 0.34 (s.e. 0.11), rejects the flat premium of full capitalization. The gradient appears in every rate regime and survives all robustness checks, including alternative donuts and bandwidths, credit-enhancement controls, a general-obligation sample, issuer effects, year clustering and the binned estimator of Dagostino (2025) (Appendix C.8).

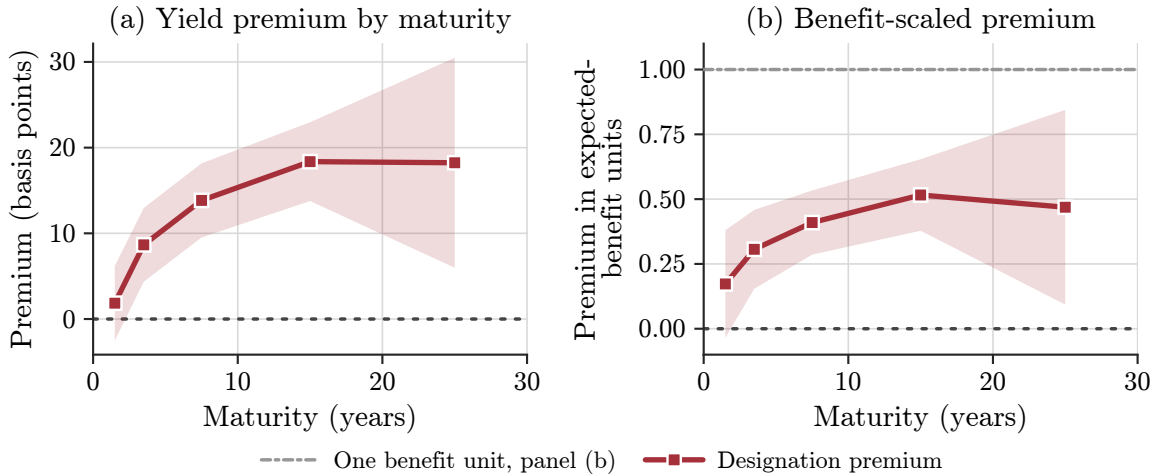


Figure 11: The bank-only benefit and municipal yields. Panel (a): 2SLS yield premium in basis points. Panel (b): the premium in expected-benefit units, with each yield divided by the expected annual benefit $\bar{X}_b$ and the same controls. Funding costs are municipal-holdings-weighted. Expected rates use the Kim–Wright expected-short-rate component, capped at ten years (Appendix C.8). Shading is the issuer-clustered 95% interval. The dashed line marks one benefit unit.

The rising bank profile is the reverse of household pass-through in P1, which falls from 1.07 to 0.30 (Figure 9), as the model predicts for institutional and household tax benefits. The 2017 corporate tax cut provides a second institutional test: it reduced P&C insurers' exemption benefit, and implied taxes then fell more at longer maturities in states where these insurers held more municipal debt. Appendix C.10 presents this second test.

6.3 P3. Rate responses by maturity

P3 asks whether municipal yields respond more to a given treasury yield change at longer maturities. Section 4.6 derives this maturity pattern relative to the model's jointly priced taxable market C . The empirical test uses treasury yields to measure rate movements. The same prediction holds for this reference in the analytical illustration.¹⁵ The representative-tax benchmark instead makes municipal zero yields a constant fraction ξ_H of treasury zero yields, implying a constant relative response. The observed response is about 0.35 at one year, below ξ_H . The test concerns how this relative response varies with maturity.

I compare consecutive trades in the same municipal bond. With $\Delta y_{b,t}^M$ the change in its yield between the two trades and $\Delta y_{n_b,t}^{\text{Tsy}}$ the change in the treasury yield matched to its remaining maturity n_b , I estimate

$$\Delta y_{b,t}^M = \beta_0 \Delta y_{n_b,t}^{\text{Tsy}} + \beta_1 (\Delta y_{n_b,t}^{\text{Tsy}} \times n_b) + \beta_2 n_b + \varepsilon_{b,t}. \quad (28)$$

The response to a treasury yield change is $\beta_0 + \beta_1 n_b$. A positive β_1 means that the same treasury yield change is associated with a larger municipal yield change at longer maturities. Comparing trades within a bond removes its time-invariant characteristics. The sample covers January 2005 to August 2020 and includes non-callable, federally tax-exempt municipal bonds rated AA and above. Trades are at most 90 days apart, and the treasury yield must move by at least 10 basis points in either direction to limit transaction noise. The treasury comparison follows the bond's remaining maturity at each trade, and standard errors are clustered by the second trade's date. Appendix C.11 details the construction.

The estimates are $\hat{\beta}_0 = 0.346$ and $\hat{\beta}_1 = 0.007$ (Appendix Table 41). A 100-basis-point rise in the treasury yield goes with a municipal increase of about 35 basis points at one year and 42 at ten years. Splitting bonds at five years gives the same ordering, with responses of 0.351 and 0.426.

Traded yields are coupon yields, which need not move in proportion to treasury yields even when zero-coupon yields do. Pricing each bond's payments with municipal zero yields fixed at 0.65 times treasury zero yields gives a coupon benchmark that accounts for 30.0% of the gradient of 0.0080 in the bonds it can price and leaves a difference of 0.0056 (s.e. 0.0021). The difference stays positive when the monthly corporate curve replaces the daily treasury curve (Appendix C.11.1). All thirty combinations of treasury-move cutoffs and trade windows give a positive maturity interaction, from

¹⁵In the two-investor illustration, municipal sensitivity relative to a taxable reference whose one-period return responds one for one to the rate state increases over the thirty illustrated maturities. Appendix Proposition A.9(iii) establishes the maturity ordering for that reference in the extension with joint long-maturity holdings.

0.0047 to 0.0104, significant at five percent (Appendix Table 43). On the wider rating sample, the gradient is larger, 0.013 against 0.007 in the continuous specification and 0.113 against 0.075 with the five-year split (Appendix Tables 47 and 48).

7 Policy Counterfactuals

Tax reforms that change investors’ exemption benefits change who holds municipal bonds. The Tax Reform Act of 1986 generally removed banks’ interest deduction for new municipal purchases and was followed by a sharp decline in their holdings. The 2017 Tax Cuts and Jobs Act again changed banks’ and insurers’ relative exemption benefits. I study two reforms with existing policy precedents. The first removes a state’s exemption for residents’ interest on its own municipal bonds, which Illinois, Iowa and Wisconsin already generally tax. The second taxes municipal interest like taxable interest for every investor, as the 2009–2010 Build America Bonds program did for its taxable municipal bonds while subsidizing issuers’ interest.

The key distinction is whether investors can retain the federal benefit by switching municipal markets. Home-state repeal leaves that benefit in place, so high-tax households mainly switch between municipal markets while their replacement investments generate state-taxable interest. Equal treatment removes the federal benefit as well. Investors who previously benefited most shorten their bond portfolios, while lower-tax and tax-neutral investors take over more long bonds. The financing and revenue effects depend on both the interest these investors earn and the compensation they require for bearing rate risk.

7.1 Design and fiscal accounting

I solve each reform in the six calibrated three-year economies covering 2000–2017 and report calendar-year-weighted averages. Each economy uses the pre-2018 tax regime and period-specific investor and bond-market inputs (Section 5.2). I change the relevant tax rules and re-solve holdings and both yield curves, keeping bond supplies, investors’ total funding, risk preferences and outside exposures at their dated baseline values.¹⁶

Annual financing costs equal debt outstanding at each maturity times its yield, summed across maturities. Their increase is the replacement-financing cost, the additional annual cost of financing the same debt at the new yields.¹⁷ Each group’s pretax interest on bonds and taxable cash divides into tax receipts and after-tax interest income, with households taxed by residence and institutions under their own tax treatment. Because aggregate funding, bond supplies and the short rate are

¹⁶Tax changes are applied in each investor’s baseline payoff units, preserving preferences and liability hedges. Yield changes average each economy’s 36 monthly rate states. Dollar amounts, California’s implied state taxes and Figure 12(a) use $z = 0$. Ratios for 2000–2017 divide average dollar financing costs by average dollar receipts.

¹⁷The calculation uses continuously compounded yields to construct coupon-equivalent flows. Existing contractual payments and original-issue-discount schedules remain unchanged. The taxable market pools treasury and corporate debt. Other liability flows, issuer deductions, refinancing timing and fiscal feedback are outside this interest-income accounting (Appendix D.2).

fixed, total cash interest does not change. Every additional dollar issuers pay therefore goes either to governments or to investors. Summing across both markets and all governments gives

$$\Delta(\text{Financing costs}) = \Delta(\text{Tax receipts}) + \Delta(\text{After-tax interest income}). \quad (29)$$

I report financing costs per dollar of added receipts. For home-state repeal, I compare both own-state and all-issuer costs with the repealing state's added revenue, excluding changes in other governments' receipts. For equal treatment, I compare total municipal and taxable issuer costs with the combined increase in federal and state revenue. When costs and receipts cover all issuers and governments, equation (29) gives a ratio of one if aggregate after-tax interest income is unchanged and a ratio above one if that income rises.¹⁸

7.2 The home-state exemption

Most income-tax states exempt residents' interest on their own municipal bonds and tax interest on other states' bonds, which encourages residents to hold local debt. Illinois, Iowa and Wisconsin generally tax interest on their own bonds as well, so the policy I study already exists in some states.¹⁹

Panel (a) of Figure 12 gradually removes California households' exemption on local municipal interest in the 2015–2017 economy, keeping statutory state and federal tax rates and household composition fixed. The horizontal axis measures the remaining tax benefit of a local bond relative to an out-of-state bond for the high-tax household. The implied state tax is the percentage by which California's yield lies below the pooled municipal yield of the states without an income tax.²⁰ At the riskless two-year maturity, the implied state tax moves one for one with this benefit. Removing the exemption entirely lowers the implied state tax by 13.3 percentage points at two years but only 5.0 percentage points at thirty years. As in P1, the household's tax benefit is fully reflected in short yields and only partly in long yields, where lower-tax residents and institutions also hold the bonds.

Panel (b) removes residents' state exemption on home-state municipal bonds, separately in California, New York, New Jersey, Massachusetts, Ohio, Pennsylvania and Minnesota, seven states with high supply-weighted tax benefits. Residents then pay the same state tax on home-state as

¹⁸These ratios compare financing and interest-income flows. Welfare would also depend on changes in risk bearing, investment opportunities and public spending.

¹⁹Each of these three states exempts particular issues. The Supreme Court upheld the home-state preference in *Department of Revenue of Kentucky v. Davis* (2008). Home-state preferences also changed within the sample: Oklahoma exempts qualifying local obligations issued after July 1, 2001, and Utah has taxed out-of-state municipal interest on bonds acquired after January 1, 2003, except from states that do not tax Utah's bonds. Since 2018, the federal state-and-local-tax deduction cap has also increased the exemption's value for residents whose deductions it limits.

²⁰The implied state tax is $1 - y_{CA,T}^M/y_T^0$, where y_T^0 is the pooled no-income-tax-state municipal yield. It is the model counterpart of the P1 measure in (24), with y_T^0 in place of $\xi_t^{\text{fed}} y_T^C$. Both yields are solved at rate state $z = 0$. If a fraction f of local interest becomes state-taxable, the high-tax household's relative home-state tax benefit is $(1 - f)\tau_H^{\text{st}}/(1 - f\tau_H^{\text{st}})$. This compares its after-tax shares on local and out-of-state interest and equals its state rate when the local exemption is full.

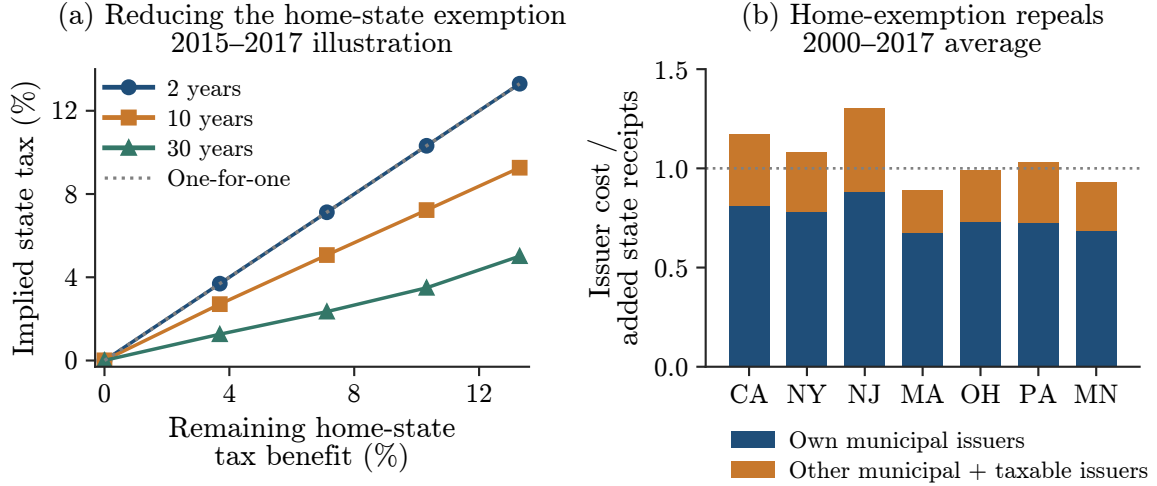


Figure 12: Home-state exemptions, municipal yields and financing costs. Panel (a) makes 0, 25, 50, 75 and 100% of local municipal interest state-taxable for both California household groups in the 2015–2017 economy, keeping all statutory tax rates fixed. It plots $1 - y_{CA,T}^M/y_T^0$ at two, ten and thirty years against the high-tax household’s remaining relative home-state tax benefit. The comparison yield y_T^0 pools the seven states without an income tax. Markers are solved equilibria at $z = 0$, and the dotted line is one-for-one. Panel (b) removes each listed state’s home-state exemption at fixed tax rates in each of the six 2000–2017 economies. It divides average replacement-financing costs for own-state municipal issuers and other municipal and pooled taxable issuers by average added state receipts from residents’ bond and taxable cash interest at $z = 0$.

on out-of-state municipal interest, while the tax treatment of institutions, nonresidents, cash and taxable bonds is unchanged. Added state receipts generally exceed local issuers’ added financing costs, which rise by \$0.68–\$0.88 per receipt dollar, below one in all seven six-period averages (Appendix Table 52). Including other municipal and taxable issuers raises the ratio to \$0.89–\$1.30, above one in California, New York, New Jersey and Pennsylvania. Both ratios use the same receipts, so their difference is the cost transmitted to other markets, of which taxable issuers bear \$0.11–\$0.31 per receipt dollar.

To see the mechanism, consider one dollar of state s ’s maturity- T debt initially held by a resident. Let y_T^0 be the pooled no-income-tax-state municipal yield and $\tau_{s,T}^{\text{impl}} = 1 - y_{s,T}^M/y_T^0$ the implied state tax. Holding y_T^0 fixed, the issuer’s yield rises in proportion to the fall in the implied state tax. The state, however, collects its full rate τ_s^{st} on the resident’s interest after repeal. For a dollar that continues to earn y_T^0 , the annual comparison is

$$\underbrace{\Delta y_{s,T}^M = -y_T^0 \Delta \tau_{s,T}^{\text{impl}}}_{\text{added cost to the issuer}} \quad \text{against} \quad \underbrace{\tau_s^{\text{st}} y_T^0}_{\text{tax on the formerly exempt dollar}}.$$

Under the representative-tax benchmark, residents determine relative yields at every maturity. Repeal lowers the implied state tax by the full rate, $-\Delta \tau_{s,T}^{\text{impl}} = \tau_s^{\text{st}}$, and the two sides are equal. In the model, this holds at the riskless short end. Long municipal bonds are also priced by lower-tax

residents and investors who do not receive the home-state exemption, so their yields reflect only part of the top-bracket resident’s benefit. Removing the exemption therefore raises long yields by less than the household benchmark predicts, consistent with the partial pass-through documented in P1. Across the six economies, California’s thirty-year implied state tax falls by 3.8 percentage points, against a mean high-tax household rate of 10.8%. This per-dollar comparison explains the limited yield response. The aggregate calculation must also account for resident ownership, because higher yields affect all local debt while the state taxes only its residents.

The federal exemption remains after home-state repeal, so residents can retain its benefit by buying other states’ municipal bonds. High-tax households therefore mainly switch between municipal markets, with relatively modest changes in the maturities they hold. Their sales of local bonds do not by themselves erode the state’s tax base, because interest on the replacement investments is also taxable. Across the seven experiments, residents’ aggregate risky-bond share changes by less than two percentage points (Appendix Table 51). Reinvestment at the same-maturity comparison yield y_T^0 preserves the benchmark receipts above. Lower-yielding investments reduce those receipts, while lower-tax households’ substitution into taxable bonds with higher gross yields increases them. The equilibrium calculation includes interest on purchases and interest forgone on sales, together with changes in yields and maturities. Appendix D.2 separates these effects.

Reallocation also spreads financing costs to other issuers because investors choose jointly across municipal and taxable markets. In the California experiment, other states’ municipal issuers bear an additional \$0.11 and taxable issuers \$0.25 per dollar of added California receipts. These spillovers raise the cost from \$0.81 for California’s own issuers to \$1.17 for all issuers. The state’s receipts can therefore exceed its own borrowers’ costs even when financing costs across all markets exceed those receipts.

7.3 Equal municipal and taxable treatment

The second reform taxes each investor’s municipal interest as it taxes that investor’s taxable interest. It removes federal and state exemption benefits for every investor group.²¹ The Build America Bonds program of 2009–2010 tested a version of it: issuers could sell taxable bonds and receive a federal subsidy of 35% of their interest. That program changed individual issues while the rest of the market stayed exempt. Equal treatment changes all municipal debt at once, so it also changes who holds long bonds and the yields they require.

I implement the reform by moving each municipal after-tax share 99.9% of the way to its taxable share and removing the same share of banks’ interest-deduction disallowance, so that banks’ municipal interest is taxed like their taxable interest. With exact equality, the riskless municipal bond and cash would be perfect substitutes, leaving investors’ split between them indeterminate. The remaining 0.1% pins down that split and keeps yields and the allocation of bonds across maturities close to those of full removal (Appendix B.2). The tax treatment of cash and taxable interest does not change.

²¹The reform is broader than the investor-specific changes in the 2017 Tax Cuts and Jobs Act (Appendix C.10).

Equal treatment raises financing costs by more than it raises tax receipts. All amounts below are average annual changes across the six 2000–2017 economies. Per dollar of added federal and state receipts, municipal issuers bear about \$1.06 in added financing costs and taxable issuers bear another \$0.13. Total costs rise by about \$1.20, leaving roughly \$0.20 per receipt dollar as additional after-tax interest income for investors (Table 5 and Figure 13(a)). In annual dollars, municipal issuers pay an additional \$41.2 billion and taxable issuers \$5.1 billion, for a total of \$46.3 billion. Governments collect \$38.7 billion and investors retain the remaining \$7.6 billion after tax, by (29). Appendix Table 53 reports each period separately.

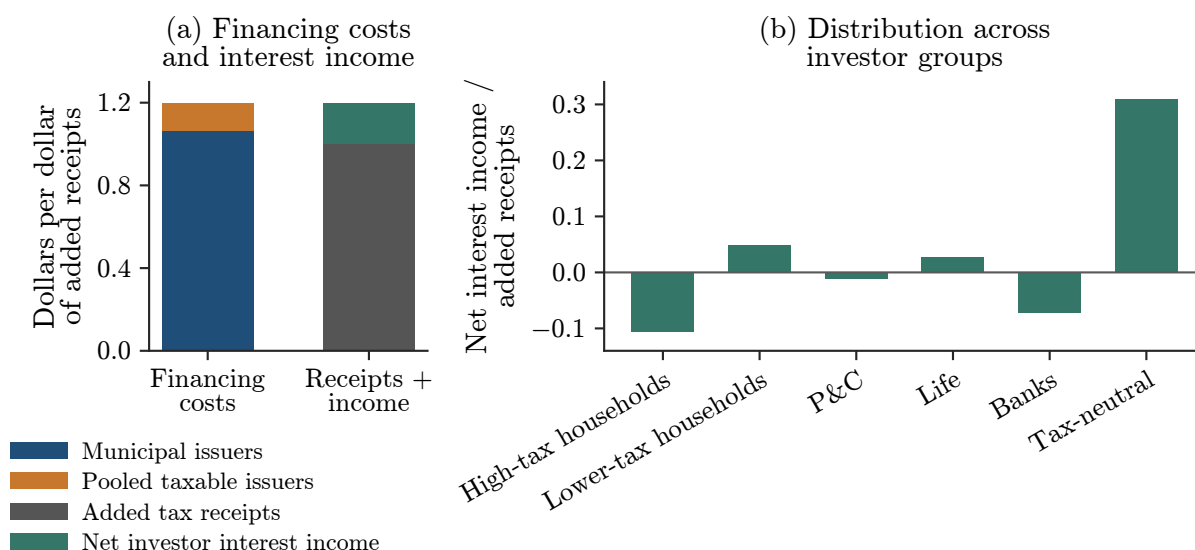


Figure 13: Average annual flows after removing 99.9% of each tax distinction in the six 2000–2017 economies. Panel (a) reports identity (29), dividing average financing costs and net investor interest income by average added federal and state receipts. Panel (b) allocates the income change across investor groups. Receipts and income include taxable cash interest. Each economy uses continuous yields and equilibrium holdings at $z = 0$. Taxable issuers pool treasury and corporate quantities.

Unlike home-state repeal, equal treatment leaves investors no municipal market in which they can retain the federal exemption. High-tax households, banks and P&C insurers shorten the average maturity of their risky bonds by 1.8 to 3.7 years, and high-tax households also increase their riskless holdings (Appendix Table 51). Shorter bonds and cash generate less interest than long bonds, limiting the revenue collected from the affected investors. Bonds beyond ten years go almost entirely to lower-tax households, life insurers and the tax-neutral group, on whom the reform imposes little additional tax. This shift in who receives long-bond interest further limits the revenue gain.

At the same time, taking over more long bonds raises these investors' rate exposure and the returns they require in both markets. Municipal yields rise nearly to taxable yields, increasing by 170.6 basis points at thirty years, while thirty-year taxable yields rise by 23.1 basis points. The transfer of rate risk therefore raises borrowing costs in both markets, while the new holders' smaller additional tax payments limit receipts.

Table 5: Average annual fiscal effects of equal municipal and taxable treatment, 2000–2017

$\Delta(\text{Financing costs})$			$\Delta(\text{Tax receipts}) + \Delta(\text{After-tax interest income})$		
	\$bn/ year	Per receipt dollar		\$bn/ year	Per receipt dollar
Municipal issuers	41.16	1.063	Federal receipts	34.29	0.886
Pooled taxable issuers	5.15	0.133	Residence-state receipts	4.42	0.114
			Total receipts	38.71	1.000
			+ Net investor interest income	7.59	0.196
Total financing costs	46.30	1.196	= Total receipts and income	46.30	1.196

Notes. Calendar-year-weighted changes across six three-year economies, each relative to its own baseline, with 99.9% of the tax distinction removed. Dollar amounts use dated-period quantities without inflation adjustment and continuous yields at $z = 0$. Per-receipt-dollar columns divide each average dollar change by average total receipts. Receipts and income include taxable cash interest, and net investor interest income is the change in after-tax interest income. Before rounding, total issuer costs equal receipts plus net investor interest income. Totals may differ because of rounding.

The \$7.6 billion net increase in annual after-tax interest income is distributed unevenly (Figure 13(b)). The tax-neutral group gains \$12.0 billion a year, lower-tax households gain \$1.9 billion and life insurers gain \$1.1 billion. High-tax households lose \$4.1 billion, banks lose \$2.8 billion and P&C insurers lose \$0.5 billion. These gains exceed the aggregate because investors who benefited more from the exemption lose income.

A direct subsidy, as under Build America Bonds, can return the added receipts to municipal issuers. Returning all \$38.7 billion of added annual receipts still leaves municipal issuers paying \$2.4 billion more per year and taxable issuers paying \$5.1 billion more. The subsidy covers most of municipal issuers' added costs but not the excess across both markets. With bond supplies and investors' total funding fixed, payments to issuers leave gross yields and investors' portfolios unchanged.

8 Conclusion

I develop an equilibrium term-structure model in which heterogeneous investor taxes create endogenous segmentation across bond markets and maturities. Investors jointly choose portfolios, balancing tax benefits against interest-rate risk. With short-sale constraints, these choices determine bond ownership and whose taxes are reflected in yields. Portfolio substitution links markets and maturities, so an investor's tax treatment can affect even bonds it never holds.

The municipal application shows why one representative investor's taxes cannot explain an entire yield curve. Top-bracket households value exemption most, yet investors with smaller exemption benefits hold substantial municipal portfolios. The model connects this ownership to relative yields: household taxes anchor the riskless short end, while longer yields also reflect lower-

tax investors' valuations and the allocation of rate risk. The quantitative and empirical evidence supports this explanation of the municipal bond puzzle.

Tax heterogeneity also shapes taxable yields because investors retain different shares of taxable interest. Joint portfolio choice transmits tax-policy changes across taxable and tax-exempt markets, as the counterfactuals demonstrate. Evaluating a reform therefore requires accounting for financing costs in both markets and the redistribution of investor income.

Municipal bonds provide clean tax variation for testing the mechanism and present a longstanding pricing puzzle. More broadly, the framework applies wherever investors differ in their tax treatment across assets and choose portfolios jointly under arbitrage frictions. Investor tax heterogeneity is a first-order determinant of bond yield curves and should be considered in term-structure models.

References

- Adelino, M., Cheong, S. C., Choi, J., and Oh, J. Y. J. (2026). Mutual fund flows and the supply of capital in municipal financing. *Review of Financial Studies*. Forthcoming.
- Adrian, T., Crump, R. K., and Moench, E. (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics*, 110(1):110–138.
- Ang, A., Bhansali, V., and Xing, Y. (2010). Taxes on tax-exempt bonds. *Journal of Finance*, 65(2):565–601.
- Asquith, P., Au, A. S., Covert, T., and Pathak, P. A. (2013). The market for borrowing corporate bonds. *Journal of Financial Economics*, 107(1):155–182.
- Auerbach, A. J. and King, M. A. (1983). Taxation, portfolio choice, and debt-equity ratios: A general equilibrium model. *Quarterly Journal of Economics*, 98(4):587–609.
- Babina, T., Jotikasthira, C., Lundblad, C. T., and Ramadorai, T. (2021). Heterogeneous taxes and limited risk sharing: Evidence from municipal bonds. *Review of Financial Studies*, 34(1):509–568.
- Bagley, J., Gissler, S., Hiteshew, K., and Ivanov, I. T. (2023). Pushing bonds over the edge: Investor demand and municipal bond liquidity. Federal Reserve Bank of Chicago Working Paper 2023-05, revised July 2023.
- Benson, E. D. and Marks, B. R. (2024). The impact of the Tax Cuts and Jobs Act on the true interest cost of bank-qualified tax-exempt bond issues. *Journal of Public Budgeting, Accounting & Financial Management*, 36(4):471–489.
- BlackRock Municipal Bond Fund, Inc. (2025). Statement of additional information: BlackRock National Municipal Fund and BlackRock Short Duration Muni Fund. SEC Form 485BPOS, <https://www.sec.gov/Archives/edgar/data/225635/000119312525248402/d56358d485bpos.htm>. Filed October 23, 2025.
- Brown, D. T. (2011). Callable bonds: What is special about munis? Working paper, SSRN 1925445.
- Brown, D. T. and Sirmans, S. (2011). Maturity clienteles in the municipal bond market: Term premiums and the muni puzzle. Working paper, SSRN 1928129.
- Cestau, D., Hollifield, B., Li, D., and Schürhoff, N. (2019). Municipal bond markets. *Annual Review of Financial Economics*, 11:65–84.
- Chalmers, J. M. R. (1998). Default risk cannot explain the muni puzzle: Evidence from municipal bonds that are secured by U.S. Treasury obligations. *Review of Financial Studies*, 11(2):281–308.
- Coppola, A. (2025). In safe hands: The financial and real impact of investor composition over the credit cycle. *Review of Financial Studies*, 38(8):2275–2325.
- Dagostino, R. (2025). The impact of bank financing on municipalities’ bond issuance and the real economy. *Journal of Financial Economics*, 166:104022.
- Dammon, R. M. and Green, R. C. (1987). Tax arbitrage and the existence of equilibrium prices for financial assets. *Journal of Finance*, 42(5):1143–1166.
- Domanski, D., Shin, H. S., and Sushko, V. (2017). The hunt for duration: Not waving but drowning? *IMF Economic Review*, 65(1):113–153.
- Dybvig, P. H. and Ross, S. A. (1986). Tax clienteles and asset pricing. *Journal of Finance*, 41(3):751–762.
- Fama, E. F. (1977). A pricing model for the municipal bond market. Unpublished manuscript, University of Chicago.
- Feenberg, D. R. and Poterba, J. M. (1991). Which households own municipal bonds? Evidence from tax returns. *National Tax Journal*, 44(4):93–103.
- Fleckenstein, M. and Longstaff, F. A. (2025). Do municipal bond investors pay a convenience premium to avoid taxes? *Review of Financial Studies*. Advance Access, hhaf100.
- Gabaix, X., Koijen, R. S. J., Mainardi, F., Oh, S. S., and Yogo, M. (2026). Asset demand of U.S. households. Working paper, SSRN 4251972, May 9, 2026.

- Garrett, D., Ordin, A., Roberts, J. W., and Suárez Serrato, J. C. (2023). Tax advantages and imperfect competition in auctions for municipal bonds. *Review of Economic Studies*, 90(2):815–851.
- Green, R. C. (1993). A simple model of the taxable and tax-exempt yield curves. *Review of Financial Studies*, 6(2):233–264.
- Green, R. C., Hollifield, B., and Schürhoff, N. (2007). Financial intermediation and the costs of trading in an opaque market. *Review of Financial Studies*, 20(2):275–314.
- Greenwood, R., Hanson, S., and Vayanos, D. (2024). Supply and demand and the term structure of interest rates. *Annual Review of Financial Economics*, 16(1):115–151.
- Greenwood, R., Hanson, S. G., and Liao, G. Y. (2018). Asset price dynamics in partially segmented markets. *Review of Financial Studies*, 31(9):3307–3343.
- Greenwood, R., Hanson, S. G., Stein, J. C., and Sunderam, A. (2023). A quantity-driven theory of term premia and exchange rates. *Quarterly Journal of Economics*, 138(4):2327–2389.
- Greenwood, R. and Vayanos, D. (2014). Bond supply and excess bond returns. *Review of Financial Studies*, 27(3):663–713.
- Gromb, D. and Vayanos, D. (2018). The dynamics of financially constrained arbitrage. *Journal of Finance*, 73(4):1713–1750.
- Guibaud, S., Nosbusch, Y., and Vayanos, D. (2013). Bond market clienteles, the yield curve, and the optimal maturity structure of government debt. *Review of Financial Studies*, 26(8):1914–1961.
- Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304.
- Haddad, V. and Sraer, D. A. (2020). The banking view of bond risk premia. *Journal of Finance*, 75(5):2465–2502.
- Hanson, S. G. and Stein, J. C. (2015). Monetary policy and long-term real rates. *Journal of Financial Economics*, 115(3):429–448.
- Harris, L. E. and Piwowar, M. S. (2006). Secondary trading costs in the municipal bond market. *Journal of Finance*, 61(3):1361–1397.
- He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *American Economic Review*, 103(2):732–770.
- Hendershott, P. H. and Kidwell, D. S. (1978). The impact of relative security supplies: A test with data from a regional tax-exempt bond market. *Journal of Money, Credit and Banking*, 10(3):337–347.
- Hendershott, P. H. and Koch, T. W. (1980). The demand for tax-exempt securities by financial institutions. *Journal of Finance*, 35(3):717–727.
- Hines, Jr., J. R. and Horton, E. (2024). Banks and tax-exempt debt arbitrage. Working Paper 32647, National Bureau of Economic Research.
- Hull, J. and White, A. (1994). Numerical procedures for implementing term structure models I: Single-factor models. *Journal of Derivatives*, 2(1):7–16.
- iShares Trust (2026). Statement of additional information, iShares National Muni Bond ETF and other series. SEC Form 485BPOS, <https://www.sec.gov/Archives/edgar/data/1100663/000119312526275730/d124972d485bpos.htm>. Filed June 18, 2026.
- Kekre, R., Lenel, M., and Mainardi, F. (2024). Monetary policy, segmentation, and the term structure. Working Paper 32324, National Bureau of Economic Research.
- Kidwell, D. S. and Koch, T. W. (1983). Market segmentation and the term structure of municipal yields. *Journal of Money, Credit and Banking*, 15(1):40–55.
- Kim, D. H. and Wright, J. H. (2005). An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates. Finance and Economics Discussion Series 2005-33, Board of Governors of the Federal Reserve System.
- Koijen, R. S. J. and Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515.

- Krishnamurthy, A. and Vissing-Jorgensen, A. (2012). The aggregate demand for Treasury debt. *Journal of Political Economy*, 120(2):233–267.
- Li, Y., O’Hara, M., and Zhou, X. A. (2024). Mutual fund fragility, dealer liquidity provision, and the pricing of municipal bonds. *Management Science*, 70(7):4802–4823.
- Li, Z. (2026). Long rates, life insurers, and credit spreads. *Review of Financial Studies*. Forthcoming.
- Livingston, M. (1982). The pricing of municipal bonds. *Journal of Financial and Quantitative Analysis*, 17(2):179–193.
- Longstaff, F. A. (2011). Municipal debt and marginal tax rates: Is there a tax premium in asset prices? *Journal of Finance*, 66(3):721–751.
- Mankiw, N. G. and Poterba, J. M. (1996). Stock market yields and the pricing of municipal bonds. Working Paper 5607, National Bureau of Economic Research.
- Miller, M. H. (1977). Debt and taxes. *Journal of Finance*, 32(2):261–275.
- Moussawi, H. (2022). Asset purchases and demand in municipal bonds. Working paper, SSRN 4017698, revised April 2025.
- Moussawi, H. (2024). The efficiency of municipal bond tax shields. Working paper, SSRN 5065844, revised August 2026.
- Mussa, M. L. and Kormendi, R. C. (1979). *The Taxation of Municipal Bonds: An Economic Appraisal*. Number 229 in AEI Studies in Tax Policy. American Enterprise Institute for Public Policy Research, Washington, DC.
- Nuveen Municipal Trust (2026). Statement of additional information, Nuveen All-American Municipal Bond Fund and other series. SEC Form 485BPOS, <https://www.sec.gov/Archives/edgar/data/1020660/000119312526322012/d113980d485bpos.htm>. Filed July 29, 2026.
- Pati, A. (2023). Heterogeneous investors and risk premiums in the municipal bond market. Essay in Ph.D. dissertation, Emory University.
- Pirinsky, C. A. and Wang, Q. (2011). Market segmentation and the cost of capital in a domestic market: Evidence from municipal bonds. *Financial Management*, 40(2):455–481.
- Poterba, J. M. (1986). Explaining the yield spread between taxable and tax-exempt bonds: The role of expected tax policy. In Rosen, H. S., editor, *Studies in State and Local Public Finance*, pages 5–52. University of Chicago Press, Chicago.
- Poterba, J. M. and Ramírez Verdugo, A. (2011). Portfolio substitution and the revenue cost of the federal income tax exemption for state and local government bonds. *National Tax Journal*, 64(2, Part 2):591–613.
- Schwert, M. (2017). Municipal bond liquidity and default risk. *Journal of Finance*, 72(4):1683–1722.
- Sialm, C. (2009). Tax changes and asset pricing. *American Economic Review*, 99(4):1356–1383.
- Siani, K. Y. (2022). Raising bond capital in segmented markets. Working paper, MIT Sloan School of Management. November 7 version.
- Siriwardane, E. N., Sunderam, A., and Wallen, J. (2025). Segmented arbitrage. *Journal of Finance*, 80(5):2543–2590.
- Skelton, J. L. (1983). Banks, firms and the relative pricing of tax-exempt and taxable bonds. *Journal of Financial Economics*, 12(3):343–355.
- Spreen, T. L. and Gerrish, E. (2022). Taxes and tax-exempt bonds: A literature review. *Journal of Economic Surveys*, 36(4):767–808.
- St. Clair, T. (2024). The supply elasticity of municipal debt: Evidence from bank-qualified bonds. *National Tax Journal*, 77(1):111–139.
- Vayanos, D. and Vila, J.-L. (2021). A preferred-habitat model of the term structure of interest rates. *Econometrica*, 89(1):77–112.

A The model: formal results and proofs

This appendix develops the formal results behind Section 4. Appendices A.1–A.3 give the general economy used for calibration, including bond-specific risk. Setting that risk to zero gives the portfolio-pricing framework in the main text. Appendix A.4 introduces the two-investor affine benchmark. The results then follow the main-text argument: participation, the three return-gap channels, rate loadings, yield-implied taxes and empirical predictions.

Appendix A.9 verifies the main-text illustration. Appendix A.10 adds bond-specific risk, which allows both investors to hold a range of long municipal maturities, and compares the figures. Appendix A.11 treats other participation motives, and Appendices A.12–A.14 extend the analysis to several municipal states, many clienteles and other markets. The analytical solutions drop the shorting constraint on cash unless stated otherwise. They also solve the problem with this constraint wherever cash remains nonnegative. Appendix B gives the numerical system with the constraint, and Table 6 collects the notation.

Table 6: Notation for the model and the proofs

Symbol	Meaning
$z, r(z) = \max\{\bar{r} + z, 0\}$	Latent rate state and the one-period log gross cash return
$\tau_i^a, \xi_i^a, \xi_i = \xi_i^C / \xi_i^M$	Investor i 's tax rate and after-tax share in market a . In the two-market benchmark, ξ_i is its relative after-tax share and $1 - \xi_i$ its exemption benefit
$\xi_H < \xi_I, \xi^*$	Benchmark after-tax shares of the high-tax household and lower-tax investor, and their tax mix $\omega_H \xi_H + \omega_I \xi_I$ with risk-tolerance shares ω_i
$p_T^a(z), A_T^a = -\partial_z p_T^a, y_T^a = -p_T^a/T$	Log price of the maturity- T bond in market a , its rate loading (the model's duration), and its yield, with $A_0^a = 0$
$C_T^a(z) = -p_T^a(z) - A_T^a(z)z$	Level term, the constant intercept in the affine benchmark
$\mu_T^a(z) = \log \mathbb{E}_z[1 + R_T^a], s_T^a = \partial_z \mu_T^a$	Mean adjusted holding return, equal to its holders' required return in the KKT conditions, and its response to the state
$R_{T,t+1}^a, \mathcal{R}_{T,t+1}^a = \mu_T^a + u_{T,t+1}^a$	Simple holding-period return and the return used in portfolio choice, with mean μ_T^a and centered log-return innovation $u_{T,t+1}^a$
$\sigma_{\eta,T}^a, v_T^a = (\sigma_{\eta,T}^a)^2$	Loading on bond-specific risk and its variance, zero for the one-period claim
$\nu_{i,T}^a$	KKT slack: required return minus offered return, zero for a positive holding
$x_{i,T}^a, Q_T^a, m_i, \gamma_i$	Holdings per unit of mass, total supply, investor mass, risk aversion. In the calibration, m_i measures baseline portfolio assets
$W_i, W_i', \widehat{W}_i', \kappa_i$	Funding per unit of mass, next-period net wealth, its discrete-time approximation, and the multiplier on the shorting constraint on cash
$\widehat{D}_i, \ell_i$	Net rate exposure and the outside liability loading. Appendix B.7 writes $h_i = -\ell_i$
$\Lambda, \widehat{\Lambda}$	Duration gap $\gamma_H \widehat{D}_H - \gamma_I \widehat{D}_I$ at the household-only allocation and at the equilibrium allocation
$\omega_{i,T}^a, \mathcal{U}_T^a$	Risk-tolerance share of an active holder and the set of active holders of the maturity- T bond in market a
$g_T = \mu_T^M - \xi_H \mu_T^C$	Gap between the municipal one-period return and the high-tax household's after-tax taxable return
$\tau_T^{Y,\text{impl}} = 1 - y_T^M / y_T^C$	Model yield-implied tax, the counterpart of <i>ImpliedTax</i> in (2)

A.1 The general economy

The quantitative model of Section 5 solves this economy with several investor groups, state municipal markets, binding shorting constraints on cash and the rate floor.

Assets and investors

Time is discrete and indexed by t . One aggregate rate state z_t follows a stationary first-order autoregression, AR(1), and the short rate, the pretax log gross return on cash, is bounded below at zero:

$$z_{t+1} = \rho z_t + \sigma_r \varepsilon_{t+1}, \quad r_t = r(z_t) := \max\{\bar{r} + z_t, 0\}, \quad \rho \in (0, 1), \quad \varepsilon_{t+1} \sim \mathcal{N}(0, 1). \quad (30)$$

Here ρ measures the persistence of the rate state and σ_r the size of its shocks. The parameter $\bar{r}$ is the unfloored cash rate at the stationary mean $z = 0$. The floor applies to the cash return while the state z_{t+1} remains Gaussian. Its effects on continuation prices are described in Appendix A.2.

Let $\mathcal{A}$ be a finite set of bond markets, indexed by a and distinguished by asset class or issuer. In the two-investor benchmark, $\mathcal{A}$ contains a taxable market C and a municipal market M ; the quantitative model has 45 municipal markets, pooling the seven states without an income tax. Each market contains zero-coupon bonds indexed by their remaining maturity $T \in \{1, \dots, T_{\max}\}$, the number of periods until the bond pays one unit. At maturity, $P_0^a = 1$. The one-period claim of the taxable market C is the cash account, with log gross return $r(z)$, so separately held taxable bonds begin at $T = 2$. A bond with T periods remaining has price $P_T^a(z_t)$ and holding-period return

$$1 + R_{T,t+1}^a = \frac{P_{T-1}^a(z_{t+1})}{P_T^a(z_t)} \exp(\sigma_{\eta,T}^a \eta_{T,t+1}^a). \quad (31)$$

The standard normal shocks $\eta_{T,t+1}^a$ are independent of each other and of ε_{t+1} , and they add bond-specific log-return variance $v_T^a := (\sigma_{\eta,T}^a)^2$. Setting $v_T^a = 0$ gives the main-text return. Every market has $\sigma_{\eta,1}^a = 0$, so a one-period claim is fully riskless. Bond-specific risk represents residual variation in holding returns, such as an issue-specific change in trading demand. It makes longer bonds imperfect substitutes without adding a state variable that affects future prices. Appendix B.8 explains the normalization and its small effect on the calibrated gap.

Investors belong to finitely many types, indexed by i . Type i has mass $m_i > 0$, risk aversion $\gamma_i > 0$, funding W_i and outside exposure ℓ_i to the rate shock from assets and liabilities beyond its bonds. It keeps the absolute after-tax share $\xi_i^a = 1 - \tau_i^a > 0$ of returns in market a , where τ_i^a is its tax rate on those returns. Because after-tax shares differ across both investors and markets, the same structure covers municipal and taxable bonds, state taxes on treasury and corporate interest, tax-favored accounts and cross-border withholding (Appendix A.14).

Portfolio choice and equilibrium

Investor i holds bonds $x_{i,T,t}^a$ and cash $x_{i,t}^0 = W_i - \sum_{a,T} x_{i,T,t}^a$, in dollars per unit of mass. Aggregate holdings are $m_i x_{i,T,t}^a$, in the same units as supply. As in Section 4.1, portfolio choice uses $\mathcal{R}_{T,t+1}^a = \mu_T^a(z_t) + u_{T,t+1}^a$, where $\mu_T^a(z_t) := \log \mathbb{E}_t[1 + R_{T,t+1}^a]$ and $u_{T,t+1}^a := \log(1 + R_{T,t+1}^a) - \mathbb{E}_t[\log(1 + R_{T,t+1}^a)]$. Thus μ is the conditional mean of the adjusted return, including the Jensen adjustment, and u is the centered log-return innovation. Substituting $R \approx \mathcal{R}$ for bonds and $e^r - 1 \approx r$ for cash gives approximated next-period wealth $\widehat{W}_i'$ and the mean-variance problem:²²

$$\begin{aligned} \max_{x_i \geq 0, x_{i,t}^0 \geq 0} \quad & \mathbb{E}_t[\widehat{W}_i'] - \frac{\gamma_i}{2} \text{Var}_t(\widehat{W}_i'), \\ \text{subject to} \quad & \widehat{W}_i' = W_i + \xi_i^C r(z_t) x_{i,t}^0 \\ & + \sum_{a,T} \xi_i^a x_{i,T,t}^a \mathcal{R}_{T,t+1}^a + \ell_i \sigma_r \varepsilon_{t+1}. \end{aligned} \quad (32)$$

Positions and cash are nonnegative, which rules out short sales and borrowing (Section 3.5). Funding and masses are fixed across states. A positive ℓ_i raises net wealth when rates rise and so offsets the rate exposure of the investor's bonds.

²²Appendices C.1.4 and D.5 give the capital-income tax correction.

Investor i 's after-tax excess return on bond (a, T) over cash is $\tilde{\mu}_{i,T}^a = \xi_i^a \mu_T^a - \xi_i^C r$. Let $\tilde{\Pi}_i(t+1)$ be the after-tax innovation of its whole portfolio, including the outside exposure. Let $\kappa_i(z) \geq 0$ denote the multiplier on the shorting constraint on cash, $\sum_{a,T} x_{i,T}^a \leq W_i$. It is zero whenever the investor holds cash. The portfolio, or Karush–Kuhn–Tucker (KKT), conditions are

$$\begin{aligned} x_{i,T}^a(z) > 0 &\implies \underbrace{\tilde{\mu}_{i,T}^a(z) = \gamma_i \xi_i^a \text{Cov}_z \left(u_{T,t+1}^a, \tilde{\Pi}_i(t+1) \right)}_{\text{marginal portfolio-risk cost}} + \underbrace{\kappa_i(z)}_{\text{shorting constraint on cash}}, \\ x_{i,T}^a(z) = 0 &\implies \tilde{\mu}_{i,T}^a(z) \leq \gamma_i \xi_i^a \text{Cov}_z \left(u_{T,t+1}^a, \tilde{\Pi}_i(t+1) \right) + \kappa_i(z). \end{aligned} \quad (33)$$

The multiplier satisfies complementary slackness, $\kappa_i(z)[W_i - \sum_{a,T} x_{i,T}^a(z)] = 0$.

Every bond other than cash clears at every rate state:

$$x_{i,T}^a(z) \geq 0, \quad \sum_i m_i x_{i,T}^a(z) = Q_T^a, \quad (34)$$

where Q_T^a is the fixed dollar supply. An equilibrium consists of prices and holdings satisfying (31), (33) and (34). Participation, which investors hold each bond, is determined by these conditions.

The reduced problem

Collecting the cash and bond terms in approximated wealth (32) gives each bond's expected after-tax excess return over cash

$$\tilde{\mu}_{i,T}^a(z_t) := \underbrace{\xi_i^a \mu_T^a(z_t)}_{\text{after-tax bond return}} - \underbrace{\xi_i^C r(z_t)}_{\text{after-tax cash return}}. \quad (35)$$

The bond and outside innovations have conditional mean zero, so expected approximated wealth separates into the return on the initial funding and the portfolio's excess returns:

$$\mathbb{E}_t[\widehat{W}_i'] = W_i(1 + \xi_i^C r(z_t)) + \sum_{a,T} x_{i,T,t}^a \tilde{\mu}_{i,T}^a(z_t).$$

The variance of wealth is that of the after-tax portfolio innovation, the random part of (32):

$$\tilde{\Pi}_i(t+1) = \sum_{a \in \mathcal{A}} \sum_T \xi_i^a x_{i,T,t}^a u_{T,t+1}^a + \underbrace{\ell_i \sigma_r \varepsilon_{t+1}}_{\text{liability hedge}}. \quad (36)$$

Dropping the portfolio-independent term $W_i(1 + \xi_i^C r(z_t))$ and eliminating cash gives the equivalent reduced problem:

$$\max_{\substack{x_i \geq 0 \\ \sum_{a,T} x_{i,T,t}^a \leq W_i}} \underbrace{\sum_{a \in \mathcal{A}} \sum_T x_{i,T,t}^a \tilde{\mu}_{i,T}^a(z_t)}_{\text{after-tax excess return}} - \underbrace{\frac{\gamma_i}{2} \text{Var}_t(\tilde{\Pi}_i(t+1))}_{\text{whole-portfolio risk cost}}. \quad (37)$$

The problem's first-order conditions are (33).

Rate risk comes from the next-period resale price. With log price $p_T^a := \log P_T^a$ and rate loading $A_T^a := -\partial_z p_T^a$, a bond with T periods remaining has rate innovation $-A_{T-1}^a(\rho z_t) \sigma_r \varepsilon_{t+1}$ to first order, at the conditional mean resale state. A one-period bond carries no rate innovation, since $A_0^a = 0$, although its price can respond to today's state. A positive ℓ_i offsets the portfolio's rate exposure, and ℓ_i can represent net outside exposure of either sign.

The discrete-time return approximation

With simple holding-period returns $R_{T,t+1}^a$, exact next-period net wealth is

$$W'_i = W_i + \xi_i^C x_i^0 (e^{r(z_t)} - 1) + \sum_{a,T} \xi_i^a x_{i,T}^a R_{T,t+1}^a + \ell_i \sigma_r \varepsilon_{t+1}.$$

Taking logs of (31) gives the change in log price plus the bond-specific shock. Following Greenwood et al. (2024), I approximate the simple bond return by this log return plus a Jensen adjustment, and $e^r - 1$ by r for cash. These substitutions give approximated wealth $\widehat{W}'_i$ in (32), with adjusted bond return

$$\mathcal{R}_{T,t+1}^a = \log(1 + R_{T,t+1}^a) + \underbrace{\log \mathbb{E}_t[1 + R_{T,t+1}^a] - \mathbb{E}_t[\log(1 + R_{T,t+1}^a)]}_{\text{Jensen adjustment}}.$$

The Jensen adjustment is known at time t . Taking the conditional expectation gives $\mathbb{E}_t[\mathcal{R}_{T,t+1}^a] = \mu_T^a(z_t) = \log \mathbb{E}_t[1 + R_{T,t+1}^a]$. Subtracting this mean cancels the adjustment:

$$u_{T,t+1}^a = \mathcal{R}_{T,t+1}^a - \mu_T^a(z_t) = \log(1 + R_{T,t+1}^a) - \mathbb{E}_t[\log(1 + R_{T,t+1}^a)].$$

In the affine benchmark, log returns are conditionally Gaussian, so

$$\mu_T^a(z_t) = \mathbb{E}_t[\log(1 + R_{T,t+1}^a)] + \frac{1}{2} \text{Var}_t[\log(1 + R_{T,t+1}^a)].$$

Together with the centered innovation u , this is the mean and innovation formulation in equations (7)–(8) of Greenwood et al. (2024). The Jensen adjustment is known at time t , so it changes the mean but leaves log-return variances and covariances unchanged. The nonlinear solver uses $\mu = \log \mathbb{E}_t[1 + R]$ directly, retaining the full Jensen adjustment even when log returns are not Gaussian.

The price recursion is exact for the chosen mean $\mu = \log \mathbb{E}_t[1 + R]$. Its use in the linear wealth equation approximates simple returns, whose conditional mean is $e^\mu - 1$. Appendix D.5 compares the two conventions for a deterministic return. I re-solve the one-market economy of Section 5 with simple returns, taxed exactly, refitting risk aversion to restore the targeted taxable yields and holding the other coefficients fixed. The thirty-year gap in that economy rises from 102.6 to 107.7 basis points.

The shorting constraint on cash and riskless claims

The shorting constraint on cash in (37) keeps cash nonnegative, so investors cannot borrow. Its multiplier $\kappa_i(z)$ in (33) is zero whenever the investor holds cash, since an extra dollar can then stay in cash. A positive multiplier requires the investor to be fully invested, so that buying more of one bond means selling another, and it raises the required pretax return by κ_i/ξ_i^a on each held bond.

Consider a riskless one-period claim in any market $a \neq C$. Its bond-specific variance $v_1^a = (\sigma_{\eta,1}^a)^2$ is zero and $p_0^a = 0$, so its return has no covariance with portfolios or liabilities. Retaining κ_i in (33), each eligible investor's KKT condition becomes

$$\mu_1^a \leq \frac{\xi_i^C r + \kappa_i}{\xi_i^a}, \quad \text{with equality for a positive holding.}$$

With positive supply, at least one investor holds the claim, so its return equals the smallest of these bounds:

$$\mu_1^a(z) = \min_i \frac{\xi_i^C r(z) + \kappa_i(z)}{\xi_i^a}. \quad (38)$$

The price recursion (45) then gives $p_1^a = -\mu_1^a$, hence $y_1^a = \mu_1^a$.

A.2 From holding returns to yields

Taking expectations of the gross holding return (31), including bond-specific risk with variance $v_T^a = (\sigma_{\eta,T}^a)^2$, gives

$$P_T^a(z) = e^{-\mu_T^a(z) + v_T^a/2} \mathbb{E}_z[P_{T-1}^a(z')], \quad P_0^a = 1. \quad (39)$$

A higher required return lowers today's price for a given expected resale value. Iterating from the maturity payment links yields to the one-period returns along the bond's life. The factor $e^{v_T^a/2}$ accounts for the mean of the bond-specific shock (see the exact recursion below), whose yield effect is small (Appendix B.8).

Write $p_T^a = \log P_T^a$ for the log price and $y_T^a = -p_T^a/T$ for the yield. In the affine benchmark of Appendix A.4,

$$\log P_T^a(z) = -A_T^a z - C_T^a, \quad y_T^a(z) = \frac{A_T^a z + C_T^a}{T}. \quad (40)$$

The coefficient A_T^a is the bond's rate loading, the model's counterpart of duration: the proportional fall in its price per unit rise in the rate state. Thus A_T^a/T is the yield's response to the rate state and C_T^a/T is its level at the mean state. The response of the one-period return to the rate state is

$$s_T^a(z) := \frac{\partial \mu_T^a(z)}{\partial z}. \quad (41)$$

Substituting affine prices into the resale equation and matching its state and constant terms gives

$$A_T^a = \underbrace{\rho A_{T-1}^a}_{\text{inherited loading}} + \underbrace{s_T^a}_{\text{current response}}, \quad C_T^a = C_{T-1}^a + \mu_T^a(0) - \frac{1}{2} \left[\sigma_r^2 (A_{T-1}^a)^2 + (\sigma_{\eta,T}^a)^2 \right]. \quad (42)$$

The first relation adds today's return response to the loading inherited through the expected resale price. The second accumulates the mean-state one-period returns less half-variance terms. Starting from $A_0^a = C_0^a = 0$, the loading $A_T^a = \sum_{j=1}^T \rho^{T-j} s_j^a$ accumulates the responses, weighted by rate persistence. With a riskless one-period claim, $\sigma_{\eta,1}^a = 0$, the same recursion gives the yield as the average of the one-period returns along the bond's life less convexity:

$$y_T^a(z_t) = \underbrace{\frac{1}{T} \sum_{j=0}^{T-1} \mathbb{E}_t[\mu_{T-j}^a(z_{t+j})]}_{\text{one-period returns as the bond ages}} - \underbrace{\frac{1}{2T} \sum_{k=2}^T \left[\sigma_r^2 (A_{k-1}^a)^2 + (\sigma_{\eta,k}^a)^2 \right]}_{\text{convexity}}. \quad (43)$$

The formulas allow zero or positive bond-specific variance. The main-text analytical illustration sets both markets' bond-specific variances to zero. The extension in Appendix A.10 gives risky municipal bonds positive variance and sets taxable bond-specific variance to zero.

The exact price recursion

The exact recursion below links the equilibrium holding returns to prices. Its affine specialization gives the loading and yield formulas used in Section 4.5.

With fixed supplies and funding, I seek an equilibrium log-price function for each market and maturity:

$$p_{T,t}^a = p_T^a(z_t), \quad p_0^a(z) \equiv 0, \quad a \in \mathcal{A}. \quad (44)$$

Let $Z' = \rho z + \sigma_r \varepsilon'$. Since μ_T^a is the log of the bond's conditional expected gross return (31), solving for today's log price gives

$$p_T^a(z) = \underbrace{\log \mathbb{E}_z[\exp\{p_{T-1}^a(Z')\}]}_{\text{log expected resale price}} + \underbrace{\frac{1}{2}(\sigma_{\eta,T}^a)^2}_{\text{bond-specific payoff term}} - \underbrace{\mu_T^a(z)}_{\text{one-period return}}. \quad (45)$$

Equation (45) discounts the expected resale payoff at the equilibrium holding-period return, so a higher return lowers today's price. This recursion links adjacent maturities, while the portfolio conditions (33) link

markets through investors' holdings. Prices and portfolios solve these equations and clearing (34) jointly at every state. Appendix B gives the numerical implementation. Equations (49)–(50) in Appendix A.3 write the portfolio system with a slack shorting constraint on cash.

The rate floor, changes in participation and binding shorting constraints on cash make prices nonlinear in the state. Resale carries these features across maturities, because today's continuation value averages future prices over states with different participation and cash constraints.

Iterating (45) from $p_0^a = 0$ expresses the price through the one-period returns the bond earns as it ages through maturities $T, T-1, \dots, 1$:

$$P_T^a(z_t) = e^{p_T^a(z_t)} = \mathbb{E}_t \left[\exp \left\{ - \sum_{j=0}^{T-1} \left(\mu_{T-j}^a(z_{t+j}) - \frac{1}{2} (\sigma_{\eta, T-j}^a)^2 \right) \right\} \right]. \quad (46)$$

Taking the negative log of (46) and dividing by T gives the yield $y_T^a = -p_T^a/T$. In the affine benchmark, this becomes the explicit average of expected one-period returns less convexity in (43). In the general economy, the yield's dependence on the state is summarized by writing the log price in loading form,

$$p_T^a(z) = -A_T^a(z)z - C_T^a(z), \quad A_T^a(z) := -\frac{\partial p_T^a(z)}{\partial z}, \quad (47)$$

so that $P_T^a = \exp\{-A_T^a z - C_T^a\}$ and, at differentiability points,

$$y_T^a(z) = \frac{A_T^a(z)z + C_T^a(z)}{T}, \quad \frac{\partial y_T^a(z)}{\partial z} = \frac{A_T^a(z)}{T}.$$

The level term C_T^a includes accumulated one-period returns and convexity, and the mean-state yield is $C_T^a(0)/T$. Outside the affine benchmark, both A_T^a and C_T^a can vary with the state. The identity $\partial_z C_T^a = -z \partial_z A_T^a$ preserves the yield derivative above.

The loading recursion links the loading A_T^a to the one-period return response $s_T^a = \partial_z \mu_T^a$ of equation (41). Above the rate floor, cash has response one, and by (38) a riskless one-period claim in another market has response ξ_i^C/ξ_i^a for its active holder i when the shorting constraint on cash is slack. At longer maturities the response also includes the change in the holders' risk premia. For locally absolutely continuous continuation prices, an integrable envelope for the Gaussian exponential moment and its state derivative permits differentiation under the expectation in (45):

$$A_T^a(z) = \underbrace{s_T^a(z)}_{\text{current return response}} + \rho \underbrace{\frac{\mathbb{E}_z[\exp\{p_{T-1}^a(Z')\} A_{T-1}^a(Z')]}{\mathbb{E}_z[\exp\{p_{T-1}^a(Z')\}]}}_{\text{continuation price response}}. \quad (48)$$

The fraction averages the continuation loading, giving more weight to states where the resale bond is more valuable. In the affine benchmark, (48) reduces to the loading recursion (42).

A.3 Equilibrium conditions and the one-period boundary

The two-investor results assume $0 < \rho < 1$, $\sigma_r > 0$, $m_i, \gamma_i > 0$, $0 < \xi_H < \xi_I$ and positive traded supplies. The one-period taxable claim is riskless cash. An allocation's support is the set of investor–bond pairs with positive holdings. A participation branch fixes the support, and a bond's active set is its holders. Affine results use floorless rates and a fixed participation branch, whereas the numerical model retains the floor and the shorting constraint on cash.

One KKT system for all markets. Index risky market–maturity coordinates by $j = (a, T)$ and write $v_j = (\sigma_{\eta, T}^a)^2$, $K_i = \text{diag}(\xi_i^a)$ and the relative after-tax share $\xi_{ij} = \xi_i^C/\xi_i^a$, collected in the vector ξ_i . In the normalized two-market benchmark, $\xi_{ij} = \xi_i$ at municipal coordinates and $\xi_{ij} = 1$ at taxable coordinates. With resale log price $p_{j,-} = p_{T-1}^a$, the exact covariance and liability credit are

$$\Sigma_{jk}(z) = \text{Cov}_z(p_{j,-}(Z'), p_{k,-}(Z')) + \mathbf{1}_{\{j=k\}} v_j, \quad c_{ij}^L(z) = -\text{Cov}_z(p_{j,-}(Z'), \ell_i \sigma_r \varepsilon'). \quad (49)$$

With a slack shorting constraint on cash, dividing the first-order condition by the positive asset after-tax share gives

$$\begin{aligned} \nu_{ij} &:= \xi_{ij}r + \gamma_i[\Sigma K_i x_i - c_i^L]_j - \mu_j \geq 0, & x_{ij} &\geq 0, & x_{ij}\nu_{ij} &= 0, \\ \sum_i m_i x_{ij} &= Q_j. \end{aligned} \tag{50}$$

Concavity of (37) makes these conditions sufficient for investor optimality. The next proposition solves today's clearing problem with future resale prices held fixed. Positive bond-specific variances give strict concavity on risky coordinates, but do not by themselves establish equilibrium uniqueness.

Proposition A.1 (Current-state clearing). *Fix the state and continuation prices, with finite $\Sigma \succeq 0$, finite liability credits and at least one eligible investor for every positive-supply coordinate.*

(i) *The KKT-and-clearing system (50) has a current-state solution.*

(ii) *If, in addition, $\Sigma = V + \sigma_r^2 AA'$ with V positive diagonal and a loading vector $A \gg 0$ on the risky coordinates, after-tax shares are positive and supported supplies are positive, the current-state clearing returns and risky holdings are unique, even with endogenous support changes. Deterministic one-period claims are treated separately. Payoff-equivalent holdings in tied boundary claims need not be unique.*

Neither part establishes the dynamic fixed point of continuation prices.

Proof. (i) Let $X = \{x \geq 0 : \sum_i m_i x_i = Q\}$ be the set of nonnegative allocations that clear, with inner product $\langle x, y \rangle_m = \sum_i m_i x_i' y_i$. Positive masses and finite supplies make X nonempty, compact and convex. For $G_i(x_i) = \xi_i r + \gamma_i(\Sigma K_i x_i - c_i^L)$, investor i 's required return on each coordinate, Brouwer gives a fixed point of the continuous map $x \mapsto \text{Proj}_X(x - G(x))$. The projection characterization implies

$$\sum_i m_i \langle G_i(x_i), y_i - x_i \rangle \geq 0 \quad (y \in X).$$

Transfers from positive holders to other eligible investors show that each positive holder minimizes G_{ij} . Thus $\mu_j = \min_i G_{ij}$ and $\nu_{ij} = G_{ij} - \mu_j$ satisfy (50). Conversely, complementarity implies the displayed inequality for every feasible transfer. Predetermined trading exclusions restrict eligibility. (ii) Strict concavity makes individual demand $x_i(\mu)$ continuous, Lipschitz and piecewise affine. On a region where investor i holds exactly the coordinates in a set B_i , its derivative is

$$J_i = \gamma_i^{-1} E_{B_i} K_{i,B_i}^{-1} \Sigma_{B_i B_i}^{-1} E_{B_i}',$$

where E_B embeds the held coordinates. Sherman–Morrison gives

$$\Sigma_{BB}^{-1} = V_B^{-1} - \frac{\sigma_r^2 V_B^{-1} A_B A_B' V_B^{-1}}{1 + \sigma_r^2 A_B' V_B^{-1} A_B}, \quad \Sigma_{BB}^{-1} A_B = \frac{V_B^{-1} A_B}{1 + \sigma_r^2 A_B' V_B^{-1} A_B} \gg 0.$$

Every selection of $J = \sum_i m_i J_i$ therefore has nonpositive off-diagonal entries and $JA \geq 0$, strictly on held coordinates. If μ^0, μ^1 both cleared, piecewise-affine continuity would give $0 = \bar{J}(\mu^1 - \mu^0)$, with $\bar{J}$ the segment-average selection matrix. Positive supply at μ^0 and demand continuity imply positive holdings on an initial segment for each asset. Finitely many selections then give $(\bar{J}A)_j > 0$ everywhere. Hence $\bar{J} \text{diag}(A)$ is strictly diagonally dominant and nonsingular, so returns coincide, and strict concavity gives identical risky holdings. On support-boundary segments, adjacent selections have the same directional derivative. $\square$

Regularity for derivatives. Derivatives below are taken at regular states, at which the cash rate and price moments are continuously differentiable, complementarity is strict on risky coordinates and the active KKT-and-clearing Jacobian is nonsingular. The implicit-function theorem then gives a locally fixed participation branch and differentiable holdings and returns. Fixed supplies imply, for each bond's active set $\mathcal{U}_j$,

$$\sum_{i \in \mathcal{U}_j} m_i \frac{\partial x_{ij}}{\partial z} = 0. \tag{51}$$

Individual strict concavity does not imply nonsingularity, because with heterogeneous after-tax shares the normalized operator contains ΣK_i , which need not be symmetric or strongly monotone.

The one-period boundary. Specializing the riskless-claim return (38) to the normalized municipal-taxable benchmark gives the short-end anchor (12) and pass-throughs (20).

Proposition A.2 (Deterministic one-period boundary). *In the normalized municipal-taxable benchmark with slack shorting constraints on cash, suppose the one-period municipal claim has $v_1^M = 0$ and positive supply, and $r(z) > 0$. Suppose the high-tax household has the minimal after-tax share, $\xi_H = \min_i \xi_i$. Then only investors with the minimal after-tax share can hold it, and*

$$y_1^M = \mu_1^M = \xi_H r(z), \quad y_1^C = r(z), \quad \tau_1^{Y, \text{impl}} = 1 - \xi_H. \quad (52)$$

If the high-tax household is the unique after-tax share minimizer, its tax passes through one-for-one and the tax of an investor with a strictly higher after-tax share has zero effect at this boundary. At a tie the household derivative requires perturbing all tied minimal after-tax shares together. At $r = 0$ the yield ratio is undefined.

Proof. With $\xi_i^M = 1$, $\xi_i^C = \xi_i$ and $\kappa_i = 0$, the riskless-claim KKT condition of Appendix A.1 is $\mu_1^M \leq \xi_i r$, with equality for positive holdings. Since $r > 0$, only investors with the minimal after-tax share can satisfy it with equality. Positive supply gives $\mu_1^M = \min_i \xi_i r = \xi_H r$ and $p_1^M = -\mu_1^M$. Taxable cash has $p_1^C = -r$. Differentiating the yield ratio proves the tax statements. $\square$

A.4 The two-investor affine benchmark across maturities

The two-investor benchmark yields the closed forms behind Sections 4.3–4.6.

The high-tax household H and the lower-tax investor I choose among municipal maturities $T = 1, \dots, T_{\max}$ and taxable maturities $T = 2, \dots, T_{\max}$, in fixed supplies Q_T^a , and cash, the one-period taxable claim. The benchmark allows bond-specific risk and different risk aversions γ_i , and keeps the relative after-tax share $\xi_i = \xi_i^C / \xi_i^M$ of Section 4.2, with $\xi_H < \xi_I$. Payoffs are in units of after-tax municipal income, so taxable returns and cash income are multiplied by ξ_i , and γ_i and ℓ_i are normalized correspondingly. I fix a participation branch, which specifies which investors hold each bond, solve its portfolio conditions and market clearing, and check that no investor benefits from changing its holdings. Closed forms use three restrictions:

- (i) the shorting constraint on cash is dropped;
- (ii) the floored cash rate is replaced by the linear rule $r(z) = \bar{r} + z$; and
- (iii) prices are computed as if the same participation branch applied at every state, which makes log prices affine, and the branch is verified at the states studied.

On a fixed branch, holdings, one-period returns and log prices are affine in z , and the price coefficients solve a fixed point. The construction below states the feasibility and uniqueness conditions. The quantitative model lets participation change across states and keeps the shorting constraint on cash and the rate floor.

Log prices then have constant coefficients,

$$p_T^{a, \text{aff}}(z) = -A_T^{a, \text{aff}} z - C_T^{a, \text{aff}}, \quad (53)$$

and one-period returns are affine as well, $\mu_T^{a, \text{aff}}(z) = \mu_T^{a, \text{aff}}(0) + s_T^{a, \text{aff}} z$, so $s_T^{a, \text{aff}}$ and $A_T^{a, \text{aff}}$ are the state coefficients of the one-period return and of the negative log price. The affine recursions (42) and the yield average (43) of Appendix A.2 hold exactly. Below and in the proofs, loadings, level terms and responses refer to this affine branch, which the superscript “aff”, where shown, marks explicitly.

Investor i ’s net rate exposure, the after-tax exposure of its whole balance sheet to the rate shock, is

$$\widehat{D}_i(z) := \underbrace{\sum_T x_{i,T}^M(z) A_{T-1}^M + \xi_i \sum_T x_{i,T}^C(z) A_{T-1}^C}_{\text{tax-weighted asset exposure}} - \underbrace{\ell_i}_{\text{outside liability exposure}}. \quad (54)$$

Municipal positions count at their loadings and taxable positions at ξ_i times theirs in these payoff units. With today's positions fixed, $\widehat{D}_i = -\partial \widehat{W}_i' / \partial z_{t+1}$, as in (10). It is dollar exposure before division by wealth. The rate component of the wealth innovation is $-\widehat{D}_i \sigma_r \varepsilon_{t+1}$, so a larger liability hedge ℓ_i lowers this exposure. A bond's rate-risk compensation equals its loading times the investor's unit risk cost, $\gamma_i \sigma_r^2 \widehat{D}_i$. Loadings are constant on an affine branch, so changes in $\widehat{D}_i$ across rate states reflect trading alone.

Bond-specific risk allows both investors to hold several of the same municipal maturities. Without it, their municipal equalities hold at the same maturity only where the difference in their rate-risk compensation on that bond equals the difference in exemption benefits.

At one period, zero risk and a slack shorting constraint on cash give $\mu_1^M = \xi_H r$ whenever $r > 0$ (Proposition A.2). The household, which keeps the smallest share of taxable interest, then holds the riskless municipal claim. At longer maturities, participation determines whose taxes and risk costs shape equilibrium returns. Proposition A.1 gives the uniqueness conditions for current-state clearing, and Appendix A.6 those for the loading solution.

The thirty-maturity main-text illustration has rate risk alone and shares municipal maturity 23. Appendix A.9 gives its parameters and verifies its portfolio conditions and price-loading recursions. Appendix A.10 adds bond-specific risk, giving joint municipal participation from maturity 17, and verifies its loading fixed point. At the household-only allocation, the household holds every municipal bond and the lower-tax investor every taxable bond, giving $A_T^M / A_T^C = \xi_H$ at every maturity (Proposition A.6). Level panels omit the vertical scale because these illustrations are not calibrations.

Constructing affine solutions

The construction fixes a participation branch, sets inactive positions to zero, sets $r(z) = \bar{r} + z$ and uses affine continuation prices on $\mathbb{R}$. It solves the active equalities jointly across maturities, retaining whole-portfolio risk, and then checks nonnegative holdings, inactive slacks and, for finite budgets, $W_i - \sum_j x_{ij}(z) \geq 0$.

For affine candidates $p_T^a(z) = -A_T^a z - C_T^a$, Gaussian integration gives exactly

$$\log \mathbb{E}_z e^{p_{T-1}^a(z')} = -\rho A_{T-1}^a z - C_{T-1}^a + \frac{1}{2} \sigma_r^2 (A_{T-1}^a)^2. \quad (55)$$

The current-state covariance $\Sigma = \sigma_r^2 A A' + \text{diag}(v)$, with A collecting predecessor loadings, and liability credit $c_i^L = \ell_i \sigma_r^2 A$ are constant. If the active KKT-and-clearing matrix is nonsingular, holdings and μ are affine in z , and the price recursion returns affine prices with updated loadings $\rho A_{T-1}^a + s_T^a$. At a coefficient fixed point, all active equalities, clearing equations and exact continuation equations hold identically, and, with $A_0^a = 0$,

$$A_T^a = \rho A_{T-1}^a + s_T^a = \sum_{j=1}^T \rho^{T-j} s_j^a, \quad s_T^a = \frac{\partial \mu_T^a}{\partial z}. \quad (56)$$

Affine holdings and slacks make the KKT-valid states an interval, possibly empty, conditional on the global affine continuation. Gaussian transitions leave any proper interval, so interval feasibility is not global equilibrium feasibility. For example, at the household-only allocation the lower-tax investor's slack on the one-period municipal claim, $(\xi_I - \xi_H)(\bar{r} + z)$, turns negative at negative rates. The cash floor changes continuation values.

For a fixed participation branch, a continuous coefficient update on a nonempty compact convex invariant set has a fixed point by Brouwer. A contraction gives uniqueness within that set. Nonsingularity, invariance and reinstated KKT feasibility must be checked separately. Appendices A.9 and A.10 verify the reported analytical illustrations, and these checks supply no global existence argument.

A.5 The participation margin

The comparison below asks whether a small municipal purchase would benefit the lower-tax investor when the household initially holds all municipal bonds. This household-only allocation, marked by the superscript hh, assigns all risky taxable bonds to the lower-tax investor. Its net rate exposure already includes its taxable portfolio and the offset from any positive outside liability exposure.

Proposition A.3 sets bond-specific risk to zero, $v_T^M = 0$, as in the equilibrium comparison of Section 4.3. Here, the comparison is evaluated at household-only holdings to identify a first-order purchase incentive.

Theorem 1 below adds positive bond-specific risk at the equilibrium allocation. In the zero-variance case, the high-tax household's active KKT equality and the lower-tax investor's condition for a nonnegative first-order purchase incentive are

$$\underbrace{\mu_T^M - \xi_H r}_{\text{offered to the household}} = \underbrace{\gamma_H \sigma_r^2 A_{T-1}^M \widehat{D}_H^{\text{hh}}}_{\text{household marginal risk cost}}, \quad (57)$$

$$\underbrace{\mu_T^M - \xi_I r}_{\text{offered to the lower-tax investor}} \geq \underbrace{\gamma_I \sigma_r^2 A_{T-1}^M \widehat{D}_I^{\text{hh}}}_{\text{lower-tax investor's marginal risk cost}}. \quad (58)$$

If the second inequality fails, an initial purchase is unprofitable at this allocation. Substituting the household's equality cancels the common return μ_T^M and leaves one condition:

$$\underbrace{\sigma_r^2 A_{T-1}^M \Lambda}_{\text{duration benefit (grows with maturity)}} \geq \underbrace{(\xi_I - \xi_H) r}_{\text{difference in exemption benefits (flat in maturity)}}, \quad \Lambda := \gamma_H \widehat{D}_H^{\text{hh}} - \gamma_I \widehat{D}_I^{\text{hh}}. \quad (59)$$

The duration benefit on the left, the lower-tax investor's saving in rate-risk cost on the bond, equals the bond's loading times $\sigma_r^2 \Lambda$, the difference between the two investors' unit costs of rate risk. Assumption 1 below requires $\Lambda > 0$: risk-weighted net rate exposure is larger for the household at this allocation. At the household-only allocation, clearing fixes every holding, so $s_T^M = \xi_H$ and (42) gives $A_{T-1}^M = \xi_H \sum_{j=0}^{T-2} \rho^j$. Starting from $A_0^M = 0$, the loading rises strictly with maturity, so the duration benefit grows from zero and crosses the difference in exemption benefits at most once.

Assumption 1 (Household duration burden at the household-only allocation). *At the household-only allocation, where the high-tax household holds all municipal bonds and the lower-tax investor all taxable bonds,*

$$\widehat{D}_H^{\text{hh}} = \sum_T A_{T-1}^M Q_T^M / m_H - \ell_H, \quad \widehat{D}_I^{\text{hh}} = \xi_I \sum_T A_{T-1}^C Q_T^C / m_I - \ell_I, \quad \Lambda := \gamma_H \widehat{D}_H^{\text{hh}} - \gamma_I \widehat{D}_I^{\text{hh}} > 0.$$

Under the assumption, the duration benefit $\sigma_r^2 A_{T-1}^M \Lambda$ is nonnegative and grows with the bond's loading.

Proposition A.3 (Participation from the long end). *In the two-investor affine benchmark without bond-specific risk, suppose $\Lambda > 0$ at the household-only allocation and $r_t > 0$. At that allocation, the maturities with a nonnegative first-order purchase incentive for the lower-tax investor form the upper interval $\{T : T \geq T^*(z_t)\}$, where*

$$T^*(z_t) = \min \left\{ T \geq 1 : A_{T-1}^M \geq \frac{(\xi_I - \xi_H) r_t}{\sigma_r^2 \Lambda} \right\}, \quad (60)$$

with $\min \emptyset = +\infty$. A strict inequality makes participation strictly profitable, and equality need not imply a positive holding.

Proof. Under the proposition's zero-variance assumption, the household holds Q_T^M / m_H of each municipal bond, and its active municipal equality sets

$$\mu_T^{M,\text{hh}} = \xi_H r + \gamma_H \sigma_r^2 A_{T-1}^M \widehat{D}_H^{\text{hh}}.$$

At zero holdings, the lower-tax investor requires $\xi_I r + \gamma_I \sigma_r^2 A_{T-1}^M \widehat{D}_I^{\text{hh}}$. Subtracting the offered return gives its KKT slack:

$$\nu_{I,T}^{\text{hh}} = (\xi_I - \xi_H) r - \sigma_r^2 A_{T-1}^M \Lambda. \quad (61)$$

The slack is nonpositive exactly when (59) holds. Sole-holder quantities are state invariant, giving $s_T^M = \xi_H$ and strictly increasing $A_T^M = \xi_H \sum_{q=0}^{T-1} \rho^q$. Since $\Lambda > 0$, the slack decreases strictly with maturity, proving the upper-interval result. Zero slack gives indifference at a zero position, while negative slack makes an initial purchase strictly profitable. $\square$

The threshold T^* is set by taxes, balance sheets and loadings, and no maturity habitat is assigned. At fixed benchmark loadings, a larger difference in exemption benefits moves the threshold out and a larger Λ

moves it in. With positive bond-specific risk, the household's required return also includes compensation for the bond-specific risk of its whole holding, while a new buyer's bond-specific risk cost is zero at a zero position. This additional participation incentive appears explicitly in the equilibrium index (62) of Theorem 1 below.

In equilibrium, prices and both portfolios are re-solved. Theorem 1 below gives conditions under which the lower-tax investor's municipal holdings form an upper maturity interval with positive bond-specific risk, using the equilibrium difference $\hat{\Lambda} := \gamma_H \hat{D}_H - \gamma_I \hat{D}_I$. Section 4.3 states the comparison without bond-specific risk in (13). The theorem adds the bond-specific risk cost to that comparison. From here on, T^* denotes the start of the lower-tax investor's holdings.

On the joint-ownership branch with bond-specific risk used in the results below, the household holds municipal bonds alone below T^* and both investors hold them from T^* on. The main-text illustration instead has a single shared maturity and lower-tax-only holdings above it (Appendix A.9). The calibrated economy has several longer-maturity clienteles (Section 5).

Theorem 1 (Upper-interval participation by the lower-tax investor on an affine branch). *At a regular state of a two-investor affine solution, suppose $v_T^M > 0$ on the risky maturities considered. Define the duration gap at the equilibrium allocation, $\hat{\Lambda} = \gamma_H \hat{D}_H - \gamma_I \hat{D}_I$, and*

$$\Psi_T = (\xi_I - \xi_H)r - \sigma_r^2 A_{T-1}^M \hat{\Lambda} - \gamma_H v_T^M Q_T^M / m_H. \quad (62)$$

Then $x_{I,T}^M > 0$ if and only if $\Psi_T < 0$. If $\hat{\Lambda} > 0$ and the predecessor loadings and $v_T^M Q_T^M$ are nondecreasing, the lower-tax investor's positive support is an upper interval, possibly empty. This identifies the cutoff within a given equilibrium and does not establish uniqueness across equilibria.

Proof. If the lower-tax investor is inactive, $x_{H,T}^M = Q_T^M / m_H$. Substituting the household equality into the lower-tax investor's inequality gives $\Psi_T = \nu_{I,T} \geq 0$. If both investors hold the bond, subtracting their equalities and using clearing gives

$$\Psi_T = -v_T^M (\gamma_I + \gamma_H m_I / m_H) x_{I,T}^M < 0.$$

If the lower-tax investor alone is active, its equality and the household inequality imply $(\xi_I - \xi_H)r - \sigma_r^2 A_{T-1}^M \hat{\Lambda} \leq -\gamma_I v_T^M Q_T^M / m_I < 0$, which also gives $\Psi_T < 0$. These cases prove the equivalence. Finally, adjacent indices differ by

$$\Psi_{T+1} - \Psi_T = -\sigma_r^2 (A_T^M - A_{T-1}^M) \hat{\Lambda} - \frac{\gamma_H}{m_H} (v_{T+1}^M Q_{T+1}^M - v_T^M Q_T^M) \leq 0.$$

With $\hat{\Lambda} > 0$ and nondecreasing predecessor loadings and $v_T^M Q_T^M$, the index is nonincreasing in maturity, and the first maturity with a strictly negative index is the participation threshold T^* . Equality is a zero-holding boundary. $\square$

Joint holdings. For candidate net rate exposures, put $N_T = (\xi_I - \xi_H)r - \sigma_r^2 A_{T-1}^M \hat{\Lambda}$, the difference in exemption benefits less the duration benefit. For $v_T := v_T^M > 0$, the difference of the two active KKT equalities and clearing give

$$x_{H,T}^M = \frac{\gamma_I Q_T^M + m_I N_T / v_T}{\gamma_H m_I + \gamma_I m_H}, \quad x_{I,T}^M = \frac{\gamma_H Q_T^M - m_H N_T / v_T}{\gamma_H m_I + \gamma_I m_H}. \quad (63)$$

Both holdings are nonnegative exactly when $-\gamma_I Q_T^M / m_I \leq N_T / v_T \leq \gamma_H Q_T^M / m_H$. Each maturity's holdings depend on the net rate exposures, which are common to all maturities and must match the resulting portfolios.

A.6 The pricing channels

This subsection gives the formal results behind Sections 4.4 and 4.5. It first derives the three channels of the one-period return gap, the exemption benefit, the duration cost and the rate loading, and shows

where compensation for bond-specific risk enters when that risk is present. It then derives the tax-mix contribution and the municipal and taxable portfolio responses within the rate-loading channel.

The three channels of the one-period return gap

Fix maturity T , and let the active set $\mathcal{U}_T^M \subseteq \{H, I\}$ denote the investors with positive holdings of that municipal bond. Specializing the active portfolio condition (33) to the affine benchmark gives the main-text condition (11), now including bond-specific risk:

$$\underbrace{\mu_T^M - \xi_i r}_{\text{after-tax excess return offered to } i} = \gamma_i \left[\underbrace{\sigma_r^2 A_{T-1}^M \widehat{D}_i}_{\text{whole-portfolio rate risk}} + \underbrace{(\sigma_{\eta,T}^M)^2 x_{i,T}^M}_{\text{bond-specific risk}} \right]. \quad (64)$$

On the branches studied here the lower-tax investor alone holds the risky taxable bonds. Dividing its taxable condition by its after-tax share gives, for $T \geq 2$,

$$\mu_T^C - r = \gamma_I [\sigma_r^2 A_{T-1}^C \widehat{D}_I + \xi_I (\sigma_{\eta,T}^C)^2 x_{I,T}^C], \quad (65)$$

and $\mu_1^C = r$ for cash. Subtracting ξ_H times the taxable equality from the equality of a municipal holder separates the one-period return gap $g_T = \mu_T^M - \xi_H \mu_T^C$ into the three channels plus bond-specific compensation. Proposition 1, the formal counterpart of (15) in Section 4.4, states this decomposition.

Proposition 1 (Three channels of the one-period return gap). *In the two-investor affine benchmark, consider a state at which the lower-tax investor alone holds the risky taxable bonds, and let i denote a holder of the maturity- T municipal bond, choosing $i = I$ whenever the lower-tax investor holds it. Then*

$$\begin{aligned} g_T = & \underbrace{(\xi_i - \xi_H)r}_{\text{exemption benefit}} + \underbrace{\sigma_r^2 A_{T-1}^M (\gamma_i \widehat{D}_i - \gamma_I \widehat{D}_I)}_{\text{duration cost}} + \underbrace{\gamma_I \sigma_r^2 \widehat{D}_I (A_{T-1}^M - \xi_H A_{T-1}^C)}_{\text{rate loading}} \\ & + \underbrace{\gamma_i (\sigma_{\eta,T}^M)^2 x_{i,T}^M - \xi_H \gamma_I \xi_I (\sigma_{\eta,T}^C)^2 x_{I,T}^C}_{\text{bond-specific compensation}}. \end{aligned} \quad (66)$$

At $T = 1$ the duration-cost, rate-loading and bond-specific contributions are zero, and so is the exemption-benefit contribution when $r > 0$. Where the household holds the bond alone, the exemption-benefit contribution is zero and the duration-cost contribution is at most $(\xi_I - \xi_H)r$. Where the lower-tax investor holds it, the exemption-benefit contribution is $(\xi_I - \xi_H)r$ and the duration-cost contribution is zero.

Proof. Investor i 's active municipal equality (64) gives $\mu_T^M = \xi_i r + \gamma_i [\sigma_r^2 A_{T-1}^M \widehat{D}_i + (\sigma_{\eta,T}^M)^2 x_{i,T}^M]$. The lower-tax investor's taxable equality (65) gives $\mu_T^C = r + \gamma_I [\sigma_r^2 A_{T-1}^C \widehat{D}_I + \xi_I (\sigma_{\eta,T}^C)^2 x_{I,T}^C]$, also covering cash at $T = 1$ because $A_0^C = 0$ and $\sigma_{\eta,1}^C = 0$. Subtracting ξ_H times the second equality from the first and adding and subtracting $\gamma_I \sigma_r^2 A_{T-1}^M \widehat{D}_I$ gives the three-channel identity (66).

At $T = 1$, $A_0^a = 0$ and there is no bond-specific risk, so the duration-cost, rate-loading and bond-specific terms vanish. When $r > 0$, the household alone holds the claim (Proposition A.2), so $i = H$ and the after-tax cash-return difference also vanishes.

If the household alone holds maturity T , then $i = H$ and $x_{H,T}^M = Q_T^M / m_H$. The lower-tax investor's inactive condition gives $\Psi_T \geq 0$ in (62), with $v_T^M = (\sigma_{\eta,T}^M)^2$, or $\sigma_r^2 A_{T-1}^M \widehat{\Lambda} + \gamma_H (\sigma_{\eta,T}^M)^2 Q_T^M / m_H \leq (\xi_I - \xi_H)r$. The duration-cost contribution is the first term on the left and is therefore at most $(\xi_I - \xi_H)r$. If the lower-tax investor holds the bond, the identity with $i = I$ has a zero duration-cost contribution and an exemption-benefit contribution of $(\xi_I - \xi_H)r$. $\square$

Below the participation threshold, where $i = H$, the duration-cost contribution equals $\sigma_r^2 A_{T-1}^M \widehat{\Lambda}$. At jointly held maturities either investor's equality gives the same total gap. The tax mix and portfolio responses derived below determine the loadings in the third channel.

The rate-loading channel links maturities. Iterating the loading recursion (42) from $A_0^a = 0$,

$$A_{T-1}^M - \xi_H A_{T-1}^C = \sum_{j=1}^{T-1} \rho^{T-1-j} (s_j^M - \xi_H s_j^C),$$

so the rate loading accumulates the response gaps $s_j^M - \xi_H s_j^C$ of the shorter maturities the bond passes through. It is zero at two periods, since the one-period responses are $s_1^M = \xi_H$ and $s_1^C = 1$ when $r > 0$.

Figure 5 shows the main-text decomposition without bond-specific risk. At thirty periods, the rate-loading term accounts for 54% of the return gap and the exemption-benefit term for the remainder. Figure 14 gives the counterpart with bond-specific risk, using the same channel names.

Tax mix and portfolio responses

The rate-loading channel depends on how one-period returns respond to the rate state. A bond's one-period return responds for two reasons. Each holder's after-tax return on cash rises by its after-tax share per unit of the rate state. Trading also changes how much net rate exposure each holder carries, and with it the compensation each requires. Aggregating the holders' equalities separates the tax-mix and portfolio responses. Multiplying investor i 's equality (64) by m_i/γ_i , summing over the active holders and differentiating with respect to z removes the bond-specific terms, because fixed supply implies $\sum_i m_i \partial x_{i,T}^M(z)/\partial z = 0$. What remains is

$$\left(\sum_{i \in \mathcal{U}_T^M} \frac{m_i}{\gamma_i} \right) s_T^M - \sum_{i \in \mathcal{U}_T^M} \frac{m_i \xi_i}{\gamma_i} = \sigma_r^2 A_{T-1}^M \sum_{i \in \mathcal{U}_T^M} m_i \frac{\partial \hat{D}_i(z)}{\partial z}. \quad (67)$$

Define the active holders' risk-tolerance shares and aggregate exposure response as

$$\omega_{i,T}^M := \frac{m_i/\gamma_i}{\sum_{j \in \mathcal{U}_T^M} m_j/\gamma_j}, \quad E_T^M := \frac{\sum_{i \in \mathcal{U}_T^M} m_i \partial \hat{D}_i(z)/\partial z}{\sum_{i \in \mathcal{U}_T^M} m_i/\gamma_i}.$$

The weight $\omega_{i,T}^M$ is type i 's risk tolerance, m_i/γ_i , as a share of the active holders' total. The term E_T^M is the response of their aggregate net rate exposure to the rate state, divided by that total risk tolerance, and E_T^C is defined in the same way from the holders of the taxable bond. Dividing (67) by total risk tolerance gives Proposition A.4, the formal counterpart of the response formulas (16) in Section 4.5.

Proposition A.4 (Holding-return response to the rate state). *Consider a regular state of the two-investor affine benchmark: price moments are differentiable, active risky holdings and inactive KKT slacks are strictly positive, and the stacked active KKT-and-clearing Jacobian is nonsingular, so the active set $\mathcal{U}_T^M \subseteq \{H, I\}$ is locally fixed. Then*

$$s_T^M = \underbrace{\sum_{i \in \mathcal{U}_T^M} \omega_{i,T}^M \xi_i}_{\text{holders' tax mix}} + \underbrace{\sigma_r^2 A_{T-1}^M E_T^M}_{\text{portfolio-risk response}} \quad (68)$$

exactly. The tax-mix term depends on who holds maturity T . The portfolio-risk response depends on how equilibrium trading changes those holders' exposure across their whole portfolios. This decomposes one market's return response. Comparing the two markets below separates the portfolio response into its municipal and taxable contributions to the rate-loading channel.

Proof of Proposition A.4. Divide (67) by the active holders' total risk tolerance. Fixed supply cancels the bond-specific derivative by (51). The definitions of $\omega_{i,T}^M$ and E_T^M then give (68). $\square$

Proposition A.5 extends the same aggregation to nonlinear prices. It collects the active holders' tax treatment and their covariance-based risk compensation before taking the state derivative.

Proposition A.5 (Exact active-clientele aggregation). *At a regular state, let $\Gamma_j = \sum_{i \in \mathcal{U}_j} m_i/\gamma_i$ be the active holders' total risk tolerance, $\bar{\xi}_j = \Gamma_j^{-1} \sum_{i \in \mathcal{U}_j} m_i \xi_{ij}/\gamma_i$ and $\Omega_j = \Gamma_j^{-1} \sum_{i \in \mathcal{U}_j} m_i [\Sigma K_i x_i - c_i^L]_j$. Here $\bar{\xi}_j$ is their tax mix and Ω_j their risk-tolerance-weighted risk compensation. Then*

$$\begin{aligned} \mu_j &= \bar{\xi}_j r + \Omega_j, \\ s_j &= \bar{\xi}_j \partial_z r + \frac{1}{\Gamma_j} \sum_{i \in \mathcal{U}_j} m_i [(\partial_z \Sigma) K_i x_i + \Sigma K_i \partial_z x_i - \partial_z c_i^L]_j. \end{aligned} \quad (69)$$

In the normalized municipal-taxable benchmark, above the cash floor the response difference is

$$s_T^M - \xi_H s_T^C = (\bar{\xi}_T^M - \xi_H) + \partial_z \Omega_T^M - \xi_H \partial_z \Omega_T^C. \quad (70)$$

Proof. Weighting the active equalities $\mu_j - \xi_{ij} r = \gamma_i [\Sigma K_i x_i - c_i^L]_j$ by m_i/γ_i , summing and dividing by Γ_j gives the first line. Differentiating on the locally fixed participation branch gives the second line. In the two-market normalization, $\bar{\xi}_j = \bar{\xi}_T^M$ at municipal and $\bar{\xi}_j = 1$ at taxable coordinates. Subtraction gives (70). $\square$

When the shorting constraint on cash binds, each active equality also contains κ_i/ξ_i^a by (33), so the return identity in (69) adds $\Gamma_j^{-1} \sum_{i \in \mathcal{U}_j} m_i \kappa_i/(\gamma_i \xi_i^a)$ and the response identity adds its state derivative. A per-dollar charge such as the bank financing-tax cost appears in the return identity in the same units and, when it does not vary with the rate state, leaves the response identity unchanged on a given participation branch.

In the affine benchmark, substituting the covariance and liability credits into the active KKT equalities gives the coordinate-level response

$$s_j = \bar{\xi}_j + \frac{\sigma_r^2 A_j}{\Gamma_j} \sum_{i \in \mathcal{U}_j} m_i \frac{\partial \hat{D}_i}{\partial z} + \frac{v_j}{\Gamma_j} \sum_{i \in \mathcal{U}_j} m_i \xi_i^a \frac{\partial x_{ij}}{\partial z}, \quad (71)$$

where $A_j = A_{T-1}^a$ is the predecessor loading. For municipal bonds, whose after-tax shares are normalized to one, (51) cancels the final term.

The tax-mix term is ξ_H under household-only holdings, ξ_I under lower-tax-only holdings and, under joint holdings, the tax mix

$$\xi^* := \omega_H \xi_H + \omega_I \xi_I, \quad \omega_i := \frac{m_i/\gamma_i}{m_H/\gamma_H + m_I/\gamma_I},$$

where ω_i is investor i 's share of total risk tolerance. Thus the lower-tax investor's after-tax share enters this term only at maturities it holds.

To sign the portfolio response, consider a branch on which the lower-tax investor alone holds risky taxable bonds, both investors hold municipal maturities in a nonempty set $J \subseteq \{2, \dots, T_{\max}\}$, and the household alone holds every other municipal maturity. Proposition 2 shows that at a regular state with $r(z) > 0$, positive loadings and positive bond-specific variance on J ,

$$d := \frac{\partial \hat{D}_H(z)}{\partial z} > 0, \quad \frac{\partial \hat{D}_I(z)}{\partial z} = -\frac{m_H}{m_I} d < 0.$$

A higher rate state raises the difference in exemption benefits, $(\xi_I - \xi_H)r$, and trading in jointly held bonds shifts net rate exposure to the household. Its unit cost of rate risk rises, raising its required compensation on all its risky municipal holdings, including maturities the lower-tax investor never holds. The sign holds for exposure-weighted portfolios, and individual jointly held positions can move either way. At a household-only risky maturity,

$$s_T^M - \xi_H = \gamma_H \sigma_r^2 A_{T-1}^M \frac{\partial \hat{D}_H(z)}{\partial z} > 0. \quad (72)$$

The taxable leg moves the other way. Because the lower-tax investor's net rate exposure falls as the rate state rises, the response of its taxable return is

$$s_T^C = 1 - bA_{T-1}^C, \quad b := \gamma_I \sigma_r^2 \frac{m_H}{m_I} d > 0,$$

which lies below one at every risky maturity. The coefficient b measures the feedback of portfolio reallocation into taxable returns.

If joint municipal participation begins at T^* and continues to the long end, the municipal responses are

$$s_T^M = \begin{cases} \xi_H + \gamma_H \sigma_r^2 A_{T-1}^M d, & T < T^*, \\ \xi^* = \omega_H \xi_H + \omega_I \xi_I, & T \geq T^*. \end{cases} \quad (73)$$

Below the threshold, the lower-tax investor's after-tax share affects the municipal return response through the household's portfolio adjustment. From the threshold on, both investors hold the bond. Clearing cancels their aggregate exposure response, so the tax mix ξ^* remains. These are the household-only and joint-ownership cases of (16). Without bond-specific risk, the illustration also has a lower-tax-only block, treated in Appendix A.9.

Figure 6(a) plots the main-text illustration's responses. Figure 14(c) shows the joint-long-holdings responses just derived.

Taxable counterpart. Proposition 2 also uses the taxable version of (71). With $\xi_{ij} = 1$ at taxable coordinates, its tax term is one. Its bond-specific term vanishes in three cases: a singleton taxable active set, since clearing fixes the sole holder's quantity; equal active taxable after-tax shares, which factor out of (51); or $v_T^C = 0$. Heterogeneous after-tax shares with multiple holders prevent cancellation by unweighted clearing. With the lower-tax investor the sole holder of the taxable bond,

$$s_T^C = 1 + \gamma_I \sigma_r^2 A_{T-1}^C \partial_z \hat{D}_I. \quad (74)$$

Proposition 2 supplies the signs in (16) and the positive response gap in (80).

Proposition 2 (Joint-ownership case with bond-specific risk). *At a regular state with $r(z) > 0$ in the two-investor affine benchmark, suppose the lower-tax investor alone holds risky taxable bonds. Let $\emptyset \neq J \subseteq \{2, \dots, T_{\max}\}$ be the fixed set of municipal maturities held by both investors, with the high-tax household alone holding every municipal maturity outside J , including the riskless one-period claim. Assume $A_{T-1}^M, A_{T-1}^C > 0$ for every $T \geq 2$ and $v_T^M > 0$ on J . Define*

$$P_J = \sum_{T \in J} \frac{A_{T-1}^M}{v_T^M}, \quad S_J = \sum_{T \in J} \frac{(A_{T-1}^M)^2}{v_T^M}, \quad d := \frac{\partial \hat{D}_H}{\partial z}.$$

Then

$$d = \frac{(\xi_I - \xi_H)m_I P_J}{(\gamma_H m_I + \gamma_I m_H)(1 + \sigma_r^2 S_J)} > 0, \quad \frac{\partial \hat{D}_I}{\partial z} = -\frac{m_H}{m_I} d. \quad (75)$$

Writing $\omega_i = (m_i/\gamma_i)/(m_H/\gamma_H + m_I/\gamma_I)$, $\xi^* = \omega_H \xi_H + \omega_I \xi_I$ and $b = \gamma_I \sigma_r^2 (m_H/m_I)d$, the responses are

$$s_T^M = \begin{cases} \xi_H + \gamma_H \sigma_r^2 A_{T-1}^M d, & T \notin J, \\ \xi^*, & T \in J, \end{cases} \quad s_T^C = 1 - bA_{T-1}^C. \quad (76)$$

Hence $s_T^M - \xi_H s_T^C > 0$ and $A_T^M - \xi_H A_T^C > 0$ at every risky maturity $T \geq 2$, with equality in both comparisons at $T = 1$.

Proof. Clearing fixes sole-held quantities at Q_T^M/m_H for the household at maturities outside J and Q_T^C/m_I for the lower-tax investor's risky taxable bonds. Only jointly held municipal positions vary. Differentiating

clearing and summing against municipal predecessor loadings, with constant liabilities and the fixed taxable book, gives

$$m_H \partial_z \hat{D}_H + m_I \partial_z \hat{D}_I = 0. \quad (77)$$

At each $T \in J$, subtracting the lower-tax investor's differentiated municipal equality from the household's gives

$$\xi_I - \xi_H = \sigma_r^2 A_{T-1}^M (\gamma_H \partial_z \hat{D}_H - \gamma_I \partial_z \hat{D}_I) + v_T^M (\gamma_H \partial_z x_{H,T}^M - \gamma_I \partial_z x_{I,T}^M).$$

With $k = (\gamma_H m_I + \gamma_I m_H)/m_H$, clearing reduces this to $\partial_z x_{I,T}^M = -(\xi_I - \xi_H)/(k v_T^M) - \sigma_r^2 A_{T-1}^M \partial_z \hat{D}_I / v_T^M$. Since $\partial_z \hat{D}_I = \sum_{T \in J} A_{T-1}^M \partial_z x_{I,T}^M$, multiplying by A_{T-1}^M and summing over J gives

$$(1 + \sigma_r^2 S_J) \partial_z \hat{D}_I = -\frac{\xi_I - \xi_H}{k} P_J < 0.$$

Equation (77) gives (75). The responses follow from (71). Household-only quantities have zero derivatives, giving the first municipal response. At jointly held maturities, risk-tolerance aggregation cancels the aggregate exposure and bond-specific derivatives and leaves ξ^* , and (74) gives the taxable formula.

Equation (76) gives the response gap $\partial_z g_T = s_T^M - \xi_H s_T^C$ as

$$\partial_z g_T = \begin{cases} \sigma_r^2 d \left[\gamma_H A_{T-1}^M + \xi_H \gamma_I \frac{m_H}{m_I} A_{T-1}^C \right], & T \notin J, \\ \omega_I (\xi_I - \xi_H) + \xi_H b A_{T-1}^C, & T \in J. \end{cases}$$

Thus $\partial_z g_T > 0$ for every $T \geq 2$, while $\partial_z g_1 = 0$ because both predecessor loadings vanish. Iterating (56), with $0 < \rho < 1$, yields

$$A_T^M - \xi_H A_T^C = \sum_{j=2}^T \rho^{T-j} \partial_z g_j > 0, \quad T \geq 2.$$

□

Combining the two legs gives the response gap $s_j^M - \xi_H s_j^C = \partial_z g_j$, whose persistence-weighted sum over $j < T$ is the loading difference $A_{T-1}^M - \xi_H A_{T-1}^C$ in the rate-loading channel of (66). Applying the decomposition of Proposition A.4 in both markets, with the lower-tax investor the sole holder of risky taxable bonds, gives

$$s_T^M - \xi_H s_T^C = \underbrace{(\bar{\xi}_T^M - \xi_H)}_{\text{difference in tax mix}} + \underbrace{\sigma_r^2 (A_{T-1}^M E_T^M - \xi_H A_{T-1}^C E_T^C)}_{\text{difference in portfolio-risk responses}}, \quad (78)$$

where $\bar{\xi}_T^M := \sum_i \omega_{i,T}^M \xi_i$ is the tax component of the active holders. It equals ξ_H , ξ^* or ξ_I under household-only, joint or lower-tax-only holdings. The second term combines both markets' portfolio responses. Below participation, adjustment raises the household's municipal response and lowers the taxable response. Joint municipal holders' exposure responses cancel in aggregate. A sole lower-tax holder instead has a negative municipal portfolio response when its net exposure falls. The joint-long-holdings branch with bond-specific risk therefore adds $\xi^* - \xi_H$ from the threshold onward, while the main-text lower-tax-only block adds $\xi_I - \xi_H$ together with its portfolio offset. Figures 6 and 14 combine the municipal tax and portfolio components and show the taxable contribution separately. All contributions vanish at the riskless one-period claim. Equation (70) gives the exact counterpart with nonlinear covariances.

On an upper-interval branch, the rate-loading channel of Proposition 1 incorporates portfolio-risk responses from three periods on and the change in tax mix from $T^* + 1$ on.

The empirical treasury comparison in Section 6.3 divides municipal loadings by $G_T := \sum_{j=1}^T \rho^{T-j} = A_T^{\text{TSY}}$, the loading of a reference whose one-period return responds one-for-one to the rate state. Since $A_T^M / G_T = \sum_{j=1}^T (\rho^{T-j} / G_T) s_j^M$, the relative yield sensitivity averages municipal responses with positive weights. It equals ξ_H at the household-only allocation. Its difference from ξ_H reflects municipal holders' tax treatment and portfolio responses, both of which depend on participation. Figure 6(b) instead uses the equilibrium taxable market C , whose own portfolio response also contributes to $A_T^M - \xi_H A_T^C$.

A per-dollar benefit or cost that does not vary with the rate state, such as a trading-cost advantage or the calibration's fixed bank financing-tax cost, drops out of these response formulas on the same participation branch. Such a benefit or cost still changes which maturities an investor holds and the levels of one-period returns, as shown next.

State-independent holding benefits. Suppose an investor's objective includes $\sum_{a,T} \vartheta_{i,T}^a x_{i,T}^a$, with each $\vartheta_{i,T}^a$ independent of the rate state. Its active municipal condition becomes

$$\mu_T^M - \xi_i r + \vartheta_{i,T}^M = \gamma_i \left[\sigma_r^2 A_{T-1}^M \hat{D}_i + v_T^M x_{i,T}^M \right],$$

and the taxable condition similarly gains $\vartheta_{i,T}^C/\xi_i$. With covariance, financing and constraints fixed, differentiation removes state-independent benefits and costs. The preceding proof and loading equations hold for the same feasible set J of jointly held maturities, so a given loading solution is unchanged on the same participation branch. Benefits and costs affect return intercepts, gap levels (Appendix A.11) and the feasible support. State-dependent benefits add their derivatives. Fixed participation fees affect global participation without a per-unit KKT term.

Uniqueness of the loading fixed point. Appendix A.10 verifies the following sufficient condition for a unique loading fixed point in the illustration. For given d , let $\mathcal{F}_J(d)$ be the right side of (75) at the municipal loadings that (76) and the loading recursion (56) generate. If $\mathcal{F}_J$ maps $[0, \infty)$ into $[0, d_{\max}]$ and $\sup_{[0, d_{\max}]} |\mathcal{F}'_J| < 1$, every nonnegative exposure fixed point lies in $[0, d_{\max}]$, and Banach gives exactly one, with its loading path on this branch. Holdings and inactive inequalities still require verification. Equilibria with other feasible participation thresholds remain possible. The derivative of $\mathcal{F}_J$ is

$$\mathcal{F}'_J(d) = c \frac{P'_J(1 + \sigma_r^2 S_J) - \sigma_r^2 P_J S'_J}{(1 + \sigma_r^2 S_J)^2}, \quad c = \frac{(\xi_I - \xi_H)m_I}{\gamma_H m_I + \gamma_I m_H},$$

with the derivatives of P_J, S_J obtained by differentiating the same loading recursions.

The last result in this subsection gives the household-only benchmark. Without participation by the lower-tax investor, municipal loadings are exactly ξ_H times taxable loadings.

Proposition A.6 (Household-only loadings). *In the two-investor affine benchmark, if the high-tax household alone holds every municipal bond and the lower-tax investor alone holds every taxable bond, then $s_T^M = \xi_H$, $s_T^C = 1$ and $A_T^M = \xi_H A_T^C$ at every maturity.*

Proof. Clearing fixes every holding, so both net rate exposures have zero derivatives, and (71) and (56) give the result. The identity concerns loadings, and different intercept risk premia can still affect the yield ratio. $\square$

A.7 From one-period returns to the yield-implied tax

The channels act on one-period returns, while the municipal puzzle concerns yields to maturity. This subsection carries the channels into yields, the formal counterpart of the yield relation (18) in Section 4.5.

Applying (43) to both markets turns the one-period gap g_T of Proposition 1 into the yield gap:

$$\begin{aligned} y_T^M - \xi_H y_T^C &= \underbrace{\frac{1}{T} \sum_{j=0}^{T-1} \mathbb{E}_t [g_{T-j}(z_{t+j})]}_{\text{average gap along the bond's life}} \\ &\quad - \underbrace{\frac{1}{2T} \sum_{k=2}^T \left[\sigma_r^2 \left((A_{k-1}^M)^2 - \xi_H (A_{k-1}^C)^2 \right) + (\sigma_{\eta,k}^M)^2 - \xi_H (\sigma_{\eta,k}^C)^2 \right]}_{\text{convexity difference}}. \end{aligned} \tag{79}$$

The riskless one-period gap is zero. The yield-implied tax is

$$\tau_T^{Y,\text{impl}} := 1 - \frac{y_T^M}{y_T^C}, \quad \text{the model counterpart of } \textit{ImpliedTax} \text{ in (2).}$$

It equals the household wedge less the yield gap per unit of taxable yield, $\tau_T^{Y,\text{impl}} = (1 - \xi_H) - (y_T^M - \xi_H y_T^C) / y_T^C$, as in the identity (3) of Section 3.6. A positive gap accumulated in (79) pulls the implied tax below the household wedge, by more at longer maturities when the gap grows faster in proportion than the taxable yield (Proposition A.7(ii)).

The relative yield response sums the response gaps in (78). On a joint-participation branch the response gap is positive at every risky maturity, through portfolio responses below the participation threshold and through both the tax mix and portfolio responses from that threshold on (Proposition 2), so for $T \geq 2$

$$\underbrace{\frac{\partial y_T^M}{\partial z} - \xi_H \frac{\partial y_T^C}{\partial z}}_{\text{relative yield response}} = \frac{1}{T} \sum_{j=2}^T \rho^{T-j} \underbrace{\left(\frac{\partial \mu_j^M}{\partial z} - \xi_H \frac{\partial \mu_j^C}{\partial z} \right)}_{\text{holding-return response gap}} > 0, \quad (80)$$

in model-period units, and both sides are zero for the riskless one-period bond. Long municipal yields therefore move more with the short rate than the household benchmark implies. The lower-tax investor's tax affects these responses even at maturities it does not hold.

Proposition A.7 uses one restriction on taxes, risk-tolerance shares and rate persistence:

$$(S) \quad \omega_H \frac{\xi_I - \xi_H}{\xi_H} \leq \rho. \quad (81)$$

Equation (86) writes the same condition in terms of masses and risk aversions. This condition keeps the feedback b below ρ , so taxable loadings rise with maturity. Municipal loadings also rise beyond the participation threshold because the loading there lies below its long-end value, $A_{T^*-1}^M < \xi^* / (1 - \rho)$ (equation (87)). The long-end taxable response $s_\infty^{C,\text{aff}} := \lim_{T \rightarrow \infty} s_T^{C,\text{aff}}$ lies in $(0, 1)$, because the lower-tax investor's net rate exposure falls as the rate state rises.

For the yield levels in Proposition A.7(ii), define the forward rate implied by today's prices for the period from $t + T - 1$ to $t + T$:

$$f_T^a(z) = \log \frac{P_{T-1}^a(z)}{P_T^a(z)}, \quad y_T^a(z) = \frac{1}{T} \sum_{j=1}^T f_j^a(z). \quad (82)$$

With positive taxable forwards, the yield ratio is a weighted average of municipal-to-taxable forward-rate ratios. The forward rate compares adjacent-maturity prices at the same date, whereas the holding return uses next period's resale price. The affine recursions (42) give their relation:

$$f_T^a(z) = \mu_T^a(z) - (1 - \rho) A_{T-1}^a z - \frac{1}{2} [\sigma_r^2 (A_{T-1}^a)^2 + (\sigma_{\eta,T}^a)^2]. \quad (83)$$

For taxable bonds, the term premium is the yield less the average expected future short rate. Equation (43) expresses it as average expected bond risk premia less convexity. The municipal-taxable yield gap also reflects differences in holders' tax benefits.

Proposition A.7 (Maturity ordering of the loading ratio and the yield-implied tax). *In the two-investor affine benchmark, consider the equilibrium of Proposition 2 in which the high-tax household alone holds municipal bonds shorter than T^* , both investors hold municipal bonds from T^* to $T_{\max}$, and the lower-tax investor holds the taxable bonds. Assume condition (S) in (81).*

(i) Rate sensitivity. *The municipal-to-taxable loading ratio is nondecreasing in maturity. It starts at the household's after-tax share and, when the recursions are extended to infinite maturity with the household's exposure response d held at its solved value, converges to the long-end tax mix relative to the taxable response:*

$$\frac{A_1^{M,\text{aff}}}{A_1^{C,\text{aff}}} = \xi_H \quad \text{and} \quad \lim_{T \rightarrow \infty} \frac{A_T^{M,\text{aff}}}{A_T^{C,\text{aff}}} = \frac{\xi^*}{s_\infty^{C,\text{aff}}}. \quad (84)$$

(ii) **Yield levels.** Suppose also that taxable forward rates are positive and the municipal-to-taxable forward-rate ratio is nondecreasing in maturity at the stationary mean $z = 0$. Then the yield ratio y_T^M/y_T^C is nondecreasing at $z = 0$ (Proposition A.8), so the yield-implied tax $\tau_T^{Y,\text{impl},\text{aff}}$ is nonincreasing.

(iii) **Long end.** With the recursions extended to infinite maturity, suppose the loadings, the risk-premium coefficients π_T^a and the bond-specific compensation terms ι_T^a converge, with limits marked by the subscript ∞ and a positive limiting taxable forward. Then

$$\lim_{T \rightarrow \infty} \tau_T^{Y,\text{impl},\text{aff}} = 1 - \frac{\xi^* \bar{r} + A_\infty^M \pi_\infty^M - \frac{1}{2} \sigma_r^2 (A_\infty^M)^2 + \iota_\infty^M}{\bar{r} + A_\infty^C \pi_\infty^C - \frac{1}{2} \sigma_r^2 (A_\infty^C)^2 + \iota_\infty^C}, \quad (85)$$

where $\pi_T^a = \sum_i \omega_{i,T}^a \gamma_i \sigma_r^2 \widehat{D}_i(0)$ is the holders' cost of rate risk per unit of loading, weighted by risk tolerance, and ι_T^a is their compensation for bond-specific risk net of its convexity term (equation (88)).

The proof below also gives the treasury-reference ordering in Proposition A.9(iii).

The sufficient condition and the threshold loading. For Proposition 2 with the upper interval $J = \{T^*, \dots, T_{\max}\}$, the maturity ordering uses condition (S) and the bound (M) on the loading at the threshold:

$$(S) \quad \frac{\gamma_I m_H (\xi_I - \xi_H)}{\xi_H (\gamma_H m_I + \gamma_I m_H)} \leq \rho, \quad (86)$$

$$(M) \quad A_{T^*-1}^M < \frac{\xi^*}{1 - \rho}. \quad (87)$$

Condition (S) is condition (81) above with ω_H written out and is assumed. The bound (M) holds at every solution of (75) on this branch. Put $n = T^* - 1$ and $\lambda = \gamma_H \sigma_r^2 d$, so that (75) reads $\lambda = \omega_I (\xi_I - \xi_H) \sigma_r^2 P_J / (1 + \sigma_r^2 S_J)$. The recursions at household-only and jointly held maturities give $A_n^M = \xi_H \sum_{q=0}^{n-1} (\rho + \lambda)^q$ and $A_{n+l}^M = \rho^l A_n^M + (1 - \rho^l) \xi^* / (1 - \rho)$ for $l \geq 0$. If $A_n^M \geq \xi^* / (1 - \rho)$, every loading on J is at least $\xi^* / (1 - \rho)$, so $S_J / P_J \geq \xi^* / (1 - \rho)$ and $\lambda < (1 - \rho) (\xi^* - \xi_H) / \xi^* < 1 - \rho$, using $\xi^* - \xi_H = \omega_I (\xi_I - \xi_H)$. Then $A_n^M < \xi_H / (1 - \rho - \lambda) < \xi^* / (1 - \rho)$, a contradiction. Both statements concern the upper-interval branch of the two-investor affine benchmark and do not prove that this participation branch exists.

Proof of Proposition A.7(i). Step 1 concerns municipal loadings. At household-only maturities, $A_T^M = (\rho + \gamma_H \sigma_r^2 d) A_{T-1}^M + \xi_H$, so A_T^M is increasing from zero. At jointly held maturities, $A_T^M = \rho A_{T-1}^M + \xi^*$, whose increments have the sign of $\xi^* - (1 - \rho) A_{T-1}^M$. Condition (M) makes them positive. Step 2 concerns the taxable leg. Since $A_{T-1}^M \geq \xi_H$ on J , $P_J / S_J \leq 1 / \xi_H$, and (75) and (S) give $0 \leq b \leq \rho$. Therefore

$$A_T^C = \sum_{q=0}^{T-1} (\rho - b)^q, \quad s_T^C = 1 - b A_{T-1}^C \downarrow \frac{1 - \rho}{1 - \rho + b} > 0.$$

Step 3 concerns the municipal response. The municipal response rises at household-only maturities, is constant at jointly held maturities and does not jump down when the lower-tax investor begins to participate. For the last claim, the response at $T^* - 1$ must not exceed ξ^* . Substituting (75) shows that the required inequality $\omega_I (\xi_I - \xi_H) \geq \gamma_H \sigma_r^2 A_{T^*-2}^M d$ is equivalent to $\sigma_r^2 A_{T^*-2}^M P_J \leq 1 + \sigma_r^2 S_J$, which follows from $A_{T-1}^M > A_{T^*-2}^M$ for every $T \in J$, established above using (M). Since Step 2 makes s_j^C positive and nonincreasing, the response ratio s_j^M / s_j^C is nondecreasing in maturity.

Step 4 is an averaging argument. For any positive e_j and nondecreasing ratios $q_j = a_j / e_j$, the ratio weighted by rate persistence,

$$\bar{q}_T = \frac{\sum_{j \leq T} \rho^{T-j} a_j}{\sum_{j \leq T} \rho^{T-j} e_j}$$

is nondecreasing: as a positive-weight average of $q_1, \dots, q_T$, it is at most $q_T \leq q_{T+1}$, and

$$\bar{q}_{T+1} - \bar{q}_T = \frac{e_{T+1} (q_{T+1} - \bar{q}_T)}{\rho \sum_{j \leq T} \rho^{T-j} e_j + e_{T+1}} \geq 0.$$

Applying this with $a_j = s_j^M$, $e_j = s_j^C$ and (56), which gives these weights in A_T^M/A_T^C , proves the loading-ratio ordering, starting from ξ_H .

Step 5 gives the limits and the treasury reference. Extending the same branch recursions to arbitrarily long maturities with fixed d and eventual joint participation gives the limits $A_\infty^M = \xi^*/(1 - \rho)$ and $A_\infty^C = 1/(1 - \rho + b)$, hence the endpoint in (84). Taking $e_j = 1$, for which (M) suffices, proves by the same argument that A_T^M/G_T is nondecreasing, where $G_T = \sum_{j \leq T} \rho^{T-j}$ is the treasury-reference loading of Appendix A.6. Thus the same ordering holds for the empirical treasury comparison of Section 6.3. $\square$

Part (ii) of Proposition A.7 rests on the following bridge.

Proposition A.8 (Exact forward-ratio bridge). *For any price functions, affine or nonlinear, let $f_T^a(z) = -p_T^a(z) + p_{T-1}^a(z)$. If $f_T^C(z) > 0$ and $f_T^M(z)/f_T^C(z)$ is nondecreasing over the maturities considered, then $y_T^M(z)/y_T^C(z)$ is nondecreasing and $\tau_T^{Y, \text{impl}}(z)$ is nonincreasing.*

Proof. Telescoping gives $Ty_T^a = -p_T^a = \sum_{j \leq T} f_j^a$. The yield ratio averages forward ratios with positive weights $f_j^C > 0$. Adding a ratio at least as large as its predecessors cannot lower the average. On a two-year grid, T counts steps and annualization cancels. $\square$

The premise concerns full forwards, including bond-specific and convexity terms, so loading monotonicity alone does not establish it. The nonlinear maturity profiles are checked numerically in Section 5.

Part (iii) adds relative term premia, convexity and bond-specific compensation to the long-end tax mix. With $\hat{\Lambda} > 0$, small taxable feedback b and negligible convexity and bond-specific terms, the long-end implied tax is, to first order, below $1 - \xi^*$ and hence below $1 - \xi_H$ (equations (90) and (91)).

Affine forward levels and Proposition A.7(iii). For coordinate $j = (a, T)$, put $\Gamma_T^a := \sum_{i \in \mathcal{U}_j} m_i/\gamma_i$ and

$$\bar{\xi}_T^a = \bar{\xi}_j = \frac{\sum_i (m_i/\gamma_i) \xi_{ij}}{\Gamma_T^a}, \quad \pi_T^a = \frac{\sigma_r^2}{\Gamma_T^a} \sum_i m_i \hat{D}_i(0), \quad \iota_T^a = v_j \left[\frac{\sum_i m_i \xi_i^a x_{i,T}^a(0)}{\Gamma_T^a} - \frac{1}{2} \right],$$

where sums run over active holders. Here π_T^a is the holders' cost of rate risk per unit of loading and ι_T^a their compensation for bond-specific risk net of its convexity term. Aggregating their level KKT equations at $z = 0$ and subtracting the two Jensen terms in the exact Gaussian mean gives

$$C_T^a - C_{T-1}^a = \bar{\xi}_T^a \bar{r} + A_{T-1}^a \pi_T^a - \frac{1}{2} \sigma_r^2 (A_{T-1}^a)^2 + \iota_T^a. \quad (88)$$

For municipal bonds the weighted holding sum is Q_T^M . For taxable bonds heterogeneous after-tax shares remain inside it. The increment $C_T^a - C_{T-1}^a$ is the forward rate at $z = 0$. If the displayed coefficients and loadings converge and the limiting taxable forward f_∞^C is positive, Cesàro convergence gives

$$y_T^a(z) \longrightarrow f_\infty^a = \bar{\xi}_\infty^a \bar{r} + A_\infty^a \pi_\infty^a - \frac{1}{2} \sigma_r^2 (A_\infty^a)^2 + \iota_\infty^a \quad (89)$$

at fixed z , because C_T^a/T converges to the forward limit and $A_T^a z/T \rightarrow 0$. At jointly held long maturities, $\bar{\xi}_\infty^M = \xi^*$, and $\bar{\xi}_\infty^C = 1$ because relative after-tax shares on taxable bonds equal one, so the implied-tax limit $1 - f_\infty^M/f_\infty^C$ is (85). This proves part (iii).

For the first-order approximation, aggregating by risk tolerance over both investors for long municipal bonds and over the lower-tax investor alone for long taxable bonds gives

$$\pi_\infty^M - \pi_\infty^C = \omega_H \sigma_r^2 \hat{\Lambda}, \quad \hat{\Lambda} = \gamma_H \hat{D}_H - \gamma_I \hat{D}_I, \quad (90)$$

where $\hat{D}_i = \hat{D}_i(0)$ is the net rate exposure at the stationary mean. The premium difference does not bound the convexity and bond-specific terms. With small taxable loading feedback ($b \ll 1 - \rho$), negligible convexity

and bond-specific terms and a positive denominator, substituting $A_\infty^M = \xi^*/(1 - \rho)$, $A_\infty^C \approx 1/(1 - \rho)$ and (90) into (89) gives the first-order endpoint

$$\tau_\infty^{Y,\text{impl},\text{aff}} \approx \underbrace{(1 - \xi^*)}_{\text{one minus tax mix}} - \underbrace{\frac{\xi^* \omega_H \sigma_r^2 \hat{\Lambda}}{(1 - \rho)\bar{r} + \pi_\infty^C}}_{\text{municipal term-premium excess}}. \quad (91)$$

With $\hat{\Lambda} > 0$, this first-order endpoint lies below one minus the tax mix, $1 - \xi^*$, and hence below the one-period implied tax $1 - \xi_H$.

A.8 From mechanism to empirical predictions

The mechanism predicts whose taxes are priced at each maturity. This subsection gives sufficient conditions on the branch with bond-specific risk and joint long-maturity holdings. Appendix A.9 treats the main-text illustration separately. P1 and P2 concern yield levels, while P3 concerns rate sensitivities. To keep these objects distinct, I pair the yield-implied tax of Appendix A.7 with a loading-implied tax:

$$\tau_T^{Y,\text{impl}} = 1 - \frac{y_T^M}{y_T^C}, \quad \tau_T^A = 1 - \frac{A_T^M}{A_T^C}.$$

To measure pass-through, I vary one investor's after-tax share at a fixed rate state and re-solve the equilibrium. Holdings, risk costs and both yield curves adjust while non-tax primitives stay fixed. The pass-throughs of investor i 's exemption benefit $1 - \xi_i$ into the two implied taxes are

$$\mathcal{P}_i^Y(T) := \frac{d\tau_T^{Y,\text{impl}}}{d(1 - \xi_i)}, \quad \mathcal{P}_i^A(T) := \frac{d\tau_T^A}{d(1 - \xi_i)}, \quad i \in \{H, I\},$$

where an argument ∞ denotes the limit as $T \rightarrow \infty$. Under the representative-tax benchmark of Section 3.3, $y_T^M = \xi_H y_T^C$ and $A_T^M = \xi_H A_T^C$ at every maturity, so in both measures household pass-through is one and the lower-tax investor's pass-through is zero at every maturity. With heterogeneous taxes, the household alone holds the riskless short bond, whose return is $\xi_H r$. Both investors hold long bonds, whose returns depend on both after-tax shares through the tax mix ξ^* . For the empirical treasury comparison of Section 6.3, I also use a taxable reference whose one-period return moves one-for-one with the rate state, giving loading $A_T^{\text{tsy}} = G_T$ (Appendix A.6).

The long-end yield inequalities in (21) hold in the main-text illustration. Proposition A.9 gives sufficient conditions for the same inequalities on the joint-long-holdings branch with positive bond-specific risk.

Proposition A.9 (Tax pass-through at the short and long ends). *Consider the equilibrium of Proposition A.7 (Proposition 2), with $0 < \rho < 1$, $\sigma_r > 0$, a positive cash rate and a slack household shorting constraint on cash. With its recursions extended to infinite maturity, suppose long yields and their tax derivatives converge locally uniformly, with a positive long taxable yield.*

- (i) (Short end) Both implied taxes equal the household wedge, $\tau_1^{Y,\text{impl}} = \tau_1^A = 1 - \xi_H$, so $\mathcal{P}_H^Y(1) = \mathcal{P}_H^A(1) = 1$ and $\mathcal{P}_I^Y(1) = \mathcal{P}_I^A(1) = 0$.
- (ii) (Long end, P1 and P2) Suppose the after-tax shares satisfy the tax-difference condition

$$2\omega_H e + \omega_I e^2 \leq 1, \quad e := \frac{\xi_I - \xi_H}{\xi_H}.$$

Then

$$\mathcal{P}_H^A(\infty) < \omega_H, \quad \mathcal{P}_I^A(\infty) > \omega_I.$$

The difference $\mathcal{P}_i^Y(\infty) - \mathcal{P}_i^A(\infty)$ shrinks to zero as the mean rate $\bar{r}$ rises, so for every $\bar{r}$ above a bound set by the other primitives, and at every feasible participation threshold,

$$\mathcal{P}_H^Y(\infty) < \omega_H < 1, \quad \mathcal{P}_I^Y(\infty) > \omega_I > 0.$$

(iii) (Response ratios, P3) The municipal-to-taxable zero-coupon response ratio A_T^M/A_T^C is nondecreasing in maturity, equals ξ_H at the short end and converges to ξ^*/s_∞^C , with $0 < s_\infty^C < 1$. For the unit-response reference, A_T^M/A_T^{Tsy} has the same ordering and short endpoint, and converges to $\xi^* > \xi_H$.

Part (ii) first separates the holders' tax mix from portfolio adjustment. The tax mix alone passes each investor's after-tax share through with its risk-tolerance share, ω_H or ω_I . Under the tax-difference condition, portfolio adjustment offsets the household's tax-mix effect and reinforces the lower-tax investor's, giving $\mathcal{P}_H^A(\infty) < \omega_H$ and $\mathcal{P}_I^A(\infty) > \omega_I$. Lemma 1 proves these bounds while allowing loadings and portfolios to adjust.

Raising ξ_H narrows the tax difference, which weakens the rate-induced transfer of net rate exposure to the household. It also raises the tax mix and bond loadings, which strengthens the risk feedback. The first effect dominates when the tax difference is not too large relative to ξ_H . The condition bounds this relative difference, e , and holds whenever $\xi_I < \frac{3}{2}\xi_H$, whatever the risk-tolerance shares or the participation threshold. Lemma 1 proves the derivative bounds with both effects included.

For yield levels, write $R_\infty = A_\infty^M/A_\infty^C$. The proof shows that $\mathcal{P}_i^Y(\infty)$ equals $\mathcal{P}_i^A(\infty)$ plus two corrections divided by the long taxable yield. The first is the tax response of $y_\infty^M - R_\infty y_\infty^C$, which contains only relative risk costs, convexity and bond-specific compensation and does not depend on $\bar{r}$. The second reflects the tax response of the taxable yield itself. The taxable yield rises with $\bar{r}$, so both corrections shrink as $1/\bar{r}$, which gives the yield-level inequalities in part (ii) when the mean rate is high enough.

The tax-difference condition accommodates the tax difference between a top-bracket household and a corporate investor. The mean-rate condition limits how much risk compensation can move relative to interest income. The tax-difference condition and the yield-level inequalities hold with wide margins in the illustration with bond-specific risk in Appendix A.10, where convexity accounts for most of the difference between the lower-tax investor's two pass-throughs. In that illustration, the yield-level inequalities also hold for every nonnegative value of the taxable holders' cost per unit of rate loading, π_∞^C . Stronger hedging demand or lower mean rates can enlarge the corrections as the long taxable yield in their denominator approaches zero.

The proposition establishes the short-end and long-end contrasts that the empirical tests use. At intermediate maturities, the part of $\mathcal{P}_i^A(T)$ that comes from the tax mix alone shifts weight from the household to the lower-tax investor as more of a bond's remaining life falls at jointly held maturities (Proposition A.10). In the illustration with bond-specific risk, $\mathcal{P}_H^Y(T)$ falls and $\mathcal{P}_I^Y(T)$ rises strictly with maturity. Appendix A.9 verifies the same finite-maturity ordering for the main-text illustration. With more clienteles, each has its own participation threshold and risk-tolerance share (Appendix A.13). Section 4.6 states the resulting P1–P3 predictions, and Section 6 tests them.

Pass-throughs are the derivatives of y_T^M/y_T^C and A_T^M/A_T^C with respect to ξ_i .

Proof of Proposition A.9.

(i) The riskless short end. The one-period claims have zero variance. With a slack household shorting constraint on cash, Proposition A.2 gives $y_1^M = \xi_H r$ and $y_1^C = r$ at every rate state with $r > 0$. The one-period loadings equal the responses $s_1^M = \xi_H$ and $s_1^C = 1$ by (42), so $y_1^M/y_1^C = A_1^M/A_1^C = \xi_H$. Differentiating the two ratios gives $\mathcal{P}_H^Y(1) = \mathcal{P}_H^A(1) = 1$ and $\mathcal{P}_I^Y(1) = \mathcal{P}_I^A(1) = 0$.

(ii) The long end. Let d be the household's exposure response in (75) and $b = \gamma_I \sigma_r^2(m_H/m_I)d$ the feedback into taxable responses, as in (76). The solved long responses and loadings satisfy

$$s_\infty^M = \xi^*, \quad s_\infty^C = \frac{1 - \rho}{1 - \rho + b}, \quad A_\infty^a = \frac{s_\infty^a}{1 - \rho}.$$

Thus $R_\infty = A_\infty^M/A_\infty^C = \xi^*/s_\infty^C$. At the mean state, the solved yield formula (89) gives

$$\begin{aligned} y_\infty^M &= \xi^* \bar{r} + A_\infty^M \pi_\infty^M - \frac{1}{2} \sigma_r^2 (A_\infty^M)^2 + \iota_\infty^M, \\ y_\infty^C &= \bar{r} + A_\infty^C \pi_\infty^C - \frac{1}{2} \sigma_r^2 (A_\infty^C)^2 + \iota_\infty^C. \end{aligned}$$

The four terms are interest income, portfolio rate-risk compensation, convexity and bond-specific compensation. Dividing the two yields separates the loading ratio from the remaining yield-level difference:

$$\frac{y_\infty^M}{y_\infty^C} = R_\infty + \frac{y_\infty^M - R_\infty y_\infty^C}{y_\infty^C}.$$

All derivatives below are with respect to ξ_i , with holdings, loadings and both yields adjusting. The assumed convergence permits differentiation of the limiting ratio. The quotient rule gives

$$\begin{aligned} \mathcal{P}_i^Y(\infty) &= \mathcal{P}_i^A(\infty) + \frac{1}{y_\infty^C} \frac{d(y_\infty^M - R_\infty y_\infty^C)}{d\xi_i} \\ &\quad - \frac{y_\infty^M - R_\infty y_\infty^C}{(y_\infty^C)^2} \frac{dy_\infty^C}{d\xi_i}. \end{aligned} \tag{92}$$

The three terms are the loading-ratio response, the response of relative yield compensation and the response of the taxable denominator. The first retains the endogenous taxable response:

$$\mathcal{P}_i^A(\infty) = \frac{\omega_i}{s_\infty^C} - \frac{\xi^*}{(s_\infty^C)^2} \frac{ds_\infty^C}{d\xi_i}.$$

Lemma 1 below proves $\mathcal{P}_H^A(\infty) < \omega_H$ and $\mathcal{P}_I^A(\infty) > \omega_I$ under the tax-difference condition, with uniform positive margins.

The calculation after the lemma shows that $y_\infty^M - R_\infty y_\infty^C$ and its tax derivatives are independent of $\bar{r}$. Also, $y_\infty^C = A_\infty^C(1 - \rho)\bar{r} + O(1)$ and $dy_\infty^C/d\xi_i = O(\bar{r})$. Both correction terms in (92) therefore vanish at rate $1/\bar{r}$. The loading and derivative bounds in the lemma are uniform over all feasible cutoffs, so a lower bound on $\bar{r}$ that depends only on the other primitives makes both yields positive and keeps the corrections smaller than the loading margins. This proves $\mathcal{P}_H^Y(\infty) < \omega_H < 1$ and $\mathcal{P}_I^Y(\infty) > \omega_I > 0$.

(iii) The yield-response ratios. At a common maturity, dividing both price loadings by T leaves their ratio unchanged. Proposition A.7(i) gives the ordering and endpoints for the equilibrium taxable market C : the ratio starts at ξ_H and converges to ξ^*/s_∞^C . For the unit-response reference of Appendix A.6, $A_T^{\text{Tsy}} = G_T$, with $A_1^{\text{Tsy}} = 1$ and $A_\infty^{\text{Tsy}} = 1/(1 - \rho)$. Its ratio therefore starts at ξ_H and converges to ξ^* . Appendix A.7 proves monotonicity for both comparisons. $\square$

Lemma 1 (The loading response). *Under the tax-difference condition in Proposition A.9, $\mathcal{P}_H^A(\infty) < \omega_H$ and $\mathcal{P}_I^A(\infty) > \omega_I$. With other primitives fixed, both strict margins have positive lower bounds over all feasible upper-interval cutoffs.*

Proof. The proof differentiates the exposure fixed point and bounds its feedback into the loading ratio. Put $n = T^* - 1$ and $\lambda = \gamma_H \sigma_r^2 d$, the response of the household's unit risk cost to the rate state. Equation (75) gives

$$\lambda = \omega_I(\xi_I - \xi_H)\sigma_r^2 f, \quad f = \frac{P_J}{1 + \sigma_r^2 S_J}, \quad R_\infty = \xi^* \left[1 + \frac{\omega_H \lambda}{\omega_I(1 - \rho)} \right].$$

The sums P_J and S_J are those of Proposition 2. With $G_n(p) = \sum_{q=0}^{n-1} p^q$, the loading recursion gives

$$A_n^M = \xi_H G_n(\rho + \lambda), \quad A_{n+l}^M = \rho^l A_n^M + (1 - \rho^l) A_\infty^M, \quad l \geq 0.$$

By (87), $A_n^M < A_\infty^M$, so loadings rise along the jointly held block. Thus $S_J/P_J \geq A_n^M$ and

$$\lambda G_n(\rho + \lambda) < \omega_I e, \quad \lambda \frac{G'_n(\rho + \lambda)}{G_n(\rho + \lambda)} < \omega_I e,$$

where e is the relative tax difference of Proposition A.9. The first bound uses $\lambda < \omega_I(\xi_I - \xi_H)P_J/S_J$, and the second follows from $G'_n \leq G_n^2$. The tax-difference condition implies $e < 1$ and hence $b = (\omega_H/\omega_I)\lambda < \omega_H e < 1$, ensuring convergence of the taxable loadings.

To differentiate the loading sums, let $t_j = \rho^l A_n^M / A_{j-1}^M$ for $j = T^* + l$ and

$$X = \bar{t}_P - \frac{2\sigma_r^2 S_J}{1 + \sigma_r^2 S_J} \bar{t}_S.$$

Here $\bar{t}_P$ and $\bar{t}_S$ are the means of t_j with weights proportional to A_{j-1}^M / v_j^M and $(A_{j-1}^M)^2 / v_j^M$, respectively. Since $0 \leq t_j \leq 1$, $-2 < X \leq 1$. Let ∂_i denote differentiation with respect to ξ_i holding λ fixed. Differentiating the loadings gives

$$\begin{aligned} \partial_H \ln f &= \left(\frac{1}{\xi_H} - \frac{\omega_H}{\xi^*} \right) X + \frac{\omega_H}{\xi^*} \frac{1 - \sigma_r^2 S_J}{1 + \sigma_r^2 S_J}, \\ \partial_I \ln f &= \frac{\omega_I}{\xi^*} \left(\frac{1 - \sigma_r^2 S_J}{1 + \sigma_r^2 S_J} - X \right). \end{aligned}$$

The partial derivatives above hold λ fixed. To include its equilibrium response, differentiate its fixed point. With G_n and G'_n evaluated at $\rho + \lambda$,

$$\frac{d\lambda}{d\xi_i} = \frac{\lambda \partial_i \ln[(\xi_I - \xi_H)f]}{D}, \quad D = 1 - \lambda \frac{G'_n}{G_n} X > 1 - \omega_I e > 0.$$

Substitute into $dR_\infty/d\xi_i = \mathcal{P}_i^A(\infty)$. The coefficient of X is positive in the household derivative and negative in the lower-tax derivative. Using $X \leq 1$ and $\lambda G'_n/G_n < \omega_I e$ gives

$$\begin{aligned} \mathcal{P}_H^A(\infty) - \omega_H &\leq \frac{\omega_H \lambda}{\omega_I(1 - \rho)D} \left(\frac{2\omega_H}{1 + \sigma_r^2 S_J} - \frac{1}{e} + \omega_I e \right) < 0, \\ \mathcal{P}_I^A(\infty) - \omega_I &> \frac{\omega_H \lambda}{\omega_I(1 - \rho)D} (1/e - \omega_I^2 e) > 0. \end{aligned}$$

The tax-difference condition makes the household bracket at most $-2\omega_H \sigma_r^2 S_J / (1 + \sigma_r^2 S_J)$, giving a positive margin for each investor.

For uniformity, consider $n \in \{1, \dots, T_{\max} - 1\}$ in the finite loading system and trial feedbacks $x \geq 0$ with $xG_n(\rho + x) \leq \omega_I e$. This compact primitive domain contains every actual pair. Its positive loading sums give positive minima for f and $\sigma_r^2 S_J / (1 + \sigma_r^2 S_J)$, while $D \geq 1 - \omega_I e$ keeps full tax derivatives bounded. At a fixed point, $\lambda \geq \omega_I(\xi_I - \xi_H)\sigma_r^2 \min f > 0$ and $D < 1 + 2\omega_I e$. The displayed margins consequently have uniform positive lower bounds. $\square$

The relative yield compensation. Subtract the two solved long yields, weighting the taxable yield by $R_\infty = A_\infty^M / A_\infty^C$. Equation (90) gives $\pi_\infty^M - \pi_\infty^C = \omega_H \sigma_r^2 \hat{\Lambda}$, while $\xi^* - R_\infty = -bA_\infty^M$. Therefore

$$\begin{aligned} y_\infty^M - R_\infty y_\infty^C &= A_\infty^M (\omega_H \sigma_r^2 \hat{\Lambda} - b\bar{r}) \\ &\quad + \frac{1}{2} \sigma_r^2 A_\infty^M (A_\infty^C - A_\infty^M) + \iota_\infty^M - R_\infty \iota_\infty^C. \end{aligned}$$

The three terms are the difference in risk costs on the books each investor holds alone, relative convexity and bond-specific compensation. Subtract the lower-tax investor's municipal equality from the household's on the shared block and use clearing:

$$\begin{aligned} \omega_H \sigma_r^2 \hat{\Lambda} - b\bar{r} &= \frac{\omega_H \sigma_r^2}{1 + \sigma_r^2 S_J} \left[\gamma_H \left(\sum_{T \notin J} \frac{A_{T-1}^M Q_T^M}{m_H} - \ell_H \right) \right. \\ &\quad \left. - \gamma_I \left(\frac{\xi_I}{m_I} \sum_T A_{T-1}^C Q_T^C - \ell_I \right) \right]. \end{aligned}$$

Every term, and hence its tax derivative, is independent of $\bar{r}$ on the maintained branch. Common duration compensation per unit of loading cancels from $y_\infty^M - R_\infty y_\infty^C$.

The joint holders' risk-premium coefficient is also independent of $\bar{r}$. With $\Gamma := m_H/\gamma_H + m_I/\gamma_I$, the value of Γ_T^M on the jointly held block, clearing gives

$$\pi_\infty^M = \frac{\sigma_r^2}{\Gamma} \left(\sum_T A_{T-1}^M Q_T^M + \xi_I \sum_T A_{T-1}^C Q_T^C - m_H \ell_H - m_I \ell_I \right).$$

Since $1/A_\infty^C = 1 - \rho + b$, dividing the taxable yield by its loading cancels $b\bar{r}$ against the mean-rate term in π_∞^C , which by (90) and the two previous displays is $-b\bar{r}$. This establishes $y_\infty^C = A_\infty^C(1 - \rho)\bar{r} + O(1)$. The bond-specific terms are $\iota_\infty^M = v_\infty^M(Q_\infty^M/\Gamma - \frac{1}{2})$ and $\iota_\infty^C = v_\infty^C(\gamma_I \xi_I Q_\infty^C/m_I - \frac{1}{2})$.

Convexity in the long-end correction. For the illustration (Appendix A.10), dividing by y_∞^C isolates convexity:

$$\begin{aligned} \Delta\tau_\infty &:= \frac{y_\infty^M}{y_\infty^C} - R_\infty = R_\infty(1 - R_\infty)\chi_\infty + \psi_\infty, \\ \chi_\infty &= \frac{\frac{1}{2}\sigma_r^2(A_\infty^C)^2}{y_\infty^C}, \quad \psi_\infty = \frac{A_\infty^M(\omega_H\sigma_r^2\hat{\Lambda} - b\bar{r}) + \iota_\infty^M - R_\infty\iota_\infty^C}{y_\infty^C}. \end{aligned}$$

The signed convexity contribution is at most $\chi_\infty/4$, since $R_\infty(1 - R_\infty) \leq 1/4$. The derivatives of the two terms sum to $\mathcal{P}_i^Y(\infty) - \mathcal{P}_i^A(\infty)$ in (92).

The tax-mix pass-through across maturities. The following supporting result supplies the intermediate-maturity profile used above and in Appendix A.12.

Proposition A.10 (Tax-mix pass-through). *Switch off exposure-response terms in the affine pricing recursions, with support and risk-tolerance weights fixed. Put*

$$G_T = \sum_{j=1}^T \rho^{T-j}, \quad \Theta_S(T) = \frac{\sum_{j=1}^{\min(T, T^*-1)} \rho^{T-j}}{G_T}, \quad \Theta_L(T) = 1 - \Theta_S(T).$$

For the loading-implied tax τ_T^A defined above, the tax-mix pass-throughs $\mathcal{D}_i(T) := -\partial_{\xi_i} \tau_T^A$, the part of $\mathcal{P}_i^A(T)$ that remains with exposure responses switched off, are

$$\mathcal{D}_H(T) = \Theta_S(T) + \omega_H \Theta_L(T), \quad \mathcal{D}_I(T) = \omega_I \Theta_L(T). \quad (93)$$

They equal $(1, 0)$ before participation and converge monotonically to (ω_H, ω_I) .

Proof. With exposure responses switched off, the recursions give $A_T^C = G_T$ and

$$A_T^M = \xi_H \sum_{j=1}^{\min(T, T^*-1)} \rho^{T-j} + \xi^* \sum_{j=T^*}^T \rho^{T-j}.$$

Differentiating A_T^M/G_T gives (93). For $T \geq T^*$, $\Theta_S(T+1) = \Theta_S(T)\rho G_T/(\rho G_T+1) < \Theta_S(T)$ and $\Theta_S(T) \rightarrow 0$. $\square$

The pass-throughs average $(1, 0)$ and (ω_H, ω_I) , using the shares of household-only and jointly held maturities in the bond's remaining life weighted by rate persistence.

Price derivatives in the nonlinear model. On a differentiable nonlinear equilibrium branch, let θ be a tax parameter, such as an after-tax share ξ_i , and put $q_T^a = \partial_\theta p_T^a$, with the return derivative $\partial_\theta \mu_T^a$ including portfolio and covariance responses. Hold the state process, cash rule and bond-specific variances fixed. Under a common integrable envelope for the exponential moments and their derivatives, differentiation of the price recursion (45) gives

$$q_T^a = \mathcal{K}_T^a q_{T-1}^a - \partial_\theta \mu_T^a, \quad q_0^a = 0, \quad (\mathcal{K}_T^a q_{T-1}^a)(z) = \mathbb{E}^{\mathbb{Q}_{T,z}^a} q_{T-1}^a(Z'), \quad (94)$$

where F_z is the conditional law of Z' and $d\mathbb{Q}_{T,z}^a/dF_z = e^{p_{T-1}^a(Z')}/\mathbb{E}_z e^{p_{T-1}^a(Z')}$. The identity follows from $\partial_\theta \log \mathbb{E}_z e^{p_{T-1}^a} = \mathbb{E}^{\mathbb{Q}_{T,z}^a} q_{T-1}^a$. The full implied-tax response is

$$\partial_\theta \tau_T^{Y,\text{impl}} = \frac{q_T^M - (y_T^M/y_T^C)q_T^C}{T y_T^C}, \quad y_T^C > 0. \quad (95)$$

Both price responses therefore enter the implied tax measured in the data.

A.9 The main-text analytical illustration

The main-text figures have rate risk alone, with $\rho = 0.8$, $\sigma_r = 0.025$, $(\gamma_H, \gamma_I) = (8, 5)$, $(\xi_H, \xi_I) = (0.55, 0.78)$, equal unit masses, $\bar{r} = 0.04$, municipal supply 0.1 per maturity, taxable supply 0.01 at each $T \geq 2$, $\ell_I = 0$ and $\ell_H = 2.9847012075$. The solved allocation has household-only municipal holdings at $T = 1, \dots, 22$, joint holdings at $k = 23$, and lower-tax-only holdings at $T = 24, \dots, 30$. The lower-tax investor holds 47.04% of maturity 23 and all risky taxable bonds. The analytical price recursions use an affine continuation on this allocation branch. Positive holdings and inactive slacks are checked locally around the reference state. The quantitative model retains bond-specific risk and solves state-dependent participation.

Proposition 2 (Rate-risk-only case). *Consider a feasible affine branch with equal masses, zero bond-specific variance, household-only municipal holdings below maturity k , both investors holding maturity k , and lower-tax-only holdings above k . The lower-tax investor holds risky taxable bonds. Put $d = \partial_z \hat{D}_H$. If outside exposures and supplies are fixed, then*

$$d = \frac{\xi_I - \xi_H}{(\gamma_H + \gamma_I)\sigma_r^2 A_{k-1}^M} > 0, \quad \partial_z \hat{D}_I = -d. \quad (96)$$

The three municipal responses in (16) follow. Write $b = \gamma_I \sigma_r^2 d$ and suppose $0 < \rho - b < 1$. If A_T^M is increasing through k , then A_T^M/A_T^C increases from ξ_H toward ξ_I .

Proof. Only the shared municipal quantity changes on the branch. Clearing gives $\partial_z \hat{D}_H + \partial_z \hat{D}_I = 0$. Differentiating the two active municipal equalities at k gives $\xi_I - \xi_H = (\gamma_H + \gamma_I)\sigma_r^2 A_{k-1}^M d$, proving (96). Sole-holder conditions give $s_T^M = \xi_H + \gamma_H \sigma_r^2 A_{T-1}^M d$ below k and $s_T^M = \xi_I - b A_{T-1}^M$ above it. At k , risk-tolerance weighting cancels the exposure responses and gives $s_k^M = \xi^*$. The taxable response is $s_T^C = 1 - b A_{T-1}^C$.

Let $\alpha = \rho - b$ and $\beta = \rho + \gamma_H \sigma_r^2 d$. The loading recursions become

$$A_T^C = \alpha A_{T-1}^C + 1, \quad A_T^M = \begin{cases} \beta A_{T-1}^M + \xi_H, & T \leq k, \\ \alpha A_{T-1}^M + \xi_I, & T > k. \end{cases}$$

The first municipal recursion holds at k as well because (96) makes its response equal ξ^* . For $T \leq k$, $A_T^M/\xi_H = \sum_{j=0}^{T-1} \beta^j$ and $A_T^C = \sum_{j=0}^{T-1} \alpha^j$. Since $\beta > \alpha > 0$, their ratio increases with T . To bound it above, put $F_T = A_T^M - \xi_I A_T^C$. For $T \leq k$,

$$F_T = \alpha F_{T-1} + (\xi_I - \xi_H) \left(\frac{A_{T-1}^M}{A_{k-1}^M} - 1 \right).$$

The term in parentheses is nonpositive, so $F_T < 0$ from $T = 1$. For $T > k$, $F_T = \alpha F_{T-1}$. Moreover, A_T^M/A_T^C is a positive weighted average of its preceding value and ξ_I , so it continues to increase. As $T \rightarrow \infty$, $A_T^C \rightarrow 1/(1 - \alpha) > 0$ and $F_T \rightarrow 0$, proving the limit ξ_I . $\square$

Numerical and tax checks. The solved branch has $d = 8.33642$, $\hat{D}_H = 2.69314$, $\hat{D}_I = 3.44204$ and $\rho - b = 0.77395$. Active municipal equality errors are below 10^{-15} , the smallest inactive municipal slack exceeds 3.3×10^{-5} , and household taxable slacks exceed 0.0027. A separate constrained portfolio optimization at rate states $z = \pm 10^{-4}$ and $\pm 10^{-5}$ reproduces the municipal and taxable response coefficients to within

10^{-12} . Direct finite sums of response gaps and the loading recursion agree to within 10^{-12} . At $T = 30$, the return gap is 200.62 basis points, comprising 92 of exemption benefit and 108.62 of rate loading. The duration-cost contribution is 91.60 basis points at $T = 22$ and zero from $T = 23$, when the lower-tax holder's conditions are used for both bonds.

For P1 and P2, I differentiate the exposure fixed point, loading recursions, portfolio conditions and yield ratio with respect to one after-tax share at a time, fixing the remaining primitives. I check these derivatives by re-solving at tax perturbations of 10^{-4} , 10^{-5} and 10^{-6} . The shared maturity remains unchanged and the portfolio conditions hold. The analytic derivatives and the smallest-step central differences agree to within 10^{-7} across all thirty maturities. Household pass-through falls from 1 at $T = 1$ to 0.604, 0.318 and 0.242 at $T = 10$, 23 and 30. Institutional pass-through rises from 0 to 0.380, 0.650 and 0.721 at those maturities. These monotone profiles are results for the solved illustration.

For the long-end price continuation, the risk-premium coefficients remain those of the solved finite-supply economy. Since the lower-tax investor prices both markets, its exposure response enters both loading recursions through $-b = \gamma_I \sigma_r^2 \partial \hat{D}_I / \partial z$. Thus $A_\infty^C = 1/(1 - \rho + b)$ and $A_\infty^M = \xi_I/(1 - \rho + b)$. The duration-risk response changes both loadings, but the common denominator cancels in their ratio. The individual return responses are $s_\infty^C = (1 - \rho)/(1 - \rho + b) = 0.885$ and $s_\infty^M = \xi_I s_\infty^C = 0.690$, so neither response omits the portfolio adjustment. With $\pi_I = \gamma_I \sigma_r^2 \hat{D}_I$, the limiting yields are

$$y_\infty^C = \bar{r} + \pi_I A_\infty^C - \frac{1}{2} \sigma_r^2 (A_\infty^C)^2, \quad y_\infty^M = \xi_I \bar{r} + \pi_I A_\infty^M - \frac{1}{2} \sigma_r^2 (A_\infty^M)^2.$$

Re-solving the finite economy under the same tax perturbations and differentiating these limits gives $\mathcal{P}_H^Y(\infty) = -0.002$ and $\mathcal{P}_I^Y(\infty) = 0.944$. The loading-implied limits are 0 and 1, respectively. For the empirical unit-response reference, the ratio A_T^M/G_T , where $G_T = \sum_{j=0}^{T-1} \rho^j$, increases over the thirty illustrated maturities. Its infinite-maturity limit is $(1 - \rho)A_\infty^M$, not ξ_I . This reference is distinct from the jointly priced taxable market C .

A.10 A small bond-specific risk and overlapping holdings

Adding bond-specific risk to the main-text illustration broadens joint participation without materially changing the return gap. I keep its other primitives fixed and add municipal log-return standard deviation $0.004[1 + 7\sqrt{(T-1)/29}]$ for $2 \leq T \leq 30$. Taxable bonds have no bond-specific shock, and the one-period claims remain riskless. Both investors now hold municipal maturities 17 through 30, rather than sharing only maturity 23.

Figure 14 uses the same three pricing channels and the same municipal and taxable rate-loading contributions as the main text. Bond-specific compensation is at most 4.4% of the return gap. At thirty periods, the gap is about 205 basis points, comprising 92 of exemption benefit, 109 of rate loading and 4 of bond-specific compensation. At the last household-only maturity, $T = 16$, the 181-basis-point gap comprises 87 of duration cost, 89 of rate loading and 5 of bond-specific compensation. Removing the shock changes the return gap by at most 4.26 basis points over all thirty maturities. Its main role here is to support overlapping holdings.

A unique loading solution. For $a_{\min} = \xi_H \sum_{q=0}^{15} \rho^q$, predecessor loadings on the joint block satisfy $A_{T-1}^M \geq a_{\min}$ and $P_J/S_J \leq 1/a_{\min}$. The exposure map $\mathcal{F}_J$ of Appendix A.6 therefore maps $[0, \infty)$ into $[0, d_{\max}]$, where

$$d_{\max} = \frac{\xi_I - \xi_H}{(\gamma_H + \gamma_I) \sigma_r^2 a_{\min}} < 10.591841.$$

Outward-rounded forty-digit evaluation on 20,000 interval cells gives $\sup |\mathcal{F}_J'| < 0.080131 < 1$, so its fixed point is unique and satisfies

$$8.31025872 < d^* < 8.31025874.$$

The resulting $A_{16}^M \in (3.25148856, 3.25148857)$ is below $\xi^*/(1 - \rho) = 3.45769231$, verifying condition (M) in (87). Condition (S) in (81) has left side $0.16084 < \rho$.

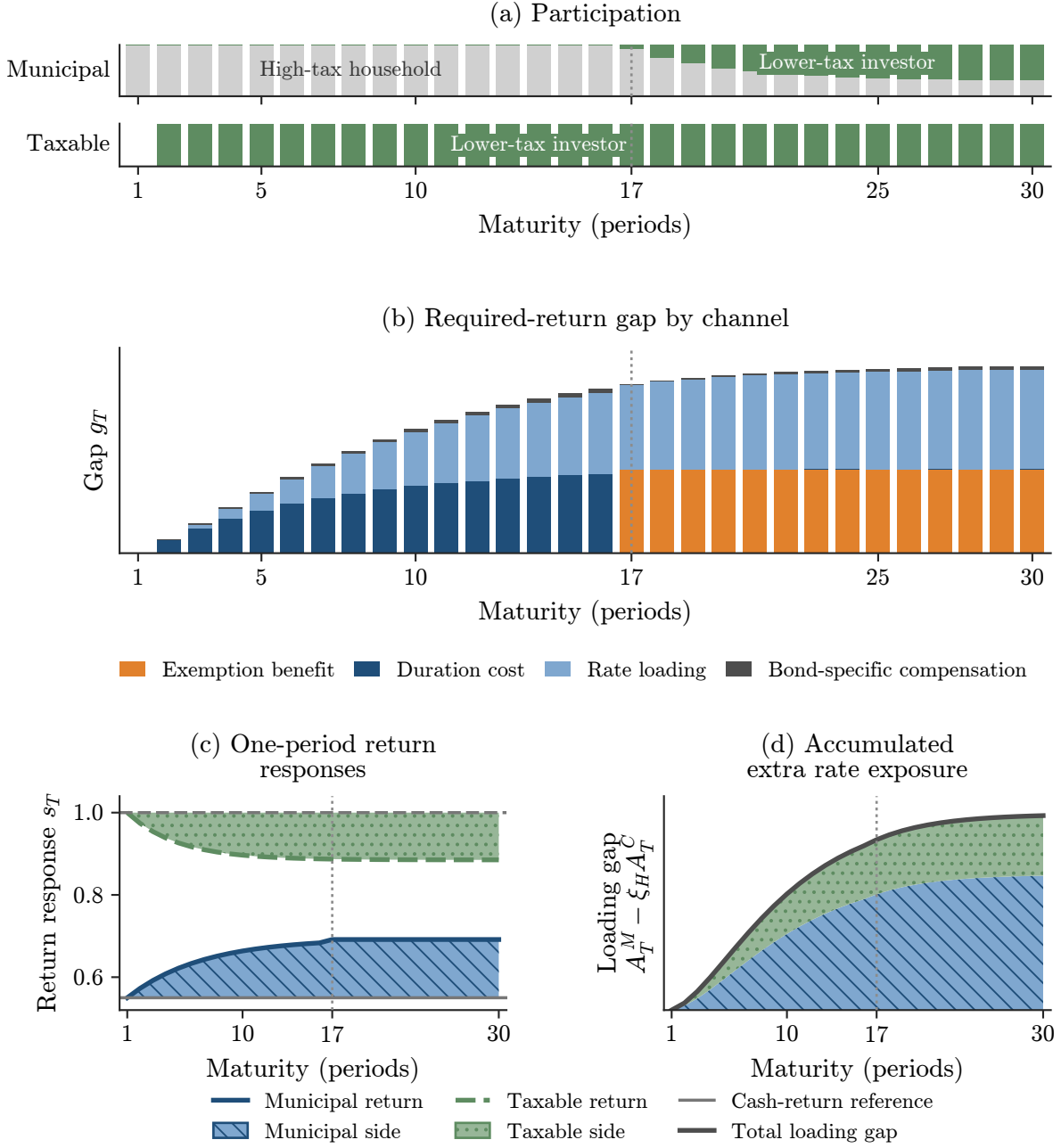


Figure 14: The analytical illustration with bond-specific risk. All other primitives match Appendix A.9. Panel (a) shows ownership. Panel (b) gives the three pricing channels of Figure 5, with small bond-specific compensation in gray. Panels (c) and (d) use the grouping of Figure 6: the municipal side combines the tax component and municipal portfolio response, while the taxable side reflects the taxable portfolio response. Panel (c) shades $s_T^M - \xi_H$ and $1 - s_T^C$. Panel (d) accumulates $s_T^M - \xi_H$ and $\xi_H(1 - s_T^C)$ into the total loading gap. The dotted line marks joint municipal holdings from $T^* = 17$.

Levels and the KKT conditions. The unchanged household outside exposure $\ell_H = 2.9847012075$ gives $\hat{\Lambda} = 4.319$, inside the interval $(4.285, 4.353]$ selecting $T^* = 17$ under Theorem 1. At $z = 0$, active KKT and return-gap identity errors are below 10^{-15} , both markets clear, inactive municipal slacks exceed 6.8×10^{-5} and household taxable slacks exceed 0.0026.

Pass-through across maturities. Re-solving the exposure fixed point, loadings and holdings for tax changes gives household pass-through declining from 1 to 0.24 and lower-tax pass-through rising from 0 to 0.73. The latter is 0.06 at $T = 2$, which the lower-tax investor does not hold. Outward-rounded evaluation verifies every adjacent comparison, with declines of at least 0.0088 for the household and increases of at least 0.0083 for the lower-tax investor. Their sum lies in $[0.97, 1.00]$, and positive holdings and inactive slacks preserve $T^* = 17$.

Long-end continuation. Continuing the price recursions with the solved economy's exposure responses and risk-premium coefficients, $Q_\infty^M = 0.1$ and $v_\infty^M = 0.001024$ gives $A_\infty^M = 3.458$, $A_\infty^C = 4.425$ and $s_\infty^C = 0.885$. The risk-premium coefficients are $(\pi_\infty^M, \pi_\infty^C) = (0.0118, 0.0108)$ and $y_\infty^C = 0.0816$. Loading-implied pass-throughs are $(-0.012, 1.012)$ and yield-implied pass-throughs $(-0.008, 0.956)$. The tax-difference condition of Proposition A.9 has left side $0.429 < 1$. These results include both markets' portfolio responses and convexity, rather than equating yield pass-through with risk-tolerance weights.

Equilibrium selection. Enumerated affine configurations are retained only when their linear level KKT-and-clearing system has a unique feasible solution. The uniqueness check applies to each configuration separately, and other feasible cutoffs remain possible, as noted in Appendix A.6. Full-state nonlinear solutions pass the independent grid checks in Appendix B.3, and their maturity patterns are numerical results.

A.11 Participation from other motives

Lower trading costs or better default diversification can induce long-maturity participation without a liability-hedging advantage (Section 4.7). The pricing comparison depends on neutrality: the motive must create no gap relative to the household benchmark when the household holds both bonds. Definition 1 states this condition for common risks and wealth costs. Separate default shocks require subtracting each market's holder-specific default compensation.

Proposition A.11 separates the participation incentive from its pricing effect. Proposition A.14 gives solved trading-cost and diversification examples, and Proposition A.15 shows why a municipal-specific preference has the opposite pricing implication. Throughout, $g_T = \mu_T^M - \xi_H \mu_T^C$ and $\xi_H < \xi_I$.

Proposition A.11 (Participation creates the gap). *In the two-investor affine benchmark without bond-specific risk, suppose every non-tax characteristic other than rate risk is neutral and zero at the one-period claim: a risk common to both markets, independent of rate risk, with municipal exposure ξ_H times the taxable exposure, or a cost that reduces after-tax wealth by the same amount in both markets.*

- (i) *If the household held every municipal and taxable bond, $g_T = 0$ at every maturity.*
- (ii) *Wherever the lower-tax investor holds the maturity- T municipal bond,*

$$g_T \geq \underbrace{(\xi_I - \xi_H)r}_{\text{exemption benefit}} + \underbrace{\gamma_I \sigma_r^2 \hat{D}_I (A_{T-1}^M - \xi_H A_{T-1}^C)}_{\text{rate loading}}.$$

Part (i) is proved in Appendix A.11.2, and part (ii) follows from Proposition A.12. The rate-loading term is positive on the branch of Proposition 2 when the lower-tax investor's net rate exposure is positive.

A.11.1 Non-tax required returns and neutral characteristics

Add to investor i 's objective in (37) a differentiable concave term $\Phi_i(x_i)$ in the same after-tax units. Dividing each KKT condition (33) by its after-tax share gives, under the normalized municipal-taxable benchmark,

$$\mu_T^M \leq \xi_i r + K_{i,T}^M, \quad \mu_T^C \leq r + K_{i,T}^C, \quad \text{with equality for a positive holding,} \quad (97)$$

where

$$K_{i,T}^a := \gamma_i \text{Cov}_z(u_{T,t+1}^a, \tilde{\Pi}_i(t+1)) + \frac{\kappa_i}{\xi_i^a} - \frac{1}{\xi_i^a} \frac{\partial \Phi_i}{\partial x_{i,T}^a}.$$

With additive shocks, $K_{i,T}^a$ splits into components $k_{ij,T}^a$ for rate risk, bond-specific risk, each additional shock or wealth characteristic, and the shorting constraint on cash. Additional shocks are independent of the rate and bond-specific shocks.

A common return characteristic is a unit-variance shock q with nonnegative maturity- T exposures e_T^M , e_T^C in both markets, taxed with returns. Its component is $k_{iq,T}^a = c_{iq} e_T^a$, where the cost per unit of exposure, $c_{iq} := \gamma_i \text{Cov}_z(q_{t+1}, \tilde{\Pi}_i(t+1))$, is the same in both markets. Affine rate risk has $q = -\varepsilon$, $e_T^a = \sigma_r A_{T-1}^a$, $c_{iq} = \gamma_i \sigma_r \hat{D}_i$ and component $\gamma_i \sigma_r^2 A_{T-1}^a \hat{D}_i$.

A wealth characteristic changes after-tax wealth by w_T^a per dollar at investor i 's intensity $\lambda_i \geq 0$, as with trading costs. This intensity is unrelated to the feedback λ in the proof of Lemma 1. For $\Phi_i = -\lambda_i \sum_{a,T} w_T^a x_{i,T}^a$, the component is $k_{iw,T}^a = \lambda_i w_T^a / \xi_i^a$. The shorting constraint on cash similarly gives κ_i / ξ_i^a .

Definition 1 (Neutral characteristic). *A common return characteristic is neutral if $e_T^M = \xi_H e_T^C$, and a wealth characteristic if $w_T^M = w_T^C$, at every maturity. For a household that bears it, a neutral characteristic satisfies $k_{H,T}^M = \xi_H k_{H,T}^C$, so it creates no gap when the household holds both bonds.*

Neutrality for each investor. A neutral common return characteristic gives $k_{iq,T}^M = c_{iq} e_T^M = c_{iq} \xi_H e_T^C = \xi_H k_{iq,T}^C$ for every investor i . A neutral wealth characteristic gives $k_{iw,T}^M = \lambda_i w_T = \xi_i (\lambda_i w_T / \xi_i) = \xi_i k_{iw,T}^C$, so

$$k_{iw,T}^M - \xi_H k_{iw,T}^C = (\xi_i - \xi_H) k_{iw,T}^C. \quad (98)$$

In the affine benchmark in which the household alone holds both bonds, rate risk is neutral. Clearing fixes its exposure, so differentiation gives $s_T^M = \xi_H$ and $s_T^C = 1$, and recursion (42), starting at $A_1^M = \xi_H$, $A_1^C = 1$, gives $A_{T-1}^M = \xi_H A_{T-1}^C$ at every maturity by induction, as in Proposition A.6.

Neutrality is stated for the household because the puzzle is defined against the household benchmark. A characteristic that weighs more on municipal bonds would add a premium even there.

A.11.2 The gap without and with participation

No holding-return gap when the household holds both bonds. Suppose the household alone holds both bonds, there is no bond-specific risk and every additional characteristic is neutral and zero at the one-period claim. Both household conditions in (97) bind, so $g_T = \sum_j (k_{Hj,T}^M - \xi_H k_{Hj,T}^C)$, and the terms κ_H and κ_H / ξ_H from its shorting constraint on cash cancel. Definition 1 removes the additional components. In the affine benchmark, the neutrality of rate risk makes the rate contribution $\gamma_H \sigma_r^2 \hat{D}_H (A_{T-1}^M - \xi_H A_{T-1}^C)$ zero, and with a constant cash rate there is no rate component. At riskless $T = 1$, the same equalities give $\mu_1^M = \xi_H \mu_1^C$, reducing to $\mu_1^M = \xi_H r$ and $\mu_1^C = r$ when the shorting constraint on cash is slack. Hence

$$g_T = 0 \quad \text{at every maturity.} \quad (99)$$

This is a statement about holding returns. The yield gap also includes the convexity difference in (79), so a zero holding-return gap need not imply a zero yield gap.

Participation incentive for any motive. At the household-only allocation of Appendix A.5, with components evaluated there, household-only municipal holdings give $\mu_T^M = \xi_H r + K_{H,T}^M$. The lower-tax investor's first-order gain from buying the maturity- T municipal bond, $\mu_T^M - \xi_I r - K_{I,T}^M$, is then nonnegative if and only if

$$\Delta_T := K_{H,T}^M - K_{I,T}^M \geq (\xi_I - \xi_H) r. \quad (100)$$

If $r > 0$, $\Delta_1 = 0$ and Δ_T is nondecreasing in T , the maturities satisfying (100) form an upper interval, possibly empty, that excludes the one-period claim. A nondecreasing Δ_T crosses the maturity-invariant tax difference at most once, and $\Delta_1 = 0 < (\xi_I - \xi_H) r$.

Rate risk gives $\Delta_T = \sigma_r^2 A_{T-1}^M \Lambda$ (Proposition A.3). At that allocation, the advantages from several motives add. This comparison holds prices and portfolios fixed. Equilibrium upper-interval participation follows from Proposition A.13(i) for the default and trading-cost motives and under the conditions of Theorem 1 for rate risk.

Proposition A.12 (The gap when the lower-tax investor participates). *In the economy of Appendix A.1 with one municipal market, suppose every additional characteristic is a neutral common return or wealth characteristic. If the lower-tax investor holds the maturity- T municipal bond at state z , then*

$$g_T \geq (\xi_I - \xi_H) r + g_T^{\text{risk}} + (\xi_I - \xi_H) \left(\sum_w k_{Iw,T}^C + \frac{\kappa_I}{\xi_I} \right),$$

$$g_T^{\text{risk}} := \gamma_I \left[\text{Cov}_z(u_{T,t+1}^M, \tilde{\Pi}_I(t+1)) - \xi_H \text{Cov}_z(u_{T,t+1}^C, \tilde{\Pi}_I(t+1)) \right],$$

where the covariances run over the rate and bond-specific shocks, with equality if it also holds the maturity- T taxable bond. In the affine benchmark without bond-specific risk, $g_T^{\text{risk}} = \gamma_I \sigma_r^2 \hat{D}_I(A_{T-1}^M - \xi_H A_{T-1}^C)$ and the shorting constraint on cash is slack. With nonnegative costs, this gives Proposition A.11(ii).

Proof. The active municipal condition gives $\mu_T^M = \xi_I r + K_{I,T}^M$. Taxable KKT gives $\mu_T^C \leq r + K_{I,T}^C$, with equality when taxable holdings are positive. Subtract ξ_H times the latter:

$$g_T \geq (\xi_I - \xi_H) r + \sum_j (k_{Ij,T}^M - \xi_H k_{Ij,T}^C).$$

Rate and bond-specific components give g_T^{risk} . Neutral common-return components cancel. Neutral wealth costs give $(\xi_I - \xi_H) k_{Iw,T}^C$ by (98). The terms κ_I and κ_I/ξ_I from the shorting constraint on cash give $(\xi_I - \xi_H) \kappa_I/\xi_I$. $\square$

The third term prices wealth costs and a binding shorting constraint on cash at the lower-tax investor's after-tax share. Nonnegative g_T^{risk} and nonnegative costs suffice for $g_T \geq (\xi_I - \xi_H) r$.

A.11.3 Three motives that produce the long-end gap

Proposition A.13 (Three motives that produce the long-end gap). *Let $r > 0$ and consider one of the three cases below. Cases (b) and (c) are set in the two-investor economy with a constant cash rate, no bond-specific risk and a slack shorting constraint on cash, so the stated characteristic is the only non-tax component. In each case the characteristic is zero at the one-period claim. In (b), investors have no outside exposure to the default shocks.*

- (a) *Liability hedging and risk-bearing capacity: the two-investor affine benchmark with $\Lambda > 0$ at the household-only allocation, without bond-specific risk for parts (i) and (iv).*
- (b) *Default diversification: municipal default risk loads on one unit-variance shock and taxable default risk on another, with correlation $\varrho \in [0, 1]$ and exposures $\xi_H e_T$ and e_T , where e_T is strictly increasing for $T \geq 1$. Write $X_i^M := \xi_H \sum_T e_T x_{i,T}^M$ and $X_i^C := \xi_i \sum_T e_T x_{i,T}^C$ for investor i 's municipal and taxable default exposures, and $k_{id,T}^M := \gamma_i \xi_H e_T (X_i^M + \varrho X_i^C)$ and $k_{id,T}^C := \gamma_i e_T (X_i^C + \varrho X_i^M)$ for its default components. Assume $\gamma_H X_H^{M,\text{hh}} > \varrho \gamma_I X_I^{C,\text{hh}}$.*
- (c) *Trading costs: the same round-trip cost $w_T \geq 0$ in both markets, strictly increasing for $T \geq 1$, and trading intensities $\lambda_I < \lambda_H$.*

Then:

- (i) *the participation advantage Δ_T in (100) is zero at $T = 1$ and nondecreasing in T . In (b) and (c), the municipal maturities the lower-tax investor holds in any equilibrium form an upper interval;*
- (ii) *$g_1 = 0$: the household anchors the one-period claim;*
- (iii) *at every maturity at which the lower-tax investor holds the municipal bond, $g_T \geq (\xi_I - \xi_H) r + g_T^{\text{risk}}$ in (a), with g_T^{risk} of Proposition A.12, $g_T \geq (\xi_I - \xi_H) r$ in (c), and in (b)*

$$g_T = (\xi_I - \xi_H) r + k_{Id,T}^M - \xi_H k_{Id,T}^C,$$

where $k_{d,T}^C$ is the default component of any investor who holds the taxable bond, so the gap net of these default components is $(\xi_I - \xi_H)r$;

(iv) with the household alone holding both bonds, $g_T = 0$ in (a) and (c), and $g_T = k_{Hd,T}^M - \xi_H k_{Hd,T}^C$ in (b).

Proof. (i) At household-only holdings, (a) gives $\Delta_T = \sigma_r^2 A_{T-1}^M \Lambda$ with $A_{T-1}^M = \xi_H \sum_{j=0}^{T-2} \rho^j$ (Proposition A.6); (b) gives $X_H^C = X_I^M = 0$ and $\Delta_T = \xi_H e_T (\gamma_H X_H^{M, \text{hh}} - \varrho \gamma_I X_I^{C, \text{hh}})$; (c) gives $\Delta_T = (\lambda_H - \lambda_I) w_T$. All vanish at $T = 1$ and increase weakly. In equilibrium (b), define investor i 's unit municipal-default cost $c_i := \gamma_i (X_i^M + \varrho X_i^C)$. The lower-tax investor's required municipal return is no greater than the household's exactly when $\xi_H e_T (c_H - c_I) \geq (\xi_I - \xi_H)r$. If the lower-tax investor holds maturity T , this comparison holds there and $c_H > c_I$. Strictly increasing e_T makes it strict beyond T , precluding household holdings and forcing lower-tax-investor holdings. Case (c) uses $(\lambda_H - \lambda_I) w_T$ identically.

(ii) Every non-tax component vanishes at the one-period claim, so the municipal conditions reduce to $\mu_1^M \leq \xi_i r$. Positive supply gives $\mu_1^M = \xi_H r$, while $\mu_1^C = r$, as in Proposition A.2.

(iii) Apply Proposition A.12 to (a), with no additional characteristic, and (c), with $g_T^{\text{risk}} = 0$ and cost $(\xi_I - \xi_H) \lambda_I w_T / \xi_I \geq 0$. In (b), active conditions give $\mu_T^M = \xi_I r + k_{Id,T}^M$ and $\mu_T^C = r + k_{d,T}^C$. With a constant rate, every taxable holder has the same default component. Subtraction proves the expression.

(iv) In (a) and (c) this is (99). In (b), the household's conditions give $\mu_T^M = \xi_H r + k_{Hd,T}^M$ and $\mu_T^C = r + k_{Hd,T}^C$. $\square$

In (a), Λ is positive when liability hedges or lower risk aversion reduce the lower-tax investor's duration costs. In (b), households bear all municipal default risk at the household-only allocation, while the lower-tax investor's taxable book is exposed to it only through ϱ , so the lower-tax investor can add it more cheaply. Independent shocks guarantee this advantage. Raw gaps also contain holder-specific default-premium differences driven by relative risk supplies. Appendix C.1.6 adjusts these with credit default swap (CDS) spreads. In (c), trading costs reduce after-tax wealth (Appendix C.1.5). For matched long bonds at 26 to 30 years, institutional round-trip costs are 0.53% of price in both markets (Table 21), supporting equal costs. With combined motives, the decomposition of Proposition A.12 still applies. Its tax-difference bound requires the remaining risk term g_T^{risk} to be nonnegative at the resulting equilibrium, and separate default shocks are treated as in (b).

A.11.4 Two solved examples without a duration motive

Consider two investors and three maturities with constant $r > 0$ and no bond-specific risk, eliminating rate-risk advantages. Both investors have risk aversion γ , masses m_H, m_I , after-tax shares $\xi_H < \xi_I$ and slack shorting constraints on cash. Supplies are Q_T^M for $T = 1, 2, 3$ and Q_T^C for $T = 2, 3$. Write $M1, M2, M3$ for the municipal bonds and $C2, C3$ for the longer taxable bonds, indexed by maturity. Each example adds only its stated non-tax characteristic.

Proposition A.14 (Participation for diversification or lower trading costs). *Consider the economy just described.*

(a) *Default diversification. Municipal and taxable default risk load on independent unit-variance shocks, with exposures $\xi_H e_T$ and e_T and $0 = e_1 < e_2 < e_3$. Let $X_H^M := \xi_H e_2 Q_2^M / m_H$ and $X_I^M := \xi_H e_3 Q_3^M / m_I$ be the investors' municipal default exposures when the household holds $M1$ and $M2$ and the lower-tax investor holds $M3$, and let*

$$X^C := \frac{e_2 Q_2^C + e_3 Q_3^C}{m_H / \xi_H + m_I / \xi_I}.$$

There is an equilibrium with this municipal allocation if and only if

$$\gamma \xi_H e_2 (X_H^M - X_I^M) \leq (\xi_I - \xi_H)r \leq \gamma \xi_H e_3 (X_H^M - X_I^M).$$

In it both investors hold taxable bonds, each with taxable default exposure X^C , and

$$g_1 = 0, \quad g_2 = \gamma \xi_H e_2 (X_H^M - X^C), \quad g_3 = (\xi_I - \xi_H)r + \gamma \xi_H e_3 (X_I^M - X^C).$$

Net of the default components of the investors who hold each bond, the gaps are 0, 0 and $(\xi_I - \xi_H)r$.

- (b) *Trading costs.* Each investor pays the after-tax cost $\lambda_i w_T$ per period on either bond of maturity T , with $0 = w_1 < w_2 < w_3$ and $0 \leq \lambda_I < \lambda_H$. If $(\lambda_H - \lambda_I)w_2 < (\xi_I - \xi_H)r \leq (\lambda_H - \lambda_I)w_3$, there is an equilibrium in which the household holds $M1$ and $M2$, the lower-tax investor holds $M3$, $C2$ and $C3$, and

$$g_1 = 0, \quad g_2 = \left(\lambda_H - \frac{\xi_H}{\xi_I} \lambda_I \right) w_2, \quad g_3 = (\xi_I - \xi_H) \left(r + \frac{\lambda_I w_3}{\xi_I} \right),$$

with $0 < g_2 < g_3$. The ratio μ_T^M / μ_T^C equals ξ_H at $T = 1$, lies strictly between ξ_H and ξ_I at $T = 2$ and equals ξ_I at $T = 3$.

Proof. (a) Independent default shocks separate market risks. Investor i 's KKT conditions are $\mu_T^M \leq \xi_i r + \gamma \xi_H e_T X_i^M$ and $\mu_T^C \leq r + \gamma e_T X_i^C$, for municipal and taxable exposures X_i^M, X_i^C . Lower taxable exposure implies lower required return and profitable purchases, so both investors hold taxable bonds with equal exposures. Clearing $\sum_i m_i X_i^C / \xi_i = e_2 Q_2^C + e_3 Q_3^C$ gives $X_H^C = X_I^C = X^C$ and $\mu_T^C = r + \gamma e_T X^C$. Each investor's split of its taxable exposure across the two maturities need not be unique. Supply-proportional holdings implement these exposures and clear both maturities. Active municipal conditions give $\mu_1^M = \xi_H r$, $\mu_2^M = \xi_H r + \gamma \xi_H e_2 X_H^M$ and $\mu_3^M = \xi_I r + \gamma \xi_H e_3 X_I^M$. Inactive KKT requires $\xi_H r < \xi_I r$ at $M1$, the left threshold inequality at $M2$ and the right at $M3$. Subtracting returns gives the gaps. Remove the holder-specific municipal component $\gamma \xi_H e_T X_i^M$ and the taxable component $\gamma e_T X^C$ to leave $\xi_H r - \xi_I r$ at $T = 2$ and $\xi_I r - \xi_H r$ at $T = 3$.

(b) Investor i requires $\xi_i r + \lambda_i w_T$ on municipal bonds and $r + \lambda_i w_T / \xi_i$ on taxable bonds. With slack shorting constraints on cash, each bond's return equals the minimum eligible requirement. Since $\lambda_I / \xi_I < \lambda_H / \xi_H$, the lower-tax investor holds all taxable bonds. It holds the maturity- T municipal bond when $\xi_I r + \lambda_I w_T \leq \xi_H r + \lambda_H w_T$, here only at $T = 3$. Subtract returns to obtain the gaps and

$$g_2 = (\lambda_H - \lambda_I)w_2 + (\xi_I - \xi_H) \frac{\lambda_I w_2}{\xi_I} < (\xi_I - \xi_H) \left(r + \frac{\lambda_I w_2}{\xi_I} \right) \leq g_3.$$

At $T = 2$ the ratio is below ξ_I because the household's required municipal return is below the lower-tax investor's. At $T = 3$ it is $(\xi_I r + \lambda_I w_3) / (r + \lambda_I w_3 / \xi_I) = \xi_I$. $\square$

Case (a) implements default diversification (Proposition A.13), requiring $X_H^M > X_I^M$. The intermediate raw gap $\gamma \xi_H e_2 (X_H^M - X^C)$ is a default-premium difference. With $\varrho \in (0, 1)$, the same allocation, where it is an equilibrium, has default-net gaps 0, 0, $(\xi_I - \xi_H)r$ (Proposition A.13(b)). Case (b)'s intermediate gap reflects higher household trading costs, independent of other holdings.

Examples with rate risk. With the affine rate process $z' = \rho z + \sigma_r \varepsilon$ added, the examples use $\xi_H = 0.6$, $\xi_I = 0.8$, $\rho = 0.8$, $\bar{r} = 4\%$ and small municipal bond-specific variance, with $\Lambda = 0$ at household-only holdings to remove the initial rate-risk advantage. For each motive, 120 economies with other parameters drawn randomly, including $\varrho \in [0, 0.9]$, are solved and screened for a participation incentive at household-only holdings. With trading costs, the lower-tax investor holds only the longest municipal bond, with $0 = g_1 < g_2 \leq g_3$ and $g_3 \geq (\xi_I - \xi_H)r$, in all 120. With default diversification, it does so in 72 economies and holds no municipal bond in 48. In the 72, the default-net gap is the tax difference plus Proposition A.12's rate and bond-specific terms. The bond-specific term is below 0.01 bp. The default-net long gap is at least the tax difference in 68 economies. In the other four, negative rate terms lower it below the tax difference by at most 0.09 bp. Rate terms across both motives span -0.09 to 1.7 bp.

A.11.5 A municipal-specific benefit

A municipal-only benefit $\zeta_T > 0$, from a mandate, issuer relationship or favorable regulation at matched maturity and credit, lowers the required municipal return without changing the taxable alternative. It therefore violates neutrality. For the lower-tax investor, it gives the participation advantage $\Delta_T = \zeta_T$ but reduces the return gap by ζ_T . A benefit that grows with maturity can thus induce long-end participation while lowering municipal returns. Unlike default compensation, it has no taxable counterpart to subtract.

Proposition A.15 (Same participation, opposite pricing). *Consider the two-investor economy without rate risk or bond-specific risk, with a constant cash rate $r > 0$, a slack shorting constraint on cash and no additional*

household characteristic. A fixed per-dollar benefit lowers the lower-tax investor's required municipal return by $\zeta_T \geq 0$, with $\zeta_1 = 0$, ζ_T nondecreasing and $\zeta_T < \xi_H r$. Compare a neutral benefit paid and taxed with returns, ζ_T on the municipal bond and ζ_T/ξ_H before tax on the taxable bond, which also lowers the investor's required taxable return by the household-equivalent ζ_T/ξ_H , with a municipal-specific benefit that leaves the required taxable return unchanged.

(i) In both cases the lower-tax investor holds the maturity- T municipal bond only if $\zeta_T \geq (\xi_I - \xi_H)r$, so participation begins at the same maturity.

(ii) With the neutral benefit, $\mu_T^C = r - \zeta_T/\xi_H$ and

$$g_T = \min\{\zeta_T, (\xi_I - \xi_H)r\} \geq 0,$$

so the gap equals the full difference in exemption benefits wherever the lower-tax investor participates. The ratio μ_T^M/μ_T^C rises with ζ_T .

(iii) With the municipal-specific benefit, $\mu_T^C = r$ and

$$g_T = \min\{0, (\xi_I - \xi_H)r - \zeta_T\} \leq 0,$$

and the ratio $\mu_T^M/\mu_T^C = \min\{\xi_H, \xi_I - \zeta_T/r\}$ is weakly below ξ_H where the lower-tax investor participates and strictly below it above the participation threshold.

Because the rate is constant, each yield averages the one-period returns $\mu_1^a, \dots, \mu_T^a$. The yield gap $y_T^M - \xi_H y_T^C$ therefore has the sign of the g_j , and the yield-implied tax is nonincreasing in maturity in case (ii) and nondecreasing in case (iii).

Proof. As in Proposition A.14(b), minimum required returns determine holdings. The household requires $\xi_H r$ and r . The lower-tax investor requires $\xi_I r - \zeta_T$ and either $r - \zeta_T/\xi_H$ or r . This proves (i). With neutrality, $g_T = \mu_T^M - \xi_H r + \zeta_T$ equals ζ_T when $\xi_H r$ is minimal, otherwise $(\xi_I - \xi_H)r$. The ratios $\xi_H r/(r - \zeta_T/\xi_H)$ below the threshold and $(\xi_I r - \zeta_T)/(r - \zeta_T/\xi_H)$ above increase in ζ_T since $\xi_I > \xi_H$. A municipal-only benefit gives $g_T = \mu_T^M - \xi_H r$. Since ζ_T is nondecreasing, μ_j^M/μ_j^C is nondecreasing in j in (ii) and nonincreasing in (iii). Adding a maturity whose ratio is at least (in (iii), at most) the running yield ratio cannot lower (raise) that ratio, which proves the yield statements. $\square$

A.12 State-specific exemptions and multiple municipal states

A household resident in state s' pays its home state's income tax on interest from a municipal bond issued by another state $s \neq s'$ but not on a home-state municipal bond. The quantitative institutional types have no modeled state-income-tax wedge on either, so state taxation is a household-specific channel (Appendix D.3). Statutory top state rates vary across states, with a supply-weighted mean of 5.8%. This cross-state variation is both a calibration input and P1's source of identification.

Multi-state aggregation. The multi-state model replaces ξ_i with $\xi_{i,s} = \xi_i^C/\xi_{i,s}^M$, which depends on the issuing state s and on the investor, whose type includes a household's state of residence. For state s 's market M_s , put $\Gamma_T^{M_s} := \sum_{i \in \mathcal{U}_T^{M_s}} m_i/\gamma_i$ and $\omega_{i,T}^s := (m_i/\gamma_i)/\Gamma_T^{M_s}$, where $\mathcal{U}_T^{M_s}$ is the set of holders of the maturity- T bond. The exact KKT aggregation (69) applies state by state, with $\xi_{ij} = \xi_{i,s}$. Applying (71) coordinate by coordinate gives, in the affine construction,

$$\begin{aligned} s_T^{M_s} &= \sum_{i \in \mathcal{U}_T^{M_s}} \omega_{i,T}^s \xi_{i,s} + \frac{\sigma_r^2 A_{T-1}^{M_s}}{\Gamma_T^{M_s}} \sum_{i \in \mathcal{U}_T^{M_s}} m_i \partial_z \hat{D}_i \\ &\quad + \frac{v_{M_s,T}}{\Gamma_T^{M_s}} \sum_{i \in \mathcal{U}_T^{M_s}} m_i \xi_{i,s}^M \partial_z x_{i,M_s}(T). \end{aligned} \tag{101}$$

Here each investor uses one payoff normalization across all markets, with after-tax shares ξ_i^a as in (32). Its net rate exposure in these units is $\hat{D}_i = \sum_{a,T} \xi_i^a x_{i,T}^a A_{T-1}^a - \ell_i$. When the municipal after-tax shares all equal

one, this reduces to (54). The final term in (101) vanishes under common active municipal after-tax shares, a singleton active set or zero bond-specific variance, as for the taxable counterpart before (74). Unweighted clearing alone does not remove it.

Proposition A.16 is the state-tax counterpart of Proposition A.10. Its implied state tax is the model counterpart of (24), and T_s^* is the first maturity at which investors other than the resident household hold state s 's bonds.

Proposition A.16 (State-tax anchor and tax-mix pass-through). *If the household H_s resident in state s uniquely minimizes $\xi_{i,s}$, the shorting constraint on cash is slack and the one-period municipal boundary is deterministic with $r > 0$, then $y_1^{M_s}/y_1^C = \xi_{H_s,s}$. Under the multiplicative mapping $\xi_{H_s,s} = \xi^{\text{fed}}(1 - \tau_s^{\text{st}})$, with federal after-tax share ξ^{fed} and state income-tax rate τ_s^{st} , define*

$$\tau_{s,T}^{\text{st,impl}} = 1 - \frac{y_T^{M_s}/y_T^C}{\xi^{\text{fed}}}, \quad \tau_{s,T}^{A,\text{st}} = 1 - \frac{A_T^{M_s}/A_T^C}{\xi^{\text{fed}}}. \quad (102)$$

Then $\tau_{s,1}^{\text{st,impl}} = \tau_s^{\text{st}}$, with unit pass-through. With exposure responses switched off in the affine recursions, if only the resident household's effective after-tax share changes and its active weight is one below T_s^* and constant at $\omega_{H_s}^s$ thereafter, the tax-mix derivative of the loading-implied state tax $\tau_{s,T}^{A,\text{st}}$ with respect to τ_s^{st} is $\Theta_S^s(T) + \omega_{H_s}^s \Theta_L^s(T)$, where Θ_S^s, Θ_L^s are the weights in (93) with T_s^* in place of T^* .

Proof. The boundary return is $\min_i \xi_{i,s} r$ (Proposition A.2). Equation (102) gives the anchor. Applying (93) with T_s^* in place of T^* gives the tax-mix derivative, because $\partial \xi_{H_s,s} / \partial \tau_s^{\text{st}} = -\xi^{\text{fed}}$ cancels the factor $1/\xi^{\text{fed}}$ in (102). $\square$

Changes to out-of-state after-tax shares or other clienteles add their derivatives. Other statutory mappings require a corresponding implied-tax transformation. The primitive boundary remains $y_1^{M_s}/y_1^C = \xi_{H_s,s}$. Reported state-tax responses use the full multi-state model, which supplies P1's full-yield pass-through profile, including level and exposure corrections.

A.13 Participation and pass-through with many clienteles

Extend the two-investor affine benchmark of Appendix A.4 to N investor types with slack shorting constraints on cash. Order after-tax shares $\xi_1 < \dots < \xi_N$, so type 1 is the household with the largest exemption benefit. Without bond-specific risk, subtract type 1's active municipal equality from type $j > 1$'s condition for a nonnegative first-order purchase incentive (Appendix A.5). Type j has that incentive at maturity T when

$$\underbrace{\sigma_r^2 A_{T-1}^M \Lambda_j}_{\text{duration benefit}} \geq \underbrace{(\xi_j - \xi_1)r}_{\text{difference in exemption benefits}}, \quad \Lambda_j := \gamma_1 \hat{D}_1 - \gamma_j \hat{D}_j. \quad (103)$$

At fixed net rate exposures the duration benefit grows with the bond's rate loading, while the difference in exemption benefits is flat in maturity. Positive bond-specific risk adds the incumbent's bond-specific compensation to the comparison, as in the KKT system (50) and the participation index (62).

Assumption 2 (Long-end participation, clientele by clientele). *Type 1 is active at every maturity, so $T_1^* = 1$. For each type $j > 1$, $\Lambda_j > 0$, and its active municipal set is an upper maturity interval $[T_j^*, T_{\max}]$ with $T_j^* \geq 2$. The thresholds need not be ordered by tax advantage.*

Many-clientele tax-mix pass-through. To isolate the tax mix under the nested supports of Assumption 2, switch off exposure-response terms and fix the support and risk-tolerance weights, as in Proposition A.10. Let $\omega_{k,j}$ be type k 's share of the holders' risk tolerance at municipal maturity j , zero when it does not hold that bond. Then $A_T^C = G_T$ and $A_T^M = \sum_{j \leq T} \rho^{T-j} \sum_k \omega_{k,j} \xi_k$. Differentiating $\tau_T^A = 1 - A_T^M/G_T$ in ξ_i gives its tax-mix pass-through $-\partial \tau_T^A / \partial \xi_i$,

$$\mathcal{D}_i(T) = \frac{\sum_{j=T_i^*}^T \rho^{T-j} \omega_{i,j}}{\sum_{j=1}^T \rho^{T-j}}, \quad (104)$$

with an empty numerator equal to zero. It is positive after participation begins and converges to ω_i if the active weight is eventually constant at ω_i , because the finite pre-participation block receives vanishing weight at the long end. Participation ordered by increasing after-tax shares makes the tax mix, the clientele-weighted after-tax share, nondecreasing in maturity, because adding an after-tax share no smaller than the incumbents' after-tax shares does not lower their risk-tolerance-weighted mean. Such ordered participation does not by itself order total slopes, loading ratios or yields, since other terms in (101) and the taxable leg remain unrestricted. Clientele-specific pass-through and endpoint comparisons need no tax ordering.

Under Assumption 2, the household anchors the riskless boundary, and (104) gives each clientele's tax-mix derivative. Full yield responses also include relative compensation and convexity (Appendix A.7). The portfolio-risk response can transmit a later clientele's tax to shorter risky bonds it never holds. P1 is the always-active household case and P2 the later-participating case. Assumption 2 restricts only this affine extension. The calibration does not impose it (Section 5).

A.14 Other markets

The mechanism requires common portfolio risk, heterogeneous relative taxes and a shorting friction. For a tax-favored bond and a comparable fully taxed bond, $\xi_i = (1 - \tau_i^{\text{tax}})/(1 - \tau_i^{\text{fav}})$, with investor i 's tax rates on the taxed and favored bonds. Symmetric taxation gives $\xi_i = 1$ and removes the relative tax incentive. The response decomposition (69) applies when after-tax shares differ. Positive taxable forwards and an increasing relative forward ratio imply a declining yield-implied tax (Proposition A.8). Condition (S) orders the affine loading ratio (Proposition A.7).

Applications include state-exempt treasury versus corporate interest; state-exempt Federal Home Loan Bank and Farm Credit issues versus Fannie Mae and Freddie Mac issues; treaty-dependent withholding on cross-border sovereign debt; and the inflation accrual on treasury inflation-protected securities (TIPS), taxed before payment in taxable accounts. Taxable accounts, individual retirement accounts (IRAs) and pensions also assign different after-tax payoffs to the same security, which retains one market price. In each application, the maturity profile depends on which clienteles hold the bonds. Municipal bonds combine a large federal wedge with tight shorting restrictions.

B Numerical solution and calibration

The quantitative economy of Section 5, with 643 investor types and 689 bond coordinates, is solved for municipal and taxable price functions of the rate state and for bond and cash positions per dollar of baseline portfolio assets. The budget W_i is measured in the same units, so total funding for type i is $m_i W_i$.

B.1 State-contingent moments and funded market clearing

The solver alternates between prices and portfolios. Given price functions, it computes resale risks and solves investors' portfolio conditions with market clearing at each rate state. The resulting required returns update prices. Iteration continues until the same price functions determine both resale values and current portfolio choices, satisfying Section 4.1 and Appendix A.1.

Rate-state grid. Let $\{z_g\}_{g=1}^G$ be the rate-state grid. It combines fifteen equally spaced nodes on $[-6\sigma_z, 6\sigma_z]$, which include $z = 0$, with the block's three observed annual rate states, for eighteen nodes in all; $\sigma_z = \sigma_r/\sqrt{1 - \rho^2}$ is the unconditional standard deviation of the rate state. An observed rate state is $z_t = r_t - \bar{r}$, where r_t is twice the continuously compounded two-year corporate zero yield in month t . The annual nodes use December of each block year, and yields are read at all 36 months of the block. The grid has no node at the rate floor $z = -\bar{r}$, the kink of the cash return, so the two-period taxable bond's continuation evaluates the cash price through the same piecewise-linear interpolant as other prices. No rate state is selected to match long corporate yields.

Continuation moments. Each coordinate j identifies a market and a maturity. Its future resale price determines both expected returns and portfolio risk. Write v_j for bond-specific return variance and $p_{j,-}(Z')$

for the log price of the one-period-shorter bond next period. Conditional on today's state z , the next state is $Z' = \rho z + \sigma_r U$ with $U \sim N(0, 1)$, as in (30). The required moments are

$$\mathcal{M}_j(z) = \log \mathbb{E}_z \exp\{p_{j,-}(Z')\} + \frac{1}{2} v_j, \quad (105)$$

$$\Sigma_{jk}(z) = \text{Cov}_z(p_{j,-}(Z'), p_{k,-}(Z')) + \mathbf{1}_{\{j=k\}} v_j. \quad (106)$$

The term $\mathcal{M}_j$ is the log expected resale value including the bond-specific mean correction. Subtracting today's log price gives $\mu_j = \mathcal{M}_j - p_j = \log \mathbb{E}_z[1 + R_j]$, the mean adjusted holding return of Appendix A.1. The matrix Σ gives covariances of the centered log-return innovations u_j . Piecewise-linear log prices, extrapolated along their end segments, permit analytic Gaussian integration of exponential moments and covariances (Appendix B.3). Thus Σ has a low-rank common component plus the diagonal bond-specific variances.

Investors and payoff units. Investor i has bond after-tax shares $K_i = \text{diag}(\xi_i^{a(j)})$ ($a(j)$ is coordinate j 's market), cash after-tax share ξ_i^C , funding W_i , cash $x_i^0 = W_i - \mathbf{1}' x_i$ and the multiplier κ_i on the shorting constraint on cash, $x_i^0 \geq 0$. After-tax shares are after-tax payoffs per dollar of pretax payoff under each block's tax law, in the normalized payoff units of Section 4.2, which set to one a household's absolute after-tax share on home-state municipal interest and an institution's on all municipal interest. Risk aversion γ_i and outside exposure ℓ_i are transformed as in Section 4.2. In these units ξ_i^C is the investor's relative after-tax share ξ_i , which includes proration for insurers, and a household's share on another state's municipal bonds is one minus its home-state rate.

The liability credit is $c_{ij}^L(z) = -\text{Cov}_z(p_{j,-}(Z'), \ell_i \sigma_r U)$ as in (49). The two-year financing-tax vector τ_i^{fin} is zero for nonbanks and taxable bonds. For banks it equals twice the annual financing-tax cost of Appendix B.7 on every municipal coordinate and at every rate state.

Portfolio conditions and clearing. Stationarity is

$$\begin{aligned} g_i &= \gamma_i K_i [\Sigma K_i x_i - c_i^L] - K_i \mu + \xi_i^C r \mathbf{1} + \tau_i^{\text{fin}} + \kappa_i \mathbf{1}, \\ g_{ij} &= \frac{\epsilon w_j}{x_{ij}}, \quad \kappa_i = \frac{\epsilon}{x_i^0}, \quad x_{ij} > 0, \quad x_i^0 > 0. \end{aligned} \quad (107)$$

Here g_{ij} , distinct from the return gap g_T , is investor i 's net cost of holding coordinate j . The second line replaces exact complementarity in the Kuhn–Tucker conditions (33) (in normalized units, plus the bank's financing-tax cost) by a log barrier on holdings and cash with weight $\epsilon > 0$ (10^{-9} in the baseline). The two-year log gross cash return is $r = \max\{\bar{r} + z, 0\}$. A municipal coordinate's barrier weight w_j equals its issuer state's share of national municipal supply at every maturity, so splitting supply across issuer states does not multiply the barrier. Taxable coordinates have unit weight. Clearing requires

$$\sum_i m_i x_i(z_g) = Q. \quad (108)$$

Appendix B.2 describes how these equations are solved.

B.2 The solution algorithm

Given candidate log price functions, continuation moments (105)–(106) define a static clearing problem with the shorting constraint on cash at each grid node. Its one-period returns imply the price update

$$\hat{p}_j(z_g) = \mathcal{M}_j(z_g) - \mu_j(z_g), \quad (109)$$

and the algorithm iterates on prices until $\hat{p} = p$ (Algorithm 1). Two boundary conditions close the recursion. A one-period municipal claim pays one next period, so its continuation log price is zero and its return is riskless. Its current price is solved in equilibrium. Cash, the one-period taxable claim, is not one of the 689 coordinates. It earns $r = \max\{\bar{r} + z, 0\}$, so its log price is $-r$, which supplies the continuation price of the two-period taxable bond.

Consistent expectations. Moments integrate every segment and extrapolated tail of the candidate prices. Covariances and the shorting constraint on cash link all maturities, so the entire price array is updated together.

Participation. Every investor may hold every bond. The barrier replaces exact zero holdings with small positive positions: $x_{ij} = \epsilon w_j / g_{ij}$. Positions with a large net cost relative to ϵ are of order 10^{-12} in the reported equilibria. Portfolios are re-solved at every node and iteration, so participation adjusts with prices. At the three annual nodes, high-tax households hold more than 99.99% of each municipal maturity through six years and less than 0.01% beyond eighteen years.

Step 1: initial prices and portfolios. The block starts from the one-market equilibrium at the same coefficients (Appendix B.6), which is solved by the same algorithm with one municipal market and household after-tax shares averaged across issuer states by supply shares. Its municipal prices are copied to every issuer-state market. Each household’s municipal holdings are placed in its home-state market, each institution’s are split across markets by supply shares, and taxable holdings are unchanged. One-period returns start at $\mathcal{M} - p$. Before the first Newton solve, one-period returns and the multipliers on the shorting constraint on cash are adjusted so that the mapped portfolios satisfy the market-level first-order conditions, and positions that the mapping leaves near zero are reset to their barrier values. Every position stays free, so the start does not restrict holdings.

Step 2: portfolios and clearing at each grid node, for given prices. At each grid node, the moments $\mathcal{M}(z_g)$ and $\Sigma(z_g)$ and the liability credits follow in closed form from the current price functions. During the solve, singular values of $\Sigma(z_g)$ below 10^{-12} times the largest are trimmed. Newton’s method then solves stationarity and clearing, (107)–(108), jointly for holdings x_i , cash x_i^0 and one-period returns μ . Each investor’s Hessian is diagonal plus low rank, so it is inverted with the Woodbury identity. Eliminating holdings leaves a 689×689 linear system in μ , the market Schur complement, which is solved by LU factorization and refined when needed to a relative residual of 10^{-12} . A fraction-to-boundary rule and backtracking keep bond and cash holdings positive. In the first price iteration, the barrier weight falls from 10^{-5} to 10^{-9} in factor-of-ten stages. Later iterations start from the node’s previous solution at $\epsilon = 10^{-9}$, and a failed solve restarts from proportional holdings with a larger barrier. Newton iteration stops when the stationarity and clearing residuals are below 10^{-11} and the largest product of a holding and its stationarity residual is below 2.5×10^{-12} . The grid nodes are solved independently and in parallel.

Step 3: price update. All prices are updated together from (109), across markets, maturities and grid nodes. A damped step moves prices a fraction $\min\{0.65, 0.025/\|\hat{p} - p\|_\infty\}$ of the way to $\hat{p}$, so no log price changes by more than 0.025. Anderson acceleration combines the last seven price iterates with least-squares weights and mixes at the damped step size. It restarts when the price error more than doubles, is used only when every weight is below 50 in absolute value, and is scaled to the same 0.025 cap.

Step 4: stopping and acceptance. Iteration stops when $\|\hat{p} - p\|_\infty < 10^{-10}$. All moments are then recomputed without the singular-value trimming, and the equilibrium is accepted only if the price, stationarity, clearing and budget residuals of Appendix B.3 meet their thresholds. Yields at the 36 monthly rate states of the block come from the solved log prices by linear interpolation in z , and the national municipal yield averages the issuer-state yields with supply weights. The comparisons with data in Figures 7 and 8 instead read each bond’s issuer-state yield at the final month of its issue quarter and average with the empirical bond, rating and quarter weights.

Convergence. The price fixed point is the only outer loop. In the six calibration blocks, the price error $\|\hat{p} - p\|_\infty$ starts near 10^{-2} and falls steadily, slowly for the first fifteen to twenty iterations and then by about a factor of ten every five iterations. It reaches 10^{-6} within forty iterations and 10^{-10} within seventy. The first iteration, which carries the barrier stages, is the most expensive. Later iterations need about two Newton steps per grid node.

Policy counterfactuals use the same algorithm, barrier and acceptance checks. The equal-treatment counterfactual (Section 7.3) uses price-update, inner, holding-weighted and linear-solve tolerances of 2×10^{-7} , 2×10^{-8} , 5×10^{-9} and 2×10^{-9} , respectively. Its 2015–2017 endpoint, which removes 99.9% of each tax distinction, starts from the solution that removes 99%, and the other five pre-2018 economies start from this accepted endpoint and independently pass the same checks. In a smaller version of the 2015–2017 economy with three municipal markets, moving from 99% to 99.9% removal changes thirty-year municipal and taxable yields by 1.9 and 0.2 basis points and each investor group’s share of bonds at 4–10, 12–20 and 22–30 years by at most 0.2 percentage points. The same move shifts up to 9.5 percentage points of a group’s assets between municipal and taxable bonds of the same maturity, which become close substitutes, so Panel B of Table 51 reports each group’s riskless share and the maturity of its risky holdings. Near a binding shorting constraint on cash, the cash Newton direction is computed from the low-rank inverse to avoid cancellation. Other policy solves use the baseline tolerances or tolerances between these and the baseline.

Algorithm 1 Full-state equilibrium of one block

Require: after-tax shares, funding, masses, supplies, exposures, rate dynamics, state grid $\{z_g\}$

- 1: Initialize p and holdings from the one-market equilibrium; $\mu \leftarrow \mathcal{M} - p$ ▷ Step 1
- 2: **repeat** ▷ outer loop over price functions
- 3: **for all** grid states z_g , in parallel **do** ▷ Step 2
- 4: Compute $\mathcal{M}(z_g)$, $\Sigma(z_g)$ and liability credits from the current p
- 5: Solve stationarity and clearing (107)–(108) for x , x^0 and μ by barrier Newton
- 6: **end for**
- 7: $\hat{p} \leftarrow \mathcal{M} - \mu$ at every grid state ▷ Step 3
- 8: Move p toward $\hat{p}$ by a damped or Anderson step, with log-price changes capped at 0.025
- 9: **until** $\|\hat{p} - p\|_\infty < 10^{-10}$ ▷ Step 4
- 10: Recompute the moments; accept only if the price, stationarity, clearing and budget checks pass
- 11: Read yields at the block’s monthly states and holdings at its annual nodes

The local rate loadings $A_T^a = -\partial_z p_T^a$ satisfy the nonlinear recursion (48) under the differentiation conditions of Appendix A.2.

B.3 Equilibrium residuals

Table 7 independently checks that prices agree with expected resale values, bond markets clear and portfolios satisfy optimality. With g_i from (107), its four diagnostics are

$$\begin{aligned} R_p &= \|p - \mathcal{M} + \mu\|_\infty, & R_c &= \|Q - \sum_i m_i x_i\|_\infty, \\ R_s &= \|g - \epsilon w/x\|_\infty, & R_{\text{comp}} &= \max\{\|x \odot g\|_\infty, \|\kappa \odot x^0\|_\infty\}. \end{aligned} \tag{110}$$

All six equilibria have zero budget and negative-holding violations, and their riskless claims have zero return variance. Acceptance requires $R_p < 3 \times 10^{-7}$, $R_s < 10^{-7}$, $R_c < 10^{-8}$ and a budget violation $\max_i (\mathbf{1}' x_i - W_i)$ below 10^{-10} , which all six blocks meet. The complementarity measure R_{comp} is reported only, since it equals ϵ at an exact barrier solution.

The fixed-coefficient 2018–2020 extension (Appendix B.7) also satisfies these checks: $R_p = 4.519 \times 10^{-11}$, $R_s = 7.777 \times 10^{-12}$, $R_c = 2.220 \times 10^{-16}$ and $R_{\text{comp}} = 1.001 \times 10^{-9}$, with zero budget and negative-holding violations and zero variance on riskless claims.

All forty-two home-state repeals, seven states in each of the six pre-2018 economies, use the baseline barrier and tolerances. Independent reconstruction gives maximum residuals $R_p = 9.788 \times 10^{-11}$, $R_s = 9.196 \times 10^{-11}$ and $R_c = 3.331 \times 10^{-16}$, with zero budget and negative-holding violations and zero variance on riskless claims. The riskless pricing frontier defined below is satisfied within 1.829×10^{-5} basis points at every observed month. Each experiment also passes separate tax-rule and fiscal-identity checks.

Table 7: Full-state equilibrium diagnostics

Block	R_p	R_s	R_c	R_{comp}
2000–2002	7.511×10^{-11}	5.884×10^{-12}	5.551×10^{-17}	1.002×10^{-9}
2003–2005	4.456×10^{-11}	8.340×10^{-12}	1.110×10^{-16}	1.002×10^{-9}
2006–2008	8.156×10^{-11}	8.547×10^{-12}	1.110×10^{-16}	1.002×10^{-9}
2009–2011	5.498×10^{-11}	8.687×10^{-12}	2.220×10^{-16}	1.002×10^{-9}
2012–2014	9.289×10^{-11}	9.864×10^{-12}	2.220×10^{-16}	1.002×10^{-9}
2015–2017	8.495×10^{-11}	8.352×10^{-12}	2.220×10^{-16}	1.002×10^{-9}

Notes. R_p , R_s , R_c and R_{comp} are the pricing, stationarity, clearing and complementarity residuals of (110). Maxima over each block's eighteen state nodes and the applicable coordinates, with $\epsilon = 10^{-9}$. Complementarity is the largest unpenalized product of a bond or cash holding and its multiplier, which equals ϵ at an exact barrier solution. The diagnostics assess only the discretized funded equations.

Riskless boundary and the household benchmark. The 45 one-period municipal claims are priced in funded clearing like other bonds. The lowest return at which a household type in $\mathcal{H}$, the set of all 45×14 household types, holds the claim of issuer state s , the household counterpart of the boundary (38), is

$$\min_{i \in \mathcal{H}} \frac{\xi_i^C r(z) + \kappa_i(z)}{\xi_i^{M_s}},$$

which includes the multiplier on the shorting constraint on cash and so can exceed $r(z)$ times the smallest after-tax-share ratio. In every equilibrium, each issuer state's one-period municipal return is within 0.01 basis points of this frontier at every grid node. The household benchmark uses the after-tax share of the high-tax household resident in the issuing state, which can differ from that of the household attaining the minimum. In the six calibrated economies, the national one-period municipal yield is within 0.01 basis points of the household benchmark at every observed month.

Analytic integration of the price interpolant. For a standard normal U and an affine segment $p = a + bU$ on $l < U < u$,

$$\mathbb{E}[e^{a+bU} \mathbf{1}_{\{l < U < u\}}] = e^{a+b^2/2} \{\Phi(u-b) - \Phi(l-b)\}.$$

Summing over segments and the extrapolated tails eliminates quadrature truncation, leaving interpolation and finite-precision error. The residuals above hold at the grid nodes and do not bound errors uniformly between them. Grid refinement assesses accuracy between nodes (Appendix B.4).

Weighted barriers and the shorting constraint on cash. Exact stationarity gives $x_{ij}g_{ij} = \epsilon w_j$ and $x_i^0 \kappa_i = \epsilon$. For concave preferences at fixed prices, the utility gap of the constrained problem is bounded by $\epsilon(\sum_j w_j + 1)$, provided the stationarity error is controlled.

B.4 Sensitivity to the grid, the barrier and a node at the rate floor

Three comparisons fix the seventeen coefficients and measured inputs in 2015–2017. The first increases the regular nodes from fifteen to twenty-nine, retaining the original nodes and endpoints. The second lowers ϵ from 10^{-9} to 10^{-10} with unchanged inner tolerances. The third adds a grid node at the rate floor (row “Exact floor knot”). I check each re-solve independently, and Table 8 reports the results.

The yield gap changes little even where holdings move more noticeably. Grid refinement reduces lower-tax households' municipal holdings by 4.02%. Adding the rate-floor node reduces those holdings by 1.69% and banks' taxable holdings by 2.80%. Changes in municipal and taxable yields partly offset each other.

Lowering the barrier leaves risky holdings and gaps stable, while riskless and near-zero holdings remain sensitive to it. Individual positions in riskless claims change by at most 77.06% of budgets across grid nodes and 24.84% at annual holdings nodes, compared with maxima of 0.1179% and 0.0709% for risky positions.

Table 8: Numerical sensitivity at fixed coefficients, 2015–2017

Comparison	Actual nodes	ϵ	Δ 30-year gap (bp)	Δ lower-tax household municipal book (%)	Δ bank taxable book (%)
Nested grid	32	1.000×10^{-9}	+0.199365	-4.020	-2.585
Smaller barrier	18	1.000×10^{-10}	+0.000138	-0.004	+0.001
Exact floor knot	19	1.000×10^{-9}	-0.228439	-1.689	-2.797

Notes. Each row compares one change with the original eighteen-node, 10^{-9} -barrier economy. Rows are not a sequence. The yield gap is the national supply-weighted municipal yield minus the original resident-household after-tax taxable benchmark, in continuously compounded yields over 36 actual months. Its baseline is 104.270250 bp. Holdings use the three annual measurement nodes, and percentage changes use each group's original book as denominator. The rows are neither confidence intervals nor the seven-block empirical-weight-matched gap. The floor is economically active in every row.

The largest riskless shift is at a positive cash rate. Complementarity falls from 1.0016×10^{-9} to 1.0134×10^{-10} , and the tiny residual-group municipal and high-tax household taxable books fall about 90%.

B.5 Risk specification

The bond-specific variance profile is $v_n = v_{30}(n/30)^{p_v}$ for maturities of $n = 4, 6, \dots, 30$ years, with $v_2 = 0$ and the same variances for municipal and taxable bonds at each maturity. The coefficients v_{30} and p_v are fitted jointly with preferences and effective exposures (Table 9) and held fixed in the comparisons above.

Effective outside exposures summarize rate risk without matching every asset position to liability cash flows. The averaged ownership and maturity conditions discipline them, and Sections 5.3–5.4 assess the fit to the yield curve. Appendix B.6 reports sensitivity to the coefficients.

B.6 Calibration conditions, parameter search and full-state evaluation

I estimate the seventeen common coefficients of Table 9 with the criterion (111) in six dated economies with one national municipal market. Table 10 lists the targeted conditions and untargeted tests. Exposure parameters use $h_i = -\ell_i$: positive values add rate risk, while negative values hedge it (Appendix B.7).

Measured inputs. The 1952–2024 annual rate estimates are a persistence of 0.853903 and an innovation standard deviation of 0.016266, and the mean 0.030206 is the 2000–2017 average of the two-year corporate-like yield. Squaring annual persistence gives two-year persistence. The two-year state is twice the annualized rate deviation, so its innovation standard deviation is $2(0.016266)\sqrt{1 + 0.853903^2}$. Thus $(\rho, \sigma_r, \bar{r}) = (0.729151, 0.042778, 0.060412)$ with floor $r = \max\{\bar{r} + z, 0\}$. Monthly two-year yields set the observed states (Appendix B.1). The pooled treasury and corporate asset uses corporate yields and household taxes, despite the different statutory treatment of treasuries. After-tax shares and bank financing-tax costs are distinct fixed inputs (Table 14), and after-tax shares alone do not identify tax receipts.

Condition construction and normalization. Table 15 lists the twenty-two conditions. Each compares a model outcome with its observed counterpart, then scales the difference so dollar stocks, maturities and yields can enter one criterion. Write g_{kj} for condition j 's signed model-minus-observed difference in block k and $o_{kj} > 0$ for its observed normalizer. For the nine ownership conditions, the difference and normalizer are the stock error in national dollars and the observed stock.

For each insurer sector and asset, the maturity condition compares model and observed mean maturities, weighting the difference by the model portfolio's value B . It divides $B(\bar{T} - \bar{T}^{\text{obs}})$ by $B^{\text{obs}}\bar{T}^{\text{obs}}$, where $\bar{T}$ is holdings-weighted maturity. The three taxable-yield conditions average the errors in annual log yields at ten, twenty and thirty years over 36 monthly states. Because observed two-year yields set the states, these are term-spread errors, normalized by the observed block-mean spreads over the two-year yield.

The remaining six conditions use maturities above two through thirty years, excluding the riskless two-year municipal stock included in the ownership and maturity conditions. For insurer i , let M_i and C_i

Table 9: Common estimated coefficients, held fixed in the full-state evaluation

Symbol	Economic role	Value
<i>Risk. Main information: taxable term spreads and the allocation of long bonds</i>		
γ	Common risk-aversion curvature, per unit of balance-sheet mass	9.6969
v_{30}	Bond-specific variance of a thirty-year bond's two-year return	0.000722
p_v	Maturity exponent of the bond-specific variance	1.2597
<i>Households. Main information: household municipal and taxable stocks</i>		
a_H	Outside-duration intercept	1.3045
b_A	Loading on the outside-asset ratio	2.704
b_D	Loading on the debt ratio	0.8379
<i>P&C insurers. Main information: P&C stocks, maturities, composition and long fractions</i>		
a_P	Low-type outside duration	-1.961
d_P	High-minus-low outside duration	3.0311
w_P	Low type's share of portfolio assets	0.0802
<i>Life insurers. Main information: life-insurer stocks, maturities, composition and long fractions</i>		
a_L	Low-type outside duration	-5.3927
d_L	High-minus-low outside duration	6.3007
w_L	Low type's share of portfolio assets	0.2616
t_L	Middle type's location within the outside-duration span	0.8625
u_L	Middle type's share of portfolio assets excluding the low type	0.0162
<i>Banks and residual group. Main information: bank stocks and taxable term spreads</i>		
h_B	Bank outside duration	1.2045
μ_R	Residual-group mean outside duration	1.0683
s_R	Residual-group outside-duration dispersion	1.7168

Notes. All seventeen coefficients are fitted jointly to the twenty-two averaged conditions of Table 10. Every coefficient affects every condition through equilibrium prices and clearing, and the headings name the conditions most informative about each block. Outside durations $h_i = -\ell_i$ are measured like net duration, in the model's two-year units. Portfolio-asset shares divide each insurer sector's measured debt securities and cash across its types. Appendix Table 11 reports the search intervals, starting values and step scales.

Table 10: Targeted conditions and untargeted tests

<i>A. Targeted: twenty-two conditions, each averaged over the six blocks of 2000–2017</i>	
Ownership stocks (9)	Municipal holdings of high-tax households, lower-tax households, P&C insurers, life insurers and banks. Taxable holdings of households, P&C insurers, life insurers and banks
Insurer maturity (4)	Holdings-weighted mean maturity of the P&C and life-insurer municipal and taxable books
Taxable term spreads (3)	Ten-, twenty- and thirty-year taxable yields over the two-year yield
Insurer composition (2)	Book-weighted mean of squared municipal portfolio shares, P&C and life insurers
Insurer long fractions (4)	Share of each insurer sector's municipal and taxable book maturing beyond twenty years
<i>B. Untargeted tests of the mechanism</i>	
Municipal yield curve	Evaluated on the observed corporate curve, the model accounts for 83% of the 117.5-basis-point thirty-year puzzle after callability and the issuance premium (Figure 8)
State-tax pass-through (P1)	With the model's own federal after-tax share, its cross-state slope falls from 1 at two years to 0.46 at thirty and lies inside the controlled pre-2018 OLS 95% band in the first four maturity bins and 0.003 above its upper bound in the longest (Figure 9)

Notes. Each averaged condition is scaled by one tenth of its observed six-block mean magnitude. Appendix Table 15 reports every scaled deviation at the estimated coefficients.

Table 11: Parameter-search intervals, initialization and coordinate scales

Coefficient	Search interval	Initial value θ_k^0	Step scale c_k
γ	[1, 10]	9	1
a_H	[-3, 3]	1	0.3
b_A	[0, 12]	2	0.4
b_D	[0, 20]	0.5	0.5
a_P	[-10, 2]	-1.5	0.5
d_P	[0.02, 12]	2.4	0.6
w_P	[0.01, 0.99]	0.22	0.08
a_L	[-12, 2]	-4	0.6
d_L	[0.02, 15]	3.2	0.8
w_L	[0.01, 0.99]	0.1	0.08
h_B	[-3, 3]	0.6	0.3
μ_R	[-4, 4]	1	0.4
s_R	[0.01, 5]	2.2	0.4
v_{30}	[0.000020, 0.006000]	0.0015	0.0004
p_v	[0, 3]	1.15	0.3
t_L	[0.01, 0.99]	0.75	0.08
u_L	[0.01, 0.99]	0.2	0.08

be municipal and taxable holdings in this window and $B_i = M_i + C_i$. The municipal-share second moment $\sum_i (M_i^2/B_i)/\sum_i B_i$ squares each insurer's municipal portfolio share and averages it with portfolio-size weights. At a given aggregate municipal share, a larger second moment means holdings are concentrated in insurers with higher municipal shares. For each sector-asset book B with holdings L beyond twenty years and observed long fraction f^{obs} , the long-maturity condition is

$$g^{\text{long}} = \frac{L - f^{\text{obs}}B}{B_{\text{window}}^{\text{obs}}}.$$

The observed portfolio value within the maturity window equals the national sector stock times its measured coverage in that window. These six conditions are scaled by the observed second moments and long fractions.

Holdings conditions are evaluated at each block's three observed end-year states and then averaged. Computing a ratio or product from averaged holdings would give a different target. Let $g_{kj}(\theta)$ be signed condition j in block k after its equilibrium solve at the common coefficient vector θ , and $o_{kj} > 0$ its observed scaling denominator. The criterion is

$$f_j(\theta) = \frac{6^{-1} \sum_{k=1}^6 g_{kj}(\theta)}{0.1 (6^{-1} \sum_{k=1}^6 o_{kj})}, \quad \mathcal{L}(\theta) = \frac{1}{2} \sum_{j=1}^{22} f_j(\theta)^2. \quad (111)$$

The book-weighted residuals above enter $\mathcal{L}$ with fixed inverse-squared-normalizer weights, which supply no covariance estimate, parameter standard errors or overidentification test.

Search coordinates. Table 11 gives bounds, initial vector θ^0 and step scales c_k , where a scaled step u_k changes coefficient k by $c_k u_k$. Derivatives and trust radii use these units. All economies start at θ^0 .

Joint search. For the scaled conditions f and their Jacobian J in u , central differences refresh J at accepted centers, using feasible one-sided differences at bounds. A bounded trust-region search compares damped Gauss-Newton, backtracked and projected-gradient candidates against the quadratic model

$$\mathcal{L} + (J'f)'u + \frac{1}{2}u'(J'J + S)u.$$

Here a structured secant update estimates the residual curvature S .

The initial search uses a finite-difference step of 0.002 and ends after 107 iterations at $\mathcal{L} = 36.357521$. A restart from that vector uses a step of 0.001 after the larger step fails the initial half-step consistency

check, and it retains the objective, bounds and equilibrium tolerances. The restart ends after 41 iterations at $\mathcal{L} = 36.286978$, the lowest retained criterion across both fits, and this vector is the quantitative baseline. Subsequent diagnostic probes do not replace it.

Search checks and interpretation. Outer-search convergence is separate from equilibrium accuracy. At the selected vector, steps of 0.002, 0.001 and 0.000125 give projected-gradient infinity norms of 14.57, 11.01 and 0.376, above the tolerance of 0.01. Changes in participation make the derivatives sensitive to the step size. The feasible probes tested improve $\mathcal{L}$ by at most 2.79×10^{-4} , which bounds only those directions. The selected calibration has no local-stationarity certificate.

Full-state evaluation. The full-state evaluation restores 45 issuer markets, separate risks and measured home-state taxes, with the same seventeen coefficients, quantities and condition definitions. These equilibria supply municipal yields, predictions and policies. Across the six blocks, the full-state thirty-year gap minus the one-market gap is $(-0.85, -0.72, 2.48, -0.22, -1.15, -0.93)$ bp, and taxable yields differ by at most 0.52 bp.

Coefficient sensitivity. Table 12 changes one coefficient group per one-market re-solve. Every change raises the criterion. Bond-specific variance moves the gap and taxable yield by less than 0.1 basis points. Lower risk aversion produces the largest change. It lowers the thirty-year taxable yield by 70 basis points and puts term spreads about 20% below targets.

Table 12: Sensitivity of the one-market calibration to the estimated coefficients

Coefficient change	30-year gap (bp)	Change (bp)	Criterion $\mathcal{L}$
Estimated coefficients	102.6		36.29
Risk aversion $\gamma \times 0.8$	89.0	-13.6	55.02
Risk aversion $\gamma = 10$ (upper bound)	104.7	+2.1	36.74
Bond-specific variance $v_{30} \times 0.8$	102.7	0.0	36.61
Bond-specific variance $v_{30} \times 1.2$	102.7	0.0	36.59
Maturity power $p_v - 0.2$	102.6	0.0	36.55
Maturity power $p_v + 0.2$	102.7	0.0	36.47
Household exposure intercept $a_H - 0.25$	99.0	-3.7	78.82
Household exposure intercept $a_H + 0.25$	106.7	+4.1	160.84
Household asset loading $b_A \times 0.8$	100.9	-1.7	62.45
Household asset loading $b_A \times 1.2$	105.2	+2.6	400.24
Household debt loading $b_D \times 0.8$	106.2	+3.5	6,720
Household debt loading $b_D \times 1.2$	103.9	+1.3	82.17
P&C exposures $\times 0.8$	102.2	-0.4	56.35
P&C exposures $\times 1.2$	103.3	+0.6	62.51
Life-insurer exposures $\times 0.8$	101.9	-0.7	59.68
Life-insurer exposures $\times 1.2$	103.4	+0.8	46.24
Bank exposure $\times 0.8$	100.5	-2.1	119.11
Bank exposure $\times 1.2$	104.5	+1.8	241.78
Residual exposure mean $\times 0.8$	102.6	-0.1	538.79
Residual exposure mean $\times 1.2$	101.7	-0.9	65.73
Residual exposure s.d. $\times 0.8$	99.3	-3.3	92.74
Residual exposure s.d. $\times 1.2$	105.8	+3.1	1,687

Notes. The one-market economy is averaged over the six blocks of 2000–2017, with all other coefficients at their estimates. Exposure changes scale every type’s exposure in the group. The gap is the municipal yield minus the household after-tax taxable yield at thirty years. The criterion $\mathcal{L}$ is the estimation objective, half the sum of squared scaled deviations of the twenty-two conditions.

Elastic bond supply. I allow borrowing to respond to its cost in both markets, holding the estimated coefficients fixed. For market $a \in \{M, C\}$ and maturity T , supply follows

$$Q_T^a(z) = \bar{Q}_T^a \left(\frac{y_T^a(z)}{\bar{y}_T^a} \right)^{-\chi_{a,T}},$$

where y_T^a is the annual continuously compounded yield. Each schedule passes through the original quantity and fixed-supply yield at the mean rate state. These anchors remain fixed when both yield curves are resolved. Risky maturities from four to thirty years have elastic supply. The two-year riskless claims retain their baseline treatment.

St. Clair (2024) estimates that a one-percent increase in borrowing costs reduces municipal issuance by approximately 0.3–0.6 percent. I use 0.45 as the midpoint of this rounded range and impose the same elasticity in both markets. The estimate concerns annual borrowing near the bank-qualified threshold. Holding financing terms fixed maps proportional borrowing-cost changes locally to proportional yield changes. Applying that response to outstanding model supply defines the comparative-static exercise.

Table 13 reports the one-market economy over the six calibration blocks. At equal elasticities of 0.3, 0.45, and 0.6, the average thirty-year gap lies between 101.0 and 102.0 basis points, compared with 102.6 under fixed supply. At 0.45, municipal yields fall by 14.1 basis points and the household after-tax taxable benchmark by 13.4, so the gap changes by only 0.7 basis points.

I also consider a larger taxable response. The November 2022 estimates of Siani (2022) imply approximately 1.5% less corporate issuance per ten-basis-point increase in credit spreads, holding the benchmark rate fixed. I match that local slope by setting $\chi_{C,T} = 15\bar{y}_T^C$, while keeping $\chi_{M,T} = 0.45$. The thirty-year taxable elasticity is 0.86–0.97 across blocks. Applying this corporate issuance response to the pooled taxable supply is an additional sensitivity assumption. The average gap is 99.2 basis points, retaining 96.7% of its fixed-supply value. Both yields adjust: municipal yields fall by 19.5 basis points and the after-tax taxable benchmark by 16.1. A larger taxable supply elasticity therefore need not widen the gap.

Table 13: Yield gaps with elastic supply in both markets

Municipal elasticity	Taxable elasticity	10 years	20 years	30 years
0	0	35.7	84.9	102.6
0.30	0.30	36.4	84.1	101.9
0.45	0.45	36.7	84.0	102.0
0.60	0.60	36.8	83.4	101.0
0.45	$15\bar{y}_T^C$	37.2	83.2	99.2

Notes. Gaps are in basis points, averaged over the original rate-state observations within each block and equally across the six blocks of 2000–2017. All other coefficients and inputs remain fixed. Elasticities are reported as positive magnitudes: higher borrowing costs reduce supply. The last row converts a corporate spread semi-elasticity to a yield elasticity at each fixed anchor, with yields in annual decimal units.

B.7 Measured inputs and investor construction

Markets. One model period is two years, and maturities in this subsection are in years. Each of the 43 states with a personal income tax and the District of Columbia forms its own municipal market. The seven states without one (Alaska, Florida, Nevada, South Dakota, Texas, Washington and Wyoming) share a single market because every investor treats their bonds alike: their residents pay no state tax on any municipal bond, and all other investors treat them as out-of-state bonds. Each of the 45 municipal markets has fifteen maturities from two to thirty years, and the taxable market fourteen from four to thirty years, for 689 market–maturity pairs. Cash, the one-period taxable claim, pays the short rate, and the two-year municipal bond is also riskless. All bond yields, including the two-year municipal yield, are determined in equilibrium. The taxable market pools treasury and corporate supply. Its yields are compared with the corporate curve, and its interest is taxed like corporate interest.

Table 14: Measured inputs, estimated coefficients and maintained measurement choices

Input	Source, construction and role
National portfolio assets and ownership	Financial Accounts vintage of September 8, 2023, with fund holdings assigned to their owners, nonprofit holdings separated and sector stocks adjusted for tax-exempt holdings. SCF direct and fund holdings split household municipal ownership. These stocks supply conditions 1–9, and tax-exempt interest flows are not used.
Household heterogeneity	Each SCF wave supplies seven balance-sheet bins within each household tax group, with equal amounts of portfolio assets in each bin. Outside financial assets exclude retirement accounts and the modeled portfolio base of deposits, direct bonds and funds. Debt and outside assets are divided by that portfolio base before bin averaging. Coefficients a_H, b_A, b_D jointly translate the ratios into exposure.
Household tax products	Dated SCF tax-unit supplements and federal calculations, with annual state-rate snapshots for high-tax households and record-level bracket assignments for lower-tax households. The products are computed before weighting. For lower-tax households, the archived bracket tables lack some filing-status cutoffs or related rules in 209 of 792 state-years. These are imputed from the state’s nearest-year thresholds or simple proxies, leaving published rates and top rates unchanged. The repeated national geography is a maintained input.
Insurer portfolios	Annual P&C and life bond-position files, matched to security tax status and remaining maturity. Positive-value positions determine four maturity-dollar and six composition conditions, 10–13 and 17–22. Three annual cross-sectional moments are equally averaged within a block. Portfolio moments inform the estimated exposure mixtures.
Institutional tax treatment	The model applies separate P&C proration, life policyholder-share proration, and bank financing-tax rules, and the residual group is tax-neutral. The measured annual bank funding-cost ratio uses interest expense relative to assets, weighted by municipal holdings. Bank-qualification and coverage assumptions are maintained inputs.
Municipal supply and geography	Dated exempt outstanding Mergent amounts, a within-bin maturity projection and state shares. The seven states without an income tax form one pooled market. National household mass in each bin is allocated by state municipal supply share.
Taxable supply	Treasury and corporate national totals from Financial Accounts, corporate maturity shares from FISC outstanding amounts. Treasury maturity shape is separately tied to the maturities measured in the Monthly Statement of the Public Debt. No national quantity is re-estimated against yields.
Rate process and yield conditions	Annual persistence and innovations from the 1952–2024 TB3MS sample, temporally aggregated to two years. The common mean is the mean of the two-year corporate-like yield over 2000–2017. The monthly corporate-like spot curve supplies conditions 14–16, converted to annual log yields. The pooled treasury and corporate market is compared with this curve and taxed like corporate interest.
Joint estimated coefficients	The seventeen coefficients in Table 9 enter the same twenty-two averaged conditions. Ownership informs household and bank exposures. Insurer books, maturities and composition inform exposure mixtures, while taxable yields and joint clearing inform risk curvature, residual exposure and the bond-specific variance profile. Equilibrium prices and allocations link all coefficients to all conditions.

Notes. Six three-year blocks cover 2000–2017. Each averaged condition is scaled by one tenth of its observed six-block mean magnitude, and the estimated coefficients are held fixed in the full-state evaluation and the policy counterfactuals. Licensed inputs remain subject to their original access restrictions.

Table 15: The twenty-two targeted conditions at the estimated coefficients: signed average deviations in units of 10 percent of the observed magnitude

	Condition	Scaled deviation
1	High-tax households: municipal stock	2.518
2	Lower-tax households: municipal stock	-2.672
3	P&C: municipal stock	-0.199
4	Life: municipal stock	0.140
5	Banks: municipal stock	0.591
6	Households: taxable stock	-0.561
7	P&C: taxable stock	-2.711
8	Life: taxable stock	0.218
9	Banks: taxable stock	-0.277
10	P&C: municipal maturity dollars	1.227
11	P&C: taxable maturity dollars	-1.904
12	Life: municipal maturity dollars	2.269
13	Life: taxable maturity dollars	-0.572
14	Taxable term spread, ten years	0.276
15	Taxable term spread, twenty years	0.165
16	Taxable term spread, thirty years	0.145
17	P&C: municipal-share second moment	2.833
18	P&C: long municipal fraction	-0.341
19	P&C: long taxable fraction	-5.630
20	Life: municipal-share second moment	-0.379
21	Life: long municipal fraction	-0.197
22	Life: long taxable fraction	0.386
	Half sum of squared deviations	36.2870

Notes. Each entry averages six signed dated conditions in original units and divides by one tenth of the observed six-block mean magnitude of the compared quantity, so an entry of one is a residual equal to 10 percent of that condition's specified observed normalizer. Entries are evaluated at the estimated coefficients in the economy with one national municipal market and are not t -statistics. Rows 1–9 are stock differences, scaled by the observed stocks. Rows 10–13 multiply the model stock by its mean-maturity difference and are scaled by the observed book times its mean maturity. Rows 14–16 are annual log-yield differences. Since the model's two-year yield equals the observed one in every month, they are term-spread errors, scaled by the observed mean term spread over the two-year yield. Rows 17 and 20 compare book-weighted second moments. The four long-fraction conditions are $(L - f^{\text{obs}}B)/B_{\text{window}}^{\text{obs}}$, with B and $B_{\text{window}}^{\text{obs}}$ the model and observed books in the greater-than-two-to-thirty-year window and L its greater-than-twenty-year holdings. Rows 17–22 are scaled by their observed values. The two-year municipal stock is part of the ownership and maturity-dollar conditions but not of that distributional window. Conditions are evaluated before averaging. Products of averaged books and ratios generally give a different object.

Calibration blocks, rates and quantities. The inputs set directly from data come in six three-year calibration blocks, 2000–2002 through 2015–2017, centered on the 2001, 2004, 2007, 2010, 2013 and 2016 Survey of Consumer Finances (SCF) waves, and each block uses its own values. The rate state z of Appendix A.1 follows a first-order autoregression whose annual persistence, 0.854, and innovation standard deviation, 1.63 percentage points, are estimated on 1952–2024 short rates. The mean rate is $\bar{r} = 0.0604$ per two-year period, twice the 2000–2017 average of the two-year corporate zero yield, 3.02% a year. Actual monthly two-year corporate zero yields set the observed rate states (Appendix B.1). Financial Accounts stocks fix each sector's portfolio assets, funding and bond quantities. Mergent outstanding amounts fix the municipal maturity profile and issuer-state shares, and outstanding amounts in the Mergent Fixed Income Securities Database (FISD) fix the corporate maturity profile. Tax rates follow statutes and each SCF tax unit's filing status and income, and households' balance-sheet ratios come from the SCF, as described below. Appendix B.6 gives the remaining details, and Appendix Table 14 lists all measured inputs and their sources.

Investor types. Table 3 lists the six groups. Baseline portfolio assets m_i are type i 's debt securities and cash in the Financial Accounts. A fraction W_i of these assets falls within the markets included in the model, giving an investment budget of $m_i W_i$. This budget includes municipal bonds, corporate bonds excluding asset-backed securities, treasury notes and bonds, and cash excluding bank cash. One dollar of portfolio assets is one unit of investor mass, so positions are measured per dollar and multiplied by m_i to obtain aggregate holdings, as in Section 5.1. With common risk aversion, an active holder's tax-mix weight is proportional to its portfolio assets (Appendix Proposition A.5). Each type also has relative after-tax share ξ_i and outside duration $h_i := -\ell_i$, the negative of the exposure in Appendix equation (36). Positive h_i describes outside assets that lose value when rates rise and add to bond risk. Negative h_i describes a gain when rates rise, such as a fall in liability value, that offsets bond losses.

Households are grouped by each SCF tax unit's federal rate under that year's law. High-tax households face a federal rate of at least 35%. In 2015–2017 they are about 5% of families and hold 46% of household portfolio assets. Lower-tax households, banks and insurers hold over half of measured municipal bonds. Within each household group, tax units are sorted by outside financial assets less debt, relative to the deposits, bonds and bond funds that fund the portfolio, into seven bins with equal amounts of portfolio assets. Bins share group tax rates but have outside duration $h_k = a_H + (b_A Out_k - b_D Debt_k)/2$, where Out_k and $Debt_k$ are bin k 's ratios of outside financial assets and of debt to that portfolio base. The three common coefficients a_H , b_A and b_D are estimated, and the factor of one half only normalizes b_A and b_D . For state taxes, high-tax households pay their home state's top rate, and lower-tax households pay each tax unit's bracket rate. The baseline multiplies federal and state after-tax shares for taxable interest tax unit by tax unit. It gives no federal deduction credit for state taxes on out-of-state municipal interest, and Appendix D.4 varies that convention.

The federal calculation includes a Social Security taxation proxy: the tax induced when exempt interest makes more Social Security benefits taxable. The proxy equals the federal tax change from a small increase in ordinary income, which also makes more benefits taxable, less the tax on that increment itself. This proxy is part of the household after-tax shares and explains why their absolute home-municipal shares can be slightly below one.

Institutions face the pre-2018 corporate rate of 35% and no modeled state tax. For property and casualty (P&C) insurers, proration taxes 15% of municipal interest at 35%, so $\xi_P = 0.65/0.9475 = 0.686$. For life insurers, the policyholders' share of municipal interest, one half before 2018, is taxed at 35%, so $\xi_L = 0.65/0.825 = 0.788$. Banks pay no tax on municipal interest and 35% on taxable interest, giving $\xi_B = 0.65$, but cannot deduct the interest expense that funds municipal holdings. Their financing-tax cost, 35% of the ratio of interest expense to assets weighted by municipal holdings, is 1.02% a year in 2000–2002 and 0.11% in 2015–2017. Including it as τ_i^{fin} in the stationarity condition (107) raises the bank's required return by the same amount on every municipal bond. The calibration applies non-bank-qualified treatment (Appendix D.1). The tax-neutral residual group, $\xi = 1$, holds other cash and taxable bonds, mainly through foreign investors, pensions, governments, money market funds and nonprofits. Estimated institutional outside durations comprise two P&C and three life-insurer types with separate shares of sector portfolio assets, one bank type and a seven-point Gauss–Hermite approximation to a normal distribution of outside durations for the residual group.

Issuer-state markets. Each household type is replicated in each of the 45 markets, with its national portfolio assets allocated by the issuer state's share of municipal supply and its relative after-tax share set at that state's rate. The quantitative assignment applies a home-state exemption in every income-tax jurisdiction and taxes out-of-state municipal interest at the resident state's rate. Institutions stay national and choose across all markets, whose independent bond-specific risks offer geographic diversification. The full economy has $45 \times 14 + 13 = 643$ investor types: fourteen household types in each market and thirteen institutional types.

Estimated coefficients. Seventeen coefficients, common to all six calibration blocks, govern what the data do not measure directly and are estimated jointly (Appendix Table 9). Three set the price and amount of risk. Risk aversion γ prices rate and bond-specific risk for every investor. A bond with n years to maturity has two-year bond-specific return variance $v_{30}(n/30)^{p_v}$ in every municipal market and the taxable market for $n \geq 4$, with the two-year claim riskless (Appendix B.5). Three household coefficients, a_H , b_A and b_D , map measured balance sheets into outside durations. Eight set the outside durations and shares of

sector portfolio assets of the P&C and life-insurer types, one sets the bank's outside duration, and two set the mean and dispersion of the residual group's outside durations. Outside durations are measured in the units of a bond's rate loading: an outside duration h has the rate risk of investing all of a type's portfolio assets in a bond whose log price falls by h per unit rise in the rate state.

The 2018–2020 economy. For the 2018–2020 curve comparison, I hold all seventeen coefficients fixed and update the measured inputs. The extension uses the 2019 SCF composition with annually inflation-adjusted incomes, mean 2018–2020 end-year stocks and dated federal tax rules. Institutions face the 21% corporate rate, P&C proration of 25% and a life-insurer policyholders' share of 30%. The annual bank financing-tax cost is 0.117%, and insurer portfolio fractions are averaged over the available 2018–2019 holdings. This seventh economy supplies the post-2017 curve observations without adding calibration targets. Appendix B.3 reports its equilibrium residuals.

B.8 Bond-specific return risk and its normalization

The calibrated model adds the independent shock in (31) to the main-text return. It can represent an issue-specific change in trading demand affecting sale proceeds without moving aggregate rates. Its position-dependent variance cost makes bonds imperfect substitutes and supports joint holdings, as Figure 14 illustrates. The shock is zero for the one-period claim. In the one-market calibration, 20% changes in its variance move the thirty-year gap by less than 0.1 basis points (Table 12).

Writing the shock in realized resale proceeds, $P_{T-1}^a(z')e^{\sigma_{\eta,T}^a\eta}$, instead of multiplying the holding return is algebraically identical. It leaves every return, conditional moment and portfolio condition unchanged. No additional persistent price state is introduced.

Its lognormal mean can also be normalized to one without changing holding returns. Replace the shock factor by $e^{\sigma_{\eta,T}^a\eta - (\sigma_{\eta,T}^a)^2/2}$ and shift log prices to

$$\check{p}_T^a(z) = p_T^a(z) - \frac{1}{2} \sum_{k=1}^T (\sigma_{\eta,k}^a)^2, \quad \check{P}_T^a(z) = e^{\check{p}_T^a(z)}.$$

The adjacent-maturity shifts exactly offset the normalization:

$$\frac{\check{P}_{T-1}^a(z')e^{\sigma_{\eta,T}^a\eta - (\sigma_{\eta,T}^a)^2/2}}{\check{P}_T^a(z)} = \frac{P_{T-1}^a(z')e^{\sigma_{\eta,T}^a\eta}}{P_T^a(z)}.$$

With fixed dollar supplies, portfolios, optimality and clearing are unchanged. For T periods of $\Delta = 2$ years, the annualized yield shift is

$$\check{y}_T^a - y_T^a = \frac{1}{2T\Delta} \sum_{k=1}^T (\sigma_{\eta,k}^a)^2.$$

The fitted variance profile is common to both markets, so both thirty-year continuously compounded yields rise by 0.855 basis points and their difference is unchanged. Corporate normalization uses their ratio, so the normalized thirty-year gap rises from 97.39 to 97.59 basis points, a change of 0.20 basis points. The riskless boundary is unchanged.

C Adjustments and empirical tests

This appendix documents the empirical work behind Sections 3 and 6, beginning with the adjustments in Appendix C.1.

C.1 Comparability and bond-side adjustments

Removing the call option and issuance premium puts municipal and corporate yields on an option-free, seasoned basis and defines the municipal puzzle. The capital-gains, liquidity and default adjustments then

measure how much of this puzzle reflects differences in bond characteristics. Subtracting these three differentials leaves the residual to compare with the model in Sections 3.7 and 5.4.

C.1.1 Object and conventions

Common unit and sample. Each adjustment is a basis-point contribution to

$$\text{Spread}(T) = y_T^{M,\text{raw}} - \bar{\xi}_H(T) y_T^C. \quad (112)$$

The raw municipal zero-coupon curve $y_T^{M,\text{raw}}$ precedes both comparability adjustments and includes callable and noncallable issues, and y_T^C is the corporate curve of Section 3.6. The benchmark $\bar{\xi}_H(T) y_T^C$ averages each bond's household after-tax corporate yield at its dated federal and state tax rates, and $\bar{\xi}_H(T)$ is this average divided by the averaged corporate yield. These rates include the net investment income tax (NIIT) and, from 2018, the cap on state and local tax (SALT) deductions. The household wedge $1 - \bar{\xi}_H(T)$ averages 41.3%.

Both curves and the benchmark are averaged equally within maturity $\times$ issue-quarter $\times$ half-notch rating cells, then across ratings within quarters and then across quarters, the three-stage averaging used throughout. Figure 7(d) shows the resulting puzzle in implied-tax units.

Table 26 reports every adjustment in basis points of the full-sample, all-bond $\text{Spread}(T)$, while each adjustment keeps its own eligible sample and estimator. The credit default swap (CDS) sample used for default starts in 2004 and contains 1,559,971 curve observations. Table 23 reports its spread under the same dated household tax convention, and the default differential estimated there is added to the full-sample accounting.

Additive accounting. The accounting estimates each adjustment separately and adds the contributions in the same basis-point units. It does not estimate interactions between adjustments, so their order of removal leaves the sum unchanged.

Appendices C.1.2–C.1.6 give the five estimators in the order of (23), and Appendix C.1.7 combines them.

C.1.2 Adjustment I: embedded call options

An issuer's right to redeem a bond early is costly to its holder, who therefore requires a higher yield. More than 95% of bonds beyond twenty years are callable, so dropping them would leave a small long-end subsample. I instead value each bond without the call option. Its higher option-free price gives a lower yield, and the difference is the call adjustment.

The call option. Long municipal bonds conventionally allow par calls on coupon dates from year ten. Callable bonds make up 51.1% of curve observations and 51.8% of those at two to thirty years. Investment-grade corporate bonds are usually noncallable bullets or have make-whole strikes that discourage exercise as rates fall. I first fit an option-adjusted spread (OAS) over the model discount curve to reproduce each callable bond's observed price. Keeping that spread and the contractual payments fixed, I value the otherwise identical noncallable bond. The difference between its bullet value V_b^{bul} and the fitted callable value V_b^{call} is

$$CV_b = V_b^{\text{bul}} - V_b^{\text{call}} \geq 0. \quad (113)$$

Adding CV_b to the observed price gives the option-free price. The unchanged OAS retains compensation for credit, liquidity and other factors. The steps below reconstruct contractual payments and call dates, value the bond with and without the call option, and convert the price difference into a yield adjustment.

Step 1: contractual cash flows. Times are in years under the 30/360 US (NASD) day count, without an end-of-February adjustment or business-day calendar: $\text{DC}(u, v) = [360(Y_v - Y_u) + 30(M_v - M_u) + D_v^* - D_u^*]/360$ for year Y , month M and day D , where $D_u^* = \min(D_u, 30)$, and $D_v^* = 30$ if $D_v = 31$ and $D_u^* = 30$, otherwise $D_v^* = D_v$. A bond with remaining maturity T_b and coupon interval $\Delta = 0.5$ years has

$n = \lceil T_b/\Delta - 10^{-9} \rceil$ coupon dates, a first stub $s = T_b - \Delta(n - 1)$ and coupon nodes $t_k = s + \Delta(k - 1)$, $k = 1, \dots, n$.

With $f = \text{DC}(\text{settlement}, \text{first coupon})$, the first paying node $\underline{k}$ satisfies $\underline{k} - 1 = \text{round}((f - s)/\Delta)$. Each bond must satisfy $|f - s - \Delta(\underline{k} - 1)| \leq 0.02$ years, $0 \leq \underline{k} - 1 \leq 2$ and $\underline{k} < n$. Stubs are rounded to 1/360 year to form groups with common $(n, s, \underline{k})$, which keeps t_n within half a day of T_b . The accrual fraction a_k is zero before $\underline{k}$, equals $s + \Delta(\underline{k} - 1)$ at $\underline{k}$ and equals Δ afterward, and the cash flow at node k is $100ca_k + 100 \cdot \mathbf{1}\{k = n\}$ for annual coupon rate c .

If the dated date, from which interest accrues, precedes settlement, accrued interest $100c\alpha$ is added to the clean target price and to the first paying coupon, with $\alpha = \text{clip}(\text{DC}(\text{dated}, \text{first coupon}) - [s + \Delta(\underline{k} - 1)], 0, 1)$. About 29% of observations require this correction.

Step 2: recurring calls. Recorded call times $t^{\text{call}} = \text{DC}(\text{settlement}, \text{call date})$ in $(0, T_b + 0.15)$ are kept, with the lowest strike at ties. Each call maps to node $k = \text{round}((t^{\text{call}} - s)/\Delta) + 1$, which must satisfy $\underline{k} \leq k \leq n$ and $|t_k - t^{\text{call}}| \leq 0.02$ years. The strike is $K_k = \infty$ before the first accepted call date, and thereafter the last recorded strike is carried forward to every coupon date. Special-call provisions are not modeled. A bond marked callable with no accepted call date is excluded, and other bonds without calls have $CV_b = 0$.

Step 3: state-dependent pricing. The issue year selects one of seven dated economies: the six calibration blocks or a 2018–2020 economy using 2018–2019 insurer portfolio fractions and 2018–2020 quantities, rates and taxes. Supply-weighted municipal log prices are interpolated over maturities $\{0, 2, \dots, 30\}$, including $p_0 = 0$:

$$\log P_T(z) = \text{Interp}_T \left\{ \sum_s w_s p_T^{M_s}(z) \right\}, \quad w_s = \frac{\text{municipal supply in state } s}{\text{total municipal supply}}. \quad (114)$$

The OAS is measured relative to this mean curve.

The model's two-year AR(1) in (5) is converted to annual and h -year steps:

$$\rho_y = \sqrt{\rho}, \quad \sigma_y = \frac{\sigma_r}{\sqrt{1 + \rho_y^2}}, \quad \rho_h = \rho_y^h, \quad \sigma_h = \sigma_y \sqrt{\frac{1 - \rho_y^{2h}}{1 - \rho_y^2}}. \quad (115)$$

Before the pricing drift is fitted, pricing yield volatility is set to measured volatility times a scale m by rescaling log prices around $z = 0$:

$$\log P_T^{\text{scaled}}(z) = \log P_T(0) + m\kappa_T[\log P_T(z) - \log P_T(0)], \quad \kappa_T = \sigma_T^{\text{data}}/\sigma_T^{\text{model}}. \quad (116)$$

Here $\sigma_T^{\text{model}} = T^{-1}|\partial_z \log P_T|_{z=0}\sigma_y\sqrt{2/(1 + \rho_y)}$ is the model's standard deviation of twelve-month changes in the T -year yield, and σ_T^{data} is its counterpart in AA-and-above zero-coupon yields, measured in the same quarter of consecutive years after averaging rating cells to maturity-quarter means, with fifty quarters per maturity and a count-weighted $\alpha_0 + \alpha_1 e^{-T/\alpha_2}$ smoothing fit.

The option model fits the conditional mean μ_j of the next rate state at each grid node z_j . This option-pricing mean is distinct from the physical conditional mean and the holding return μ_T^a . The fit uses the model's one-step prices at $T = 1, 1.5, \dots, 30$, with fifteen-point Gauss–Hermite nodes x_q and weights w_q :

$$\mu_j = \arg \min_{\mu} \sum_T \left[\log P_{\Delta}(z_j) + \log \sum_q \frac{w_q}{\sqrt{\pi}} P_{T-\Delta}(\mu + \sigma_{\Delta}\sqrt{2}x_q) - \log P_T(z_j) \right]^2. \quad (117)$$

Prices and drifts are interpolated to 201 nodes spanning the model grid. Transitions Π_{z_j} over one coupon interval split quadrature mass between adjacent nodes and clamp tails to the end nodes z_1 and z_N . Over the first stub, $\Pi_{z_j}^{(s)}$ uses variance σ_s^2 and drift $\mu^{(s)}(z) = z + (s/\Delta)[\mu(z) - z]$. Each bond starts at $x_0 = z_t$, the observed rate state of Appendix B.1 in the final month of its issue quarter: $z_t = r_t - \bar{r}$, where r_t is

twice the continuously compounded two-year corporate zero yield. Every observed state lies inside $[z_1, z_N]$. Contractual dates remain unchanged.

Step 4: callable value and OAS. Starting from $V_n(z) = 100ca_n + 100$, the callable value is computed backward at a flat continuous spread θ :

$$C_k(z) = P_\Delta(z)e^{-\theta\Delta} \sum_j \Pi_{zj} V_{k+1}(z_j), \quad V_k(z) = 100ca_k + \min\{C_k(z), K_k\}. \quad (118)$$

The issuer calls when the continuation value $C_k(z)$ exceeds the strike K_k , and exercise is ex-coupon. The first stub is discounted using $P_s(z)e^{-\theta s} \sum_j \Pi_{zj}^{(s)} V_1(z_j)$ and interpolated at x_0 . The OAS θ^* equates this value to the dirty callable price and is found by 42 bisections on $[-0.20, 0.30]$, which leave an interval of width about 10^{-13} and a repricing error below 10^{-9} per 100 face.

Step 5: bullet value. Removing the call cap in (118), with the OAS and cash flows held fixed, gives the call value $CV_b = \max\{V^{\text{bul}}(\theta^*) - V^{\text{call}}(\theta^*), 0\}$ and the call-adjusted price $P_b^{\text{adj}} = P_b^{\text{dirty}} + CV_b$.

Step 6: zero-coupon yields. On the maturity-interpolated mean discount curve D_0 at $z = 0$, Newton's method solves for a flat continuous spread $\hat{s}$, separately for the raw dirty price treated as a bullet and for the adjusted price:

$$100c \sum_k a_k D_0(t_k) e^{-\hat{s}t_k} + 100D_0(T_b) e^{-\hat{s}T_b} = P_b^{\text{target}}, \quad (119)$$

$$y_b^{\text{target}} = 200 \left[\exp \left\{ \frac{y_0(T_b) + \hat{s}}{2} \right\} - 1 \right], \quad y_0(T_b) = -\frac{\log D_0(T_b)}{T_b}. \quad (120)$$

These nominal semiannual yields, in percent, give a call spread $y_b^{\text{raw}} - y_b^{\text{adj}}$ that is nonnegative because the adjusted price is no lower.

From bonds to the curve. The 1,828,379 A-or-above pricing bonds are drawn from 1,990,791 issues (Table 1). Bonds must have $T_b \in (1, 30]$, at least two coupons, a complete offering price, coupon and yield, fixed or original-issue-discount (OID) coupons, an offering price of at least $100 - 0.25T_b$ per 100 face and consistent state-tax flags. Calendar and price screens exclude 16,295 bonds for irregular first coupons, unsupported conventions, missing accrual dates or unusable call schedules. Eligible bonds retain 97.6%–99.8% of the three-stage weights at each maturity. The three-stage mean of their call spreads is $\text{Call}(T)$ in (23), which is subtracted from the observed curve.²³

The call table uses the screened A-or-above pricing sample, while the main puzzle retains the full observed curve and subtracts the same call adjustment. Screening changes three-stage municipal yields by at most 1.1 bp and the household benchmark by less than 0.1 bp.

Identification of the volatility level. The scale m is fitted to AA-and-above callable and noncallable bonds matched within maturity $\times$ issue-quarter $\times$ rating $\times$ half-point-coupon cells. Their mean callable-minus-noncallable yield difference, averaged across cells and then quarters, is the market gap $\text{gap}_T^{\text{mkt}}$, with across-quarter standard error se_T , and model call spreads in the same cells give the model gap $\text{gap}_T(m)$. If callability is the only difference, the correct m equates the two gaps. The fitted m is applied to A-or-above bonds without re-estimation. Runs at $m = 1$ and 1.15 give the local elasticity

$$e_T = \frac{\ln \text{gap}_T(1.15) - \ln \text{gap}_T(1)}{\ln 1.15}, \quad (121)$$

²³Premium bonds with early par calls can require a negative OAS and produce extreme call spreads, and they remain in the sample. Under full three-stage weights, call spreads above 100 basis points contribute at most 0.34 basis points to any maturity's three-stage mean, and 0.25 at thirty years.

so that $\text{gap}_T(m) \approx \text{gap}_T(1) m^{e_T}$ locally. The scale m is fitted by weighted least squares at integer maturities six through twenty with twelve or more quarters and positive market and model gaps:

$$\ln \text{gap}_T^{\text{mkt}} = \ln \text{gap}_T(1) + e_T \ln m + \varepsilon_T, \quad \text{weight } w_T = (\text{gap}_T^{\text{mkt}} / \text{se}_T)^2, \quad (122)$$

whose solution is

$$\ln \hat{m} = \frac{\sum_T w_T e_T \ln(\text{gap}_T^{\text{mkt}} / \text{gap}_T(1))}{\sum_T w_T e_T^2}. \quad (123)$$

The estimate $\hat{m} = 1.077$ has a conditional 95% band of $[0.964, 1.189]$ from 2,000 bootstrap draws of independent quarters, which preserve within-quarter dependence across maturities and fix $\text{gap}_T(1)$, e_T and w_T . The band contains measured volatility ($m = 1$).

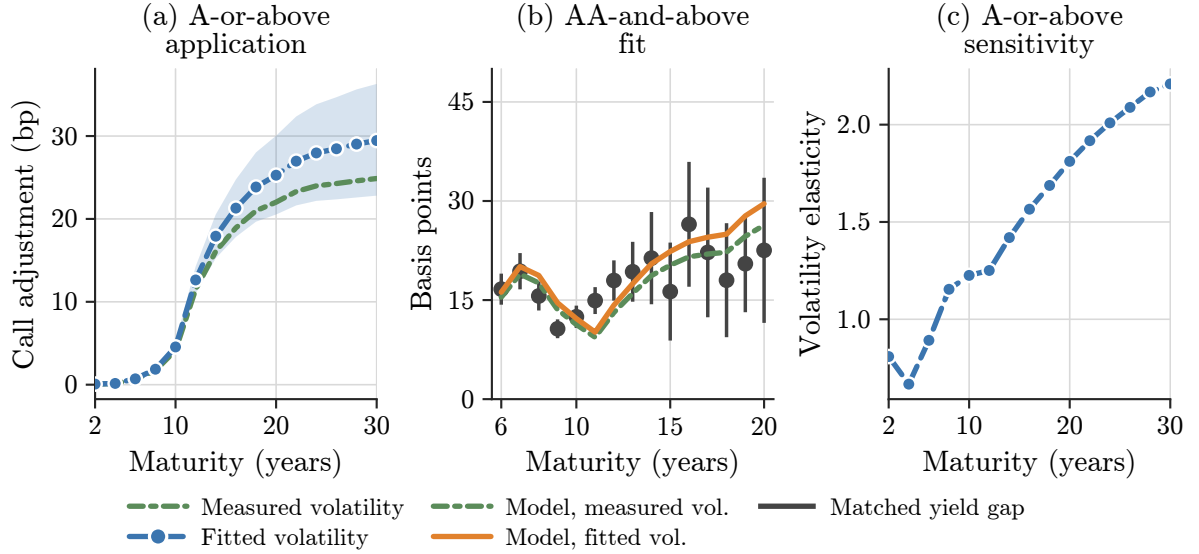


Figure 15: Call adjustment and volatility calibration. (a) A-or-above call adjustments at measured and fitted volatility, with direct-repricing endpoints at the bootstrap band of $\hat{m}$. (b) AA-and-above coupon-matched market gap (± 1.96 standard errors) and model gaps at the fitting maturities, 6–20 years. (c) All-bond elasticity e_T^{all} of the A-or-above call adjustment, distinct from the matched-cell elasticity e_T used in the volatility fit.

Result. At $\hat{m}$, the call adjustment for screened A-or-above bonds is small through ten years, 4.6 bp at ten, and reaches 29.5 bp at thirty (Table 16 and panel (a) of Figure 15), with a directly repriced range of $[22.8, 36.3]$ bp. Measured volatility alone gives 24.9 bp.

Direct-repricing validation of the elasticity. At $\hat{m}$ and both band endpoints, which are repriced directly, the all-bond elasticity approximation is within 0.156 bp of direct repricing at all fifteen reporting maturities.

Validation against the noncallable subsample. The call-adjusted all-bond and noncallable curves in shared cells (Figure 16) differ by less than three basis points through ten years. On the displayed two-year maturity grid, the largest absolute difference is eight basis points, at eighteen years. At twenty years the adjusted all-bond yield lies 3.5 basis points below the noncallable yield.

Robustness to option-pricing dynamics. I repeat the valuation with a normal one-factor Hull–White model (Hull and White, 1994), whose time-varying drift fits each dated input discount curve exactly, with separately set mean reversion and volatility. With annual mean reversion fixed at 0.10, fitting volatility

Table 16: Call adjustment by maturity for screened A-or-above bonds (bp, except the elasticity e_T^{all}). The central estimate is a direct repricing at $\hat{m}$, and the range reprices at the 2.5th and 97.5th percentiles of the bootstrap distribution of $\hat{m}$. The measured-volatility column ($m = 1$) and e_T^{all} are diagnostics.

T	measured ($m = 1$)	e_T^{all}	central ($\hat{m}$)	range
2	0.1	0.81	0.1	[0.1, 0.1]
4	0.1	0.67	0.2	[0.1, 0.2]
6	0.7	0.89	0.7	[0.6, 0.8]
8	1.7	1.15	1.9	[1.6, 2.1]
10	4.2	1.23	4.6	[4.0, 5.1]
12	11.5	1.25	12.6	[11.0, 14.3]
14	16.1	1.42	17.9	[15.3, 20.5]
16	18.9	1.57	21.3	[17.8, 24.8]
18	21.0	1.69	23.9	[19.7, 28.0]
20	22.0	1.81	25.3	[20.5, 30.0]
22	23.3	1.92	27.0	[21.6, 32.4]
24	24.0	2.01	28.0	[22.2, 33.8]
26	24.3	2.09	28.5	[22.4, 34.7]
28	24.6	2.17	29.0	[22.6, 35.6]
30	24.9	2.21	29.5	[22.8, 36.3]

to the same AA-and-above coupon-matched gaps at six to twenty years gives diffusion volatility 0.0126, with rates in decimals and time in years. Cash flows, call schedules and the OAS procedure are unchanged. Probability-weighted paired repricings across the seven dated economies give A-or-above adjustments of 4.7, 23.4 and 28.5 basis points at ten, twenty and thirty years, compared with 4.6, 25.3 and 29.5 in the baseline.²⁴ The root-mean-square error of the Hull–White matched-gap fit is 3.7 basis points, against 4.2 for the baseline model. In the same noncallable-validation cells, the adjusted curve lies 1.9 basis points below the observed noncallable curve at twenty years, compared with 3.5 in the baseline.

C.1.3 Adjustment II: the issuance premium

Spread(T) compares municipal yields at issuance with a fitted curve of seasoned corporate bonds. The adjustment Prem(T) in (23) removes the positive, placement-related difference between offering and secondary-market yields.

Offering and zero-coupon yields. The panel has 1,828,379 A-or-above bonds, each observed once at issuance. The premium is measured in offering yields and subtracted one for one from the zero-coupon curve. At five, ten and thirty years, the two yields have correlations 0.998, 0.993 and 0.937, and the conversion from coupon to zero-coupon yields raises mean yields by 2.8, 12.7 and 58.0 bp.

Measurement. Trades come from the Municipal Securities Rulemaking Board (MSRB) dealer-to-customer tape. Both sides use customer purchases, because an offering is an investor purchase, and comparing it with customer sales would add the bid–ask spread and double-count liquidity (Appendix C.1.5). Same-week seasoned bonds control for market drift:

$$\text{Premium}_b = [y_b^{\text{off}} - \bar{y}_k^{\text{seas}}(w_b^{\text{off}})] - [y_{bj}^{\text{tr}} - \bar{y}_k^{\text{seas}}(w_{bj})]. \quad (124)$$

Here j indexes eligible prints, w weeks and k maturity buckets. Each bond’s premium averages over its eligible prints, customer purchases within thirty days of the offering date excluding when-issued trades. The

²⁴The comparison reprices a probability sample of 3,136 draws, covering 2,843 distinct callable contracts, under both models. Weighted mean differences from baseline pricing are added to the baseline population adjustments, which keep the same three-stage weights and zero call values for noncallables. The conditional numerical-integration standard error of the paired thirty-year difference is 0.057 basis points.

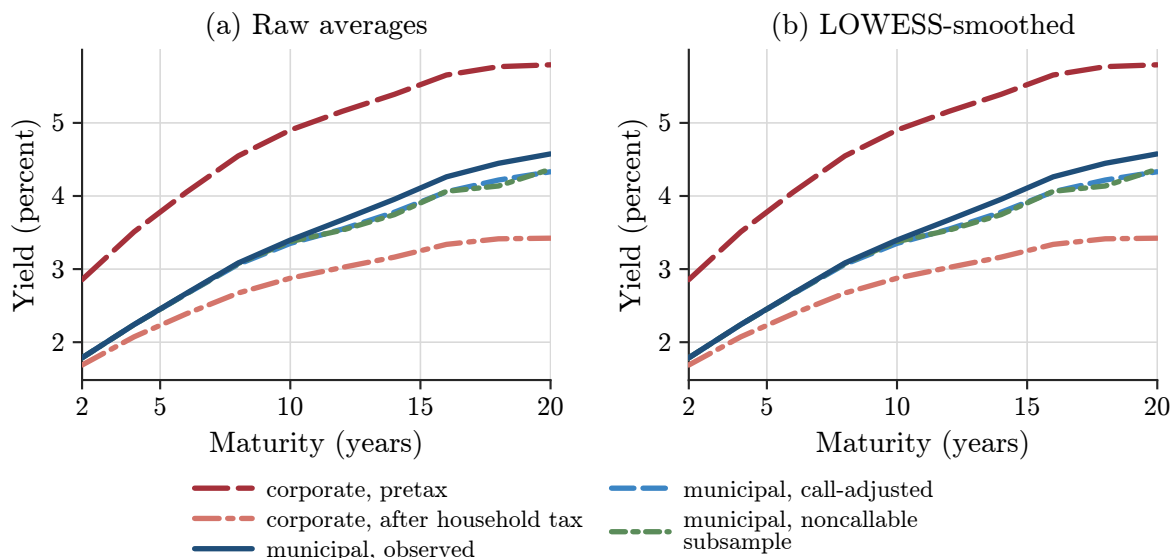


Figure 16: Noncallable validation of Table 16: raw means (a) and locally weighted scatterplot smoothing (LOWESS) (b). Lines show pretax and household after-tax corporate yields alongside observed, call-adjusted and noncallable municipal yields. The call-adjusted and noncallable curves share A-or-above maturity-quarter-rating cells containing noncallables, with fifty or more quarters per maturity and no extrapolation past twenty years. Remaining coupon, credit and other differences prevent identification of individual option values.

benchmark is the par-weighted yield of same-week customer purchases of bonds in the bucket older than three years, with at least 25 prints per cell. Differencing removes common rate changes and fixed credit differences, while changing credit and call exposure can remain. Early and seasoned prints must have at least \$250,000 par, and prints of all sizes give 21.1 bp at thirty years (Panel B of Table 17).

Sample and aggregation. The sample is restricted to the curve's own filtered population. Using all Mergent offering yields would put only 74% of estimation bonds in that universe. Trade age and benchmark weeks are measured from the offering date. The dated date differs from it by a median of fourteen days, by more than a week for 82.7% of bonds and by more than a month for 8.4%, and it precedes the offering for 13.7%.

The three-stage averaging of Appendix C.1.1 is applied within issue-quarter and rating cells and the maturity buckets 1–5, 5–8, 8–12, 12–17, 17–23 and 23–31 years. Estimates and bootstrap endpoints are interpolated from the bucket midpoints (3, 6.5, 10, 14.5, 20, 27) to the two-year grid, with ends held flat. The longest-bucket mean is 11.9 bp with par weights, which measure dollars, and 12.8 with curve weights, which measure curve observations.

Result. The issuance premium rises steeply with maturity, from 2.9 bp at two years to 12.8 bp at thirty (Table 26 and panel (b) of Figure 17). Its range comes from a 10,000-draw issue-quarter bootstrap that samples four-quarter blocks with replacement and endpoint wrapping. Lag-one autocorrelations of quarterly bucket means range from -0.34 (longest bucket) to 0.02 . The block bootstrap and a bootstrap of independent quarters give similar short-end intervals, $[1.7, 4.2]$ and $[1.8, 4.0]$ at 1–5 years. The thirty-year interval is $[8.1, 17.3]$ bp, and the intervals exclude zero at every maturity. Intervals condition on the matched sample and seasoned benchmark.

Sensitivity to the trade-age window. In all five trade-age windows of Table 17, premia beyond twelve years exceed those below twelve years. From fourteen days on, the two longest buckets give 9.3–12.8 bp. Three- and seven-day samples have fewer institutional-size prints and are noisier. Panel B shows the same gradient at higher levels.

Table 17: Issuance premium by maturity bucket (bp) under alternative trade-age windows, in days; the baseline uses thirty. Panel A, the baseline, requires prints of at least \$250,000 par in both the early print and the seasoned benchmark. Panel B uses prints of all sizes.

maturity bucket (yr)	3	7	14	30	60
<i>Panel A: institutional-size prints (baseline)</i>					
1–5	3.7	3.6	1.4	2.9	3.3
5–8	2.8	5.2	3.0	3.3	4.0
8–12	3.4	5.2	3.9	4.0	4.6
12–17	6.2	9.0	7.7	7.5	7.8
17–23	9.1	7.3	9.3	10.8	12.2
23–31	9.5	7.2	11.9	12.8	12.0
<i>Panel B: prints of all sizes</i>					
1–5	4.5	6.7	5.6	3.6	2.2
5–8	5.9	10.1	7.6	6.1	5.0
8–12	7.2	10.5	9.0	7.9	7.4
12–17	15.5	16.2	13.8	12.8	12.4
17–23	15.0	17.3	16.7	16.6	16.7
23–31	18.1	18.7	20.0	21.1	21.0

Notes. Each cell is the drift-controlled difference-in-differences premium for bonds with at least one print inside the window, averaging each bond’s prints inside the window and then averaging bonds with the curve’s three-stage weights (Appendix C.1.1). Column heads are windows in days.

Timing, conventions and size. Across the six maturity buckets, premia average 5.8 bp within three days of the offering, 6.4–7.0 at three to thirty days and 8.5 and 11.6 at thirty to sixty and sixty to ninety days. Changing composition across these age bins prevents a within-bond interpretation. For par and discount bonds the two longest buckets give 10.9 and 13.2 bp, and for premium bonds 9.3 and 10.9, so yield-to-worst differences do not explain the gradient. Long-end premia are positive in every trade-size decile of all-size prints and smallest for the largest trades (panel (c) of Figure 17). Small-trade markups also raise the all-size estimates.

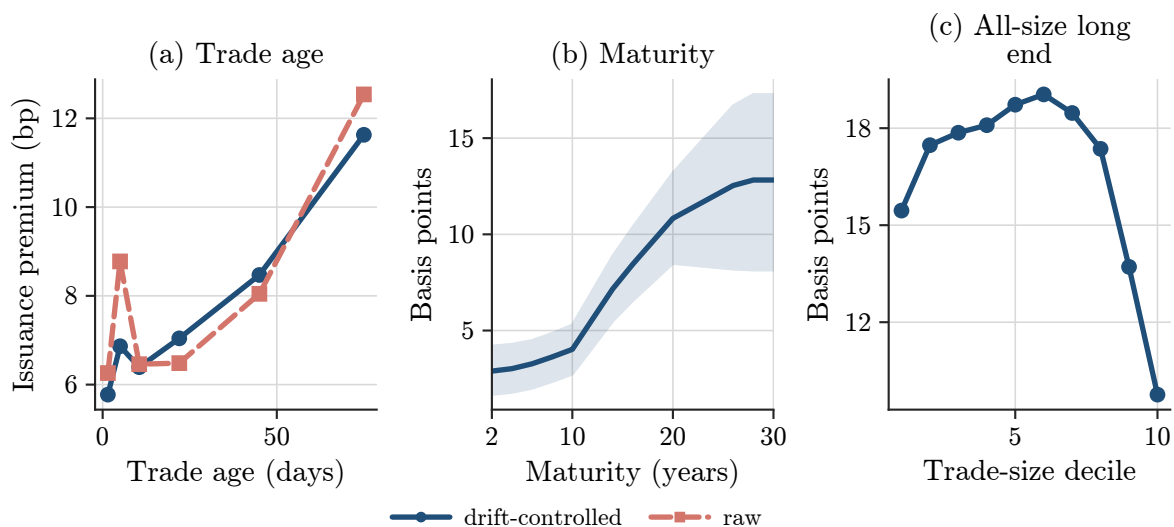


Figure 17: Issuance-premium diagnostics. (a) Institutional-size estimates by trade age, with and without drift controls. (b) Institutional-size maturity profile with block-bootstrap band. (c) All-size long-end estimates by trade-size decile.

Interpretation. The drift-adjusted post-offering yield change captures the mismatch between primary and seasoned pricing that matters for the curve comparison. Changing dealer markups, ongoing distribution, liquidity seasoning and price discovery may also contribute, so I treat the change as a proxy for the issuer’s placement cost.

Sample coverage. Only bonds with an institutional-size print within thirty days of the offering contribute, so the estimate covers 4% of the target curve, and the sign of the resulting selection bias is not determined here. The trade tape begins in 2005 and the curve spans 2000–2020, so the premium is measured on later issues and applied to the whole curve. It also inherits any staleness in recorded offering yields.

C.1.4 Adjustment III: capital-gains taxation

Municipal interest is exempt, but resale gains and market discount can be taxable. Corporate gains can also face different rates from interest. The model’s wealth equation (32) instead applies interest-income tax treatment to the whole holding-period return. The adjustment $CG(T)$ measures the effect of separating these tax treatments.

The comparison. For each bond pair and sale horizon, I find the municipal issuance-yield shift that equates the sellers’ after-tax returns under two mappings. Mapping B uses the model’s tax treatment. Mapping A separates interest, market discount and capital gains under the assumptions below. Both calculations include the next buyer’s resale valuation and the original seller’s taxes. The difference between the shifts isolates the tax correction from existing curve differences. First-sale probabilities weight this correction across horizons, with zero adjustment for bonds held to maturity.

Taxed amounts and rates. Appendix D.5 gives the accrual and basis rules. Municipal interest and OID are exempt. Market discount and gains are taxable. OID raises basis, original-issue premium (OIP) lowers it without a municipal deduction, and acquisition premium reduces OID. Taxable interest and OID are ordinary income, and premium amortization assumes the Section 171 election. Original buyers use basis A_h (Step 1), and secondary buyers start from the purchase price.

Market discount below a quarter point per full remaining year is de minimis and treated as capital gain. Larger discount is ordinary income at disposition. With n remaining years and face 100, the de minimis threshold is

$$\bar{D}_n = 0.0025 \times 100 \times n. \quad (125)$$

All seven economies use the fixed federal rates $\tau_{oi} = 0.37$ on ordinary income and $\tau_{cg} = 0.20$ on capital gains, excluding NIIT and state tax, unlike the dated household rates in (112). Residual gains and losses are treated symmetrically,

$$\mathcal{T}(G) = \tau_{cg}G, \quad (126)$$

so losses are fully deductible in this calculation.

Step 1: the instruments. Each market’s par and zero-coupon instruments use its own discount factors d_n , from the model’s supply-weighted municipal mean log prices or corporate prices at $z = 0$, interpolated annually from two-year knots. At even $T = 4, \dots, 30$ and sale horizons $h = 2, 4, \dots, T-2$, the two instruments are

$$\text{PAR: } c = \frac{1 - d_T}{\sum_{n=1}^T d_n}, \quad \text{offer} = 100, \quad A_j \equiv 100; \quad \text{ZERO: } \text{offer} = 100 d_T, \quad A_j = \text{offer} \cdot e^{y_0 j}, \quad (127)$$

with $y_0 = -\log(d_T)/T$, where A_j is the original buyer’s accreted basis after j years.

For a zero purchased at issue, the interest accruing during year $j + 1$ is

$$A_{j+1} - A_j = A_j(e^{y_0} - 1), \quad A_T = 100.$$

At a sale in year h , the original buyer's residual capital gain is $P_h - A_h$, so accrued interest is not taxed again as a gain.²⁵

Step 2: rate states at the resale date. Sale-date rate states use composite Gauss–Legendre nodes z_i and weights ω_i , split where either market's tax regime changes, and the grid-truncated AR(1) density of the state h years after $z = 0$:

$$w_i \propto \omega_i \frac{1}{\sigma_h} \varphi\left(\frac{z_i}{\sigma_h}\right). \quad (128)$$

Here φ is the standard normal density and σ_h is given by (115). Results with twenty-four and with twelve nodes per interval differ by less than 0.001 bp across 2,940 cells, and the residuals of (136), in annualized log return, are at most 10^{-12} .

Step 3: the corporate anchor. The seasoned corporate equilibrium price $P_h^{\text{corp}}(z)$ at the sale date is held fixed. For each mapping $\mathcal{M} \in \{B, A\}$, the remaining annual after-tax cash flows $X_j^{\mathcal{M}}$, contractual payments less the tax due under $\mathcal{M}$, determine a buyer's continuous after-tax return at that price:

$$\sum_{j=1}^{T-h} X_j^{\mathcal{M}} e^{-r_{\mathcal{M}}^*(h,z)j} = P_h^{\text{corp}}(z). \quad (129)$$

Flows are annual, with continuous discounting and no day-count or semiannual convention. For a corporate buyer holding to maturity, the two mappings differ only in their de minimis treatment of market discount D :

$$\text{rate}_{\mathcal{M}}(D) = \begin{cases} 0, & D \leq 0, \\ \tau_{cg}, & \mathcal{M} = A \text{ and } D < \bar{D}_n, \\ \tau_{oi}, & \text{otherwise.} \end{cases} \quad (130)$$

Step 4: the municipal resale price under each mapping. A municipal buyer holding to maturity under $\mathcal{M}$ must earn $r_{\mathcal{M}}^*$. Benchmark municipal flows CF_k are exempt, giving

$$P^B(h, z) = \sum_{k=1}^{T-h} CF_k e^{-r_B^*(h,z)k} \equiv G_B(h, z). \quad (131)$$

The actual price and tax depend jointly on the market discount $A_h - P$. Let G_A be the gross present value at the actual anchor and $q = e^{-r_A^*(h,z)(T-h)}$ the terminal discount factor. Taxing the discount once at maturity at rate τ gives $P = G_A - q\tau(A_h - P)$, which is linear in P , so

$$P(\tau) = \frac{G_A(h, z) - q\tau A_h}{1 - q\tau}, \quad (132)$$

and, with $n = T - h$ remaining years, the regime is selected in this order:

$$P^A(h, z) = \begin{cases} G_A, & G_A \geq A_h \quad (\text{no discount: premium, untaxed}), \\ P(\tau_{cg}), & A_h - P(\tau_{cg}) < \bar{D}_n \quad (\text{de minimis}), \\ P(\tau_{oi}), & \text{otherwise.} \end{cases} \quad (133)$$

When both regime-specific prices satisfy their conditions, this order selects the de minimis price, the higher of the two.

²⁵IRS Publication 1212, *Guide to Original Issue Discount (OID) Instruments*, and Publication 550, *Investment Income and Expenses*, describe accrual and basis. Annual effective yield $e^{y_0} - 1$ corresponds to continuous issuance yield.

Step 5: the two shifts. A scalar δ shifts the municipal log-discount curve in parallel:

$$\hat{p}_n^\delta(z) = \hat{p}_n(z) - \delta n \iff y_n^\delta = y_n + \delta. \quad (134)$$

Both sellers are compared using the annualized log of the expected gross holding return:

$$R = \frac{1}{h} \log \frac{\sum_i w_i \mathcal{W}(z_i)}{P_0}, \quad (135)$$

where $\mathcal{W}(z_i)$ is terminal wealth at rate state z_i and P_0 is the offering price. Both sellers reinvest interim coupons to h on the unshifted municipal curve. Corporate wealth carries forward coupons net of τ_{oi} and the annual OID tax, and the residual gain $P_h^{\text{corp}} - A_h^C$ is taxed at τ_{oi} under B and at τ_{cg} under A . Municipal wealth equals the resale price plus exempt coupons, less τ_{cg} on its residual gain under A . Bisection solves for each mapping's shift,

$$R^{\text{muni}}(\delta_B; \text{benchmark}) = R_B^{\text{corp}}, \quad R^{\text{muni}}(\delta_A; \text{actual}) = R_A^{\text{corp}}, \quad (136)$$

and the conditional adjustment is their difference,

$$\Delta_{cg}(h, T) = 10^4 (\delta_A - \delta_B) \text{ basis points.} \quad (137)$$

The adjustment averages across instrument pairs with issuance weights, in continuously compounded annual yields. The difference $\delta_A - \delta_B$ isolates capital-income taxation from curve mismatch between the markets. Each shift recomputes only the municipal issuance coupon or zero price and basis, together with its resale price. Rate states, weights, corporate prices, cash flows and targets, anchors and the reinvestment curve stay fixed.

Instrument types and issuance weights. Issues with positive coupons and prices of at least 90 are PAR-like, including premium issues above 110, and issues priced below 90 or with zero coupons are ZERO-like. In the Mergent Fixed Income Securities Database (FISD), nonpositive prices denote missing values.

Pooled 1998–2021 amount-weighted ZERO shares are 1.3% municipal and 2.2% corporate (Table 18). The baseline shares match original maturities within three years of each node, pooling when offering amounts are insufficient, and are 0.97% and 0.31% at thirty years. Each of the four pairs in $\{\text{PAR}, \text{ZERO}\}^2$ receives the product of its corporate and municipal shares. The PAR/PAR pair receives 98.7% of the weight at thirty years and drives the adjustment.

Table 18: Four instrument pairs at $T = 30$: pooled issuance weights and each pair's maximum and mean conditional adjustment $\Delta_{cg}(h, 30)$ over sale horizons h , averaged across the seven dated economies. The baseline uses maturity-matched shares, which put PAR/PAR at 98.7% at thirty years.

Corporate / Municipal	weight (%)	max over h (bp)	mean over h (bp)
PAR / PAR	96.5	6.51	5.49
ZERO / PAR	2.2	20.07	14.90
PAR / ZERO	1.3	7.71	5.74
ZERO / ZERO	0.0	15.75	10.91

First-sale weights. The conditional adjustment $\Delta_{cg}(h, T)$ is weighted by the original buyer's probability of a first sale at age h . Only first sales count: Step 4 already prices the next buyer's tax under hold-to-maturity, so turnover weights would double-count later dispositions.

Let t_a be seasoned traded par over outstanding par at age a , measured from the MSRB dealer-to-customer tape (71.6 million records, 2005–2021) and pooled over original maturities within ± 4 years of T . To exclude primary distribution, I omit when-issued trades and the first quarter after the dated date. A trading segment,

a share $\phi(T)$ of par, carries all resales, and the remainder is held to maturity. With aggregate turnover fixed, the first-sale hazard at age a is

$$\lambda_a = \min(t_a/\phi, 0.99), \quad (138)$$

where the cap binds only at the first two ages of thirty-year bonds in the central calibration, at which t_a/ϕ would otherwise reach 1.003 and 0.991. The first-sale density and held-to-maturity mass are

$$f_\phi(a) = \phi \lambda_a \prod_{j < a} (1 - \lambda_j), \quad S_\phi(T) = 1 - \phi \left(1 - \prod_{a < T} (1 - \lambda_a)\right), \quad \sum_a f_\phi(a) + S_\phi(T) = 1, \quad (139)$$

and the unconditional adjustment is $\text{CG}(T) = \sum_h \pi_h(T) \Delta_{\text{cg}}(h, T)$, where $\pi_h(T)$ sums f_ϕ over the ages assigned to sale horizon h and the held-to-maturity branch carries a zero adjustment. Each h receives ages $h - 1$ and h . The first node also receives age 0, and the last node, with two years remaining, also receives the final pre-maturity age $T - 1$.

Calibrating the trading segment. Repeated trades in a narrow segment can generate high turnover without nearly universal resale. Cumulative seasoned volume in each CUSIP (bond identifier), capped at its outstanding par, bounds the par ever sold (Table 19). The parameter ϕ is fitted over the observed window and then extended to full life. At thirty years, par weights over original maturities within four years give a ceiling of 0.602 (single-bin 0.617), so $\phi = 0.602$ and $S_\phi = 0.398$, with first sales concentrated at young ages.

Table 19: The measured ceiling on the fraction of par ever sold, for selected one-year bins of original maturity. Cumulative traded par counts repeat trades of the same dollar, so the ceiling is an upper bound, and the capital-gains range also scales it by 0.5 and 0.75. Never traded is the par share of CUSIPs with no seasoned trade. The observation window is the par-weighted mean of the shorter of original maturity and the years from offering to 2021. Entries are par-weighted over issues offered from 2005 with at least three observable years.

orig. maturity	ceiling on $P(\text{ever sold})$	never traded	obs. window (y)	CUSIPs
4	0.466	0.195	3.9	124520
6	0.571	0.155	5.6	121412
8	0.556	0.227	7.0	118265
10	0.625	0.132	8.0	110693
14	0.558	0.171	9.4	73189
18	0.546	0.216	9.7	46382
22	0.586	0.219	10.0	13056
26	0.612	0.193	10.6	6551
30	0.617	0.159	10.3	22151

Result. The thirty-year adjustment is 2.3 bp, with range [1.2, 2.5] bp (Table 20 and panel (b) of Figure 18). Because the tax in (126) is linear, the expected tax on the residual gain depends only on its mean.

The range is the envelope of eighteen variants, each averaged equally over the seven 2000–2020 economies: three tape windows (the full window and its two halves), three segment multipliers (0.5, 0.75 and 1.0) and two weighting schemes (pooled and maturity-matched). For multipliers below one, the segment share is refitted on the full window before first-sale weights are measured on each half window. The central estimate uses the ceiling, and the effects of the window and the weights have either sign.

With the same economies, tax rates and trade tape, the thirty-year adjustment is 4.8 bp conditional on an observed sale and 3.9 bp with a homogeneous hazard ($\phi = 1$). At long maturities the homogeneous hazard exceeds the measured resale ceiling, and its adjustment lies above the band. At short maturities the fitted segment share is $\phi = 1$, so the reported and homogeneous-hazard estimates coincide.

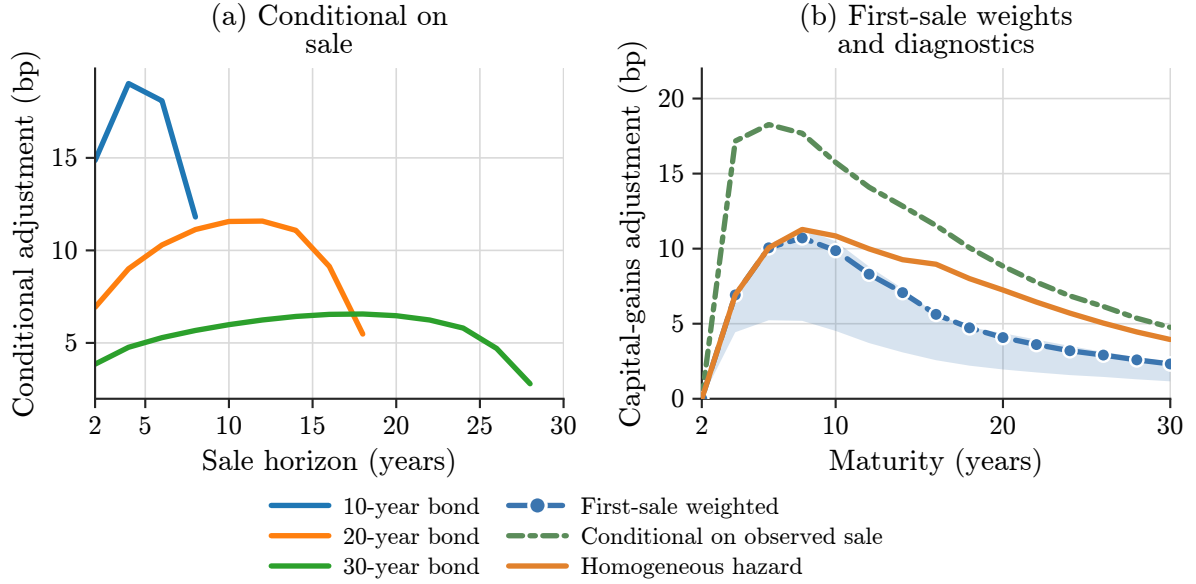


Figure 18: Capital-gains adjustment, averaged over the seven dated economies. (a) Pair-weighted conditional adjustment $\Delta_{cg}(h, T)$ by sale horizon for ten-, twenty- and thirty-year bonds. (b) Adjustment weighted by first-sale probabilities, with the range of its eighteen variants, and the conditional-on-observed-sale and homogeneous-hazard diagnostics.

Table 20: Capital-gains adjustment by maturity (bp, except ϕ and $P(\text{hold})$). The reported column uses first-sale probabilities, with the eighteen-variant envelope as its range. The cond. on sale and homogeneous columns are diagnostics. The maximum is over sale horizons of the dated-average, pair-weighted conditional curve. Columns ϕ and $P(\text{hold})$ give the trading-segment share and the held-to-maturity probability. At two years no sale horizon precedes maturity on the two-year grid, so the adjustment is zero.

T	cond. on sale	max over h	ϕ	$P(\text{hold})$	homogeneous	reported	range
2	0.0	0.0	—	1.00	0.0	0.0	[0.0, 0.0]
4	17.2	17.2	1.00	0.60	6.9	6.9	[4.4, 7.0]
6	18.3	19.2	1.00	0.45	10.1	10.1	[5.2, 10.3]
8	17.7	19.9	0.88	0.40	11.3	10.7	[5.2, 11.3]
10	15.7	19.0	0.82	0.37	10.8	9.9	[4.5, 10.5]
12	14.1	17.8	0.73	0.39	10.0	8.3	[3.7, 8.8]
14	12.8	16.2	0.67	0.40	9.3	7.1	[3.1, 7.3]
16	11.5	14.6	0.58	0.44	9.0	5.6	[2.6, 5.6]
18	10.1	13.0	0.57	0.44	8.0	4.7	[2.2, 4.8]
20	8.8	11.6	0.57	0.43	7.2	4.1	[2.0, 4.4]
22	7.8	10.4	0.58	0.42	6.4	3.6	[1.8, 3.9]
24	6.9	9.3	0.59	0.41	5.7	3.2	[1.6, 3.5]
26	6.1	8.5	0.60	0.40	5.0	2.9	[1.5, 3.1]
28	5.4	7.4	0.60	0.40	4.5	2.6	[1.3, 2.8]
30	4.8	6.6	0.60	0.40	3.9	2.3	[1.2, 2.5]

C.1.5 Adjustment IV: liquidity

Construction. The imputed round-trip cost (RTC) is $(\max - \min) / \max$ price within CUSIP $\times$ day $\times$ size cells that contain both customer buys and customer sells. The extreme prices need not belong to one dealer’s round trip. The comparison uses 2005–2020 MSRB municipal trades and Trade Reporting and Compliance Engine (TRACE) corporate trades of at least \$250,000, made more than three months after issue and more than one year before maturity. Remaining maturity is rounded down to whole years and assigned to the six five-year bins in Table 21. Costs are averaged to bond-months before matching. Trades at the TRACE size caps of 1MM+ and 5MM+ are retained at their lower bounds, and trade size is not otherwise used.

Cells are matched by month $\times$ rating family $\times$ maturity bin, with ten or more bonds per market. Agency letter grades map AAA to notch one, AA to two–four and A to five–seven on the twenty-two-notch scale. Averages weight ratings equally within months, then months within years, then years.

Let c_m and c_c be the matched municipal and corporate costs in percent of price. One purchase and resale over a holding period of h years requires annual compensation of c_m/h and c_c/h in the two markets, giving the yield differential in basis points

$$\Delta y^{\text{liq}} = \underbrace{\frac{100 c_m}{h}}_{\text{municipal cost per year}} - \underbrace{\frac{100 c_c}{h}}_{\text{corporate cost per year}} = \frac{100(c_m - c_c)}{h}.$$

Both markets use $h = 5$. Trading costs reduce after-tax wealth in both markets, so the differential is not multiplied by an interest-income after-tax share. A positive difference raises the yields municipal holders require and is subtracted from the puzzle.

Regressing A-or-above offering yields on bond-average round-trip costs, with maturity-bin $\times$ rating $\times$ offer-year effects, gives 15.4 bp per percentage point of price in round-trip cost, compared with 20 at $h = 5$, so this regression map scales $h = 5$ estimates by 0.772. Standard errors are the standard deviation of sixteen annual differences divided by $\sqrt{16}$. The band spans $h = 2$, $h = 10$, the regression map and the central estimate ± 1.96 standard errors.

The six bin estimates at 3, 8, 13, 18, 23, 28 years are interpolated linearly to 2, 4, $\dots$, 30, with end values held flat, to give $\text{Liq}(T)$ in (23). Thirty years uses the 26–30 bin, and the interpolation avoids artificial jumps across maturities.

Result. Institutional-size municipal round trips are slightly cheaper than corporate ones before twenty years and costlier afterward (Table 21 and Figure 19). At 26–30 years, costs of 0.5335% and 0.5310% imply 0.050 bp annually at $h = 5$, which is reported as 0.1 bp, with a band of $[-2.2, 2.3]$ bp. The higher-cost 21–25 bin produces the peak adjustment of 2.9 bp at twenty-four years.

At $h = 5$, standard errors are 0.8–2.3 bp. The 21–25 estimate is about 1.5 standard errors from zero and depends on the holding period. Appendix C.2 gives separate issuer-level regressions on trading frequency and trading costs, which are not part of the band.

Table 21: Liquidity by remaining-maturity bin (years), matched trades of 2005–2020. Round-trip costs (RTC) and their municipal-minus-corporate difference in percent of price, and adjustments in bp under three holding periods h and the regression map.

Bin	Municipal RTC	Corporate RTC	Difference	$h = 2$	$h = 5$	$h = 10$	regr.
1–5	0.225	0.246	-0.022	-1.08	-0.43	-0.22	-0.33
6–10	0.382	0.403	-0.021	-1.05	-0.42	-0.21	-0.33
11–15	0.582	0.588	-0.006	-0.30	-0.12	-0.06	-0.09
16–20	0.609	0.619	-0.010	-0.49	-0.20	-0.10	-0.15
21–25	0.684	0.506	0.178	8.92	3.57	1.78	2.75
26–30	0.533	0.531	0.003	0.13	0.05	0.03	0.04

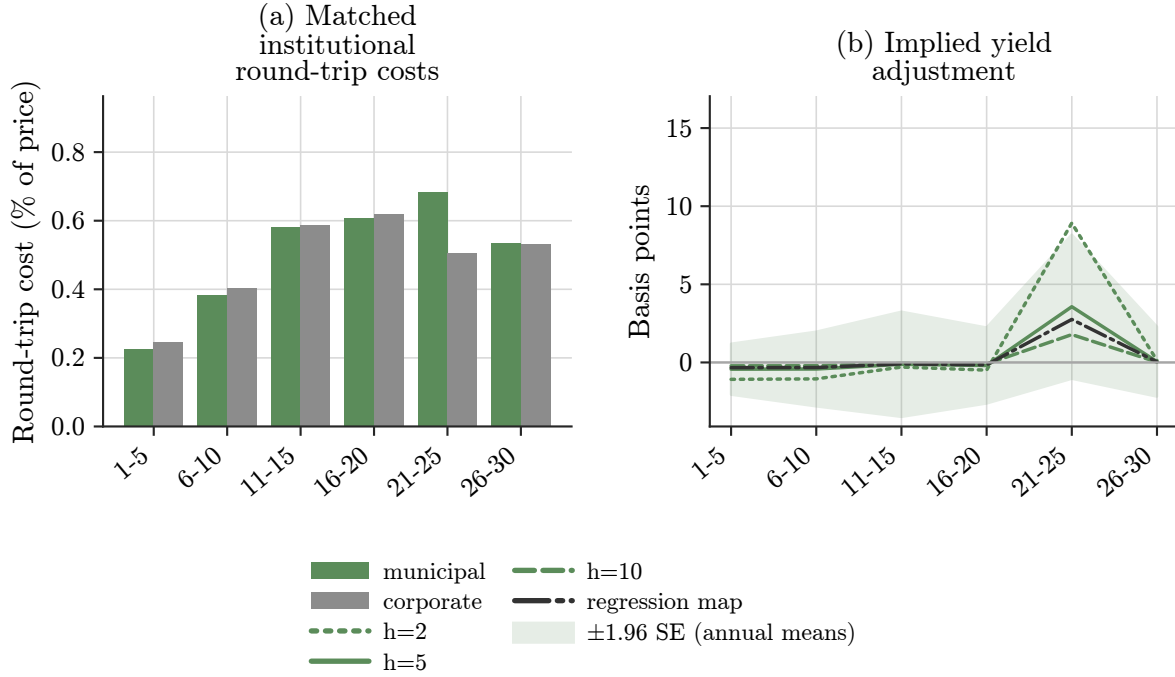


Figure 19: Matched institutional-size customer trades of at least \$250,000, 2005–2020. Panel (a): imputed round-trip costs in common month, rating and maturity cells, percent of price. Panel (b): the municipal-minus-corporate cost difference annualized under three holding periods and under the regression map. Shading is ± 1.96 standard errors calculated across annual means at $h = 5$.

Sensitivity to the cell rule. Table 22 varies the minimum number of bonds in each market’s cells, using the same underlying trades and size screen. A one-bond minimum admits thin cells and changes the 26–30-year estimate from 0.05 to -6.53 bp. A twenty-bond minimum leaves only 19 common months at 11–15 years and changes the estimate there from -0.12 to -2.99 bp. The ten-bond minimum balances coverage against the influence of thin cells.

Table 22: Liquidity adjustment at $h = 5$ (bp) under alternative minimum bond counts per market cell.

bin	≥ 1	≥ 5	≥ 10 (used)	≥ 20
1–5	-0.59	-0.50	-0.43	-0.39
6–10	-0.79	-0.49	-0.42	-0.27
11–15	-0.12	0.42	-0.12	-2.99
16–20	-0.09	-0.30	-0.20	-0.31
21–25	2.05	2.20	3.57	1.81
26–30	-6.53	-5.81	0.05	0.30

C.1.6 Adjustment V: default

Method. The default adjustment measures the part of the municipal yield gap attributed to credit risk. Following the simple taxable-equivalent convention in [Schwert \(2017\)](#), municipal interest divided by the household after-tax share is compared with taxable interest. State-government and pooled AA corporate Markit CDS spreads proxy credit compensation on this scale. Let s_b^M and s_b^C denote the matched spreads for curve observation b , and $\xi_{H,b}$ its dated household after-tax share. Converting both legs to the household after-tax units used in the puzzle gives

$$\text{Def}(T) = \text{Avg}_T[\xi_{H,b}s_b^M] - \text{Avg}_T[\xi_{H,b}s_b^C]. \quad (140)$$

Both CDS quotes are pretax. Both receive the same multiplier because their credit compensation is expressed on the taxable-equivalent scale before conversion to the units of $y_T^M - \bar{\xi}_H(T)y_T^C$. A positive differential is subtracted from the puzzle. The dated after-tax share includes federal and state rates, NIIT from 2013 and the 2018 change in state-tax deductibility. No further corporate-tax or loss-deduction factor is applied.

Matching and extrapolation. State general-obligation (GO) CDS are matched to A-or-above bonds by rating, maturity and quarter, and pooled quarterly AA corporate CDS form the corporate leg. After-tax shares are applied before Avg_T averages within maturity-quarter-whole-notch-rating cells, then across ratings and quarters, retaining each leg’s nonmissing support. The AA-and-above sample of [Appendix C.12.1](#) uses the same construction. Beyond quoted tenors, the cumulative spread $T\text{CDS}(T)$ is extrapolated linearly, with negative filled spreads floored at zero and quoted spreads left unchanged.

Coverage and sensitivity. In the A-or-above CDS panel, exactly observed cells cover 49% at 1–10 years, 47% at 11–20 and 13% at 21–30, and the remaining cells use filled curves (panel (c) of [Figure 20](#)). The sensitivity range is the pointwise minimum and maximum of five variants, each with the same household tax mapping: the baseline, cells with at least half of their source observations quoted or interpolated (quality-capped), the 2007-onward sample and CDS–cash basis factors 0.75 and 0.50 ([Table 24](#)).

Result. The thirty-year default adjustment is 16.5 bp, with a sensitivity range of [8.2, 20.8] bp. It equals 9.9% of the CDS-matched spread and 10.3% of the full-sample denominator $\text{Spread}(T)$, and it remains positive at every maturity ([Table 23](#)).

Table 23: Default by maturity on the A-or-above CDS-matched sample. Yields in percent, and spread, default adjustment and band in bp. The spread is the municipal yield minus the corporate yield after household tax, with each observation’s dated federal and state after-tax share as in the full-sample $\text{Spread}(T)$. The band is the range of [Table 24](#).

T	municipal yield	corporate after-tax	spread (bp)	adjustment (bp)	band
2	1.56	1.51	5.7	16.48	[5.69, 17.64]
4	1.99	1.85	14.6	14.35	[6.55, 16.10]
6	2.39	2.15	24.8	12.18	[5.94, 14.20]
8	2.79	2.42	37.0	12.46	[6.23, 14.95]
10	3.12	2.63	49.5	12.46	[6.23, 15.27]
12	3.47	2.79	68.1	11.55	[5.77, 14.50]
14	3.73	2.89	84.0	11.65	[5.83, 14.85]
16	3.93	2.96	97.1	12.26	[6.13, 15.68]
18	4.09	3.01	108.4	12.10	[6.05, 15.54]
20	4.26	3.04	122.2	12.70	[6.35, 16.30]
22	4.36	3.07	129.0	13.34	[6.67, 17.06]
24	4.47	3.08	138.3	14.08	[7.04, 17.94]
26	4.59	3.09	149.3	15.62	[7.81, 19.77]
28	4.58	3.09	149.5	16.16	[8.08, 20.39]
30	4.77	3.11	165.7	16.45	[8.23, 20.75]

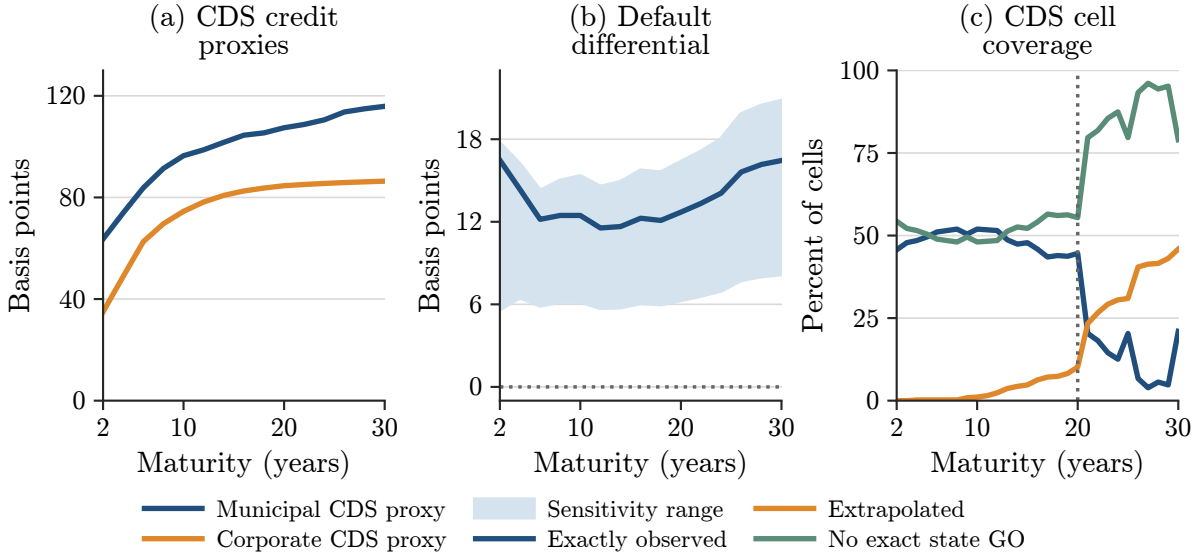


Figure 20: Default adjustment on the A-or-above CDS panel. (a) Municipal and corporate CDS proxies in taxable-equivalent units. (b) Their differential after the dated household tax mapping in (140), with the five-variant sensitivity range. (c) Exactly observed, extrapolated and no-exact-state-GO shares of rating-matched CDS cells. The last two overlap, so the three shares are not a partition. The vertical rule marks twenty years.

Table 24: Default sensitivity in basis points. Every column uses the dated household tax mapping of (140). The 2007+ column starts the sample in 2007, the basis columns scale both CDS proxies by the stated factor, and the quality-capped column keeps cells with at least half of their source observations quoted or interpolated. The band in Table 23 spans these five columns.

T	2007+	basis 0.50	basis 0.75	baseline	quality-capped
2	17.64	8.24	12.36	16.48	5.69
6	14.20	6.09	9.13	12.18	5.94
10	15.27	6.23	9.35	12.46	8.35
14	14.85	5.83	8.74	11.65	8.02
18	15.54	6.05	9.07	12.10	8.76
22	17.06	6.67	10.00	13.34	11.24
26	19.77	7.81	11.71	15.62	17.47
30	20.75	8.23	12.34	16.45	18.66

C.1.7 The combined effect

Table 25 reports the accounting at selected maturities, and Table 26 gives every maturity with each adjustment's band.

The call and issuance-premium adjustments grow with the spread. The three bond-side adjustments together remain near 17–20 bp from twelve years onward. Capital gains peak at intermediate maturities because the resale-tax wedge is amortized over the remaining life and the de minimis threshold rises with it. At thirty years, the residual before the model is 98.7 bp, and the normalized model leaves an unexplained residual of 1.3 bp (Figure 8). Negative residuals before the model at the shortest maturities refer to the full curve, and an adjustment's own matched sample can give a different sign.

Table 25: What explains the municipal spread over the household benchmark: contributions in basis points to $\text{Spread}(T)$ of equation (112). The two comparability columns remove the call option and the issuance premium from the municipal yield itself, giving the puzzle of Figure 2. The three bond-side differentials are municipal-minus-corporate and do not map to a single adjusted curve.

T	comparability			bond-side differentials				Total differential	Share of puzzle	Residual
	Spread	Call	Issuance premium	Puzzle	Capital gains	Liquidity	Default			
2	9.9	0.1	2.9	6.9	0.0	-0.4	16.5	16.0		-9.1
6	27.9	0.7	3.3	23.9	10.1	-0.4	12.2	21.8	91%	2.1
10	52.2	4.6	4.0	43.6	9.9	-0.3	12.5	22.0	51%	21.6
14	87.0	17.9	7.2	61.9	7.1	-0.1	11.7	18.6	30%	43.3
18	110.3	23.9	9.6	76.9	4.7	-0.2	12.1	16.6	22%	60.2
22	128.8	27.0	11.4	90.5	3.6	2.8	13.3	19.8	22%	70.7
26	145.0	28.5	12.5	104.0	2.9	1.5	15.6	20.0	19%	84.0
30	159.8	29.5	12.8	117.5	2.3	0.1	16.5	18.8	16%	98.7

Notes. The puzzle is the spread less the call and issuance-premium adjustments. The total differential sums the three bond-side adjustments, the share divides that sum by the puzzle (omitted where the puzzle is below 20 basis points), and the residual is the residual before the model, the puzzle less the sum. Thirty-year bands, in basis points: call [22.8, 36.3], issuance premium [8.1, 17.3], capital gains [1.2, 2.5], liquidity [-2.2, 2.3], default [8.2, 20.8]. Using unrounded band endpoints gives a thirty-year residual before the model of [80.8, 121.7] bp, an envelope with no joint coverage level. Appendix C.1 gives the construction. The issuance-premium estimate uses the 4% of curve bonds with an institutional-size trade within one month of issuance.

Table 26: All five adjustments to $\text{Spread}(T)$ of equation (112) by maturity, in bp. Brackets give each adjustment's band. The total band adds the band endpoints and has no joint coverage level. The final column divides the total by $\text{Spread}(T)$ and is suppressed below 20 bp of observed spread. Table 25 divides the three bond-side adjustments by the puzzle.

T	call pt band	issuance premium pt band	capital gains pt band	liquidity pt band	default pt band	total pt band	obs. Spread(T)	% expl.
2	0.1 [0.1, 0.1]	2.9 [1.7, 4.2]	0.0 [0.0, 0.0]	-0.4 [-2.1, 1.2]	16.5 [5.7, 17.6]	19.0 [5.4, 23.1]	9.9	
4	0.2 [0.1, 0.2]	3.0 [1.8, 4.3]	6.9 [4.4, 7.0]	-0.4 [-2.2, 1.3]	14.3 [6.6, 16.1]	24.0 [10.7, 28.9]	16.9	
6	0.7 [0.6, 0.8]	3.3 [2.0, 4.5]	10.1 [5.2, 10.3]	-0.4 [-2.5, 1.7]	12.2 [5.9, 14.2]	25.8 [11.3, 31.4]	27.9	92
8	1.9 [1.6, 2.1]	3.6 [2.3, 4.9]	10.7 [5.2, 11.3]	-0.4 [-2.8, 2.0]	12.5 [6.2, 15.0]	28.2 [12.6, 35.2]	41.1	69
10	4.6 [4.0, 5.1]	4.0 [2.7, 5.3]	9.9 [4.5, 10.5]	-0.3 [-3.1, 2.5]	12.5 [6.2, 15.3]	30.6 [14.4, 38.7]	52.2	59
12	12.6 [11.0, 14.3]	5.6 [4.1, 7.1]	8.3 [3.7, 8.8]	-0.2 [-3.3, 3.0]	11.5 [5.8, 14.5]	37.9 [21.2, 47.6]	71.2	53
14	17.9 [15.3, 20.5]	7.2 [5.5, 8.9]	7.1 [3.1, 7.3]	-0.1 [-3.3, 3.0]	11.7 [5.8, 14.8]	43.7 [26.3, 54.6]	87.0	50
16	21.3 [17.8, 24.8]	8.4 [6.5, 10.4]	5.6 [2.6, 5.6]	-0.2 [-3.0, 2.6]	12.3 [6.1, 15.7]	47.5 [30.1, 59.1]	99.4	48
18	23.9 [19.7, 28.0]	9.6 [7.5, 11.8]	4.7 [2.2, 4.8]	-0.2 [-2.6, 2.2]	12.1 [6.0, 15.5]	50.1 [32.8, 62.4]	110.3	45
20	25.3 [20.5, 30.0]	10.8 [8.5, 13.2]	4.1 [2.0, 4.4]	1.3 [-2.0, 4.6]	12.7 [6.4, 16.3]	54.2 [35.3, 68.5]	123.1	44
22	27.0 [21.6, 32.4]	11.4 [8.4, 14.4]	3.6 [1.8, 3.9]	2.8 [-1.4, 7.1]	13.3 [6.7, 17.1]	58.1 [37.1, 74.8]	128.8	45
24	28.0 [22.2, 33.8]	12.0 [8.3, 15.5]	3.2 [1.6, 3.5]	2.9 [-1.3, 7.2]	14.1 [7.0, 17.9]	60.1 [37.8, 78.0]	136.9	44
26	28.5 [22.4, 34.7]	12.5 [8.2, 16.7]	2.9 [1.5, 3.1]	1.5 [-1.7, 4.6]	15.6 [7.8, 19.8]	61.0 [38.1, 78.9]	145.0	42
28	29.0 [22.6, 35.6]	12.8 [8.1, 17.3]	2.6 [1.3, 2.8]	0.1 [-2.2, 2.3]	16.2 [8.1, 20.4]	60.7 [37.9, 78.3]	145.6	42
30	29.5 [22.8, 36.3]	12.8 [8.1, 17.3]	2.3 [1.2, 2.5]	0.1 [-2.2, 2.3]	16.5 [8.2, 20.8]	61.1 [38.2, 79.1]	159.8	38

Model and observed curves. The model comparison in Section 5.4 preserves the model’s municipal-to-taxable yield ratio and applies it to the observed corporate yield. Subtracting the observed household benchmark then puts the model gap on the same basis as the empirical adjustments. With bars denoting sample averages and superscripts m and d distinguishing model and data,

$$\tilde{y}_T^M = \frac{\bar{y}_T^{M,m}}{\bar{y}_T^{C,m}} \bar{y}_T^{C,d}, \quad \tilde{\mathcal{G}}_T = \tilde{y}_T^M - B_T^d.$$

Here $B_T^d = \bar{\xi}_H(T) \bar{y}_T^{C,d}$ is the observed household benchmark of (112). All yields use nominal semiannual units, and dated after-tax products are formed before the averaging of Appendix C.1.1. This preserves the observed benchmark rather than rescaling the model’s own averaged benchmark. Figure 8 and Table 4 report this normalized comparison. For the same maturities as the table, 2, 6, 10, 14, 18, 22, 26 and 30 years, the unnormalized gaps using the model’s own taxable curve are $-0.2, 14.3, 36.2, 62.1, 81.1, 92.7, 101.1$ and 107.4 basis points, respectively. Section 5.4 explains the distinction.

The normalized root-mean-square unexplained residual is 11.10 bp (16.83 bp unnormalized), and the maximum absolute residual is 14.77 bp at sixteen years. Five of fifteen absolute residuals are below 10 bp. The signed residuals are -12.00 bp at twelve years, -11.08 at twenty and $+1.32$ at thirty.

C.2 Issuer-level liquidity and offering-yield regressions

These regressions ask whether issuers whose bonds trade more often or at lower cost have lower offering yields. They complement the matched municipal–corporate trading-cost comparison of Appendix C.1.5, which supplies the liquidity adjustment in Section 3.7.

Liquidity measures. Table 27 associates offering yields with full-life issuer liquidity proxies. Because the proxies are not predetermined at issuance, the coefficients are neither causal ex-ante premia nor formal bounds. The trading-frequency measure is the issuer mean of each bond’s average daily trades reported to the MSRB, with days counted from the bond’s dated date to maturity. The spread measure is the corresponding issuer mean of each bond’s average imputed round-trip cost (Appendix C.1.5), in bond-price units (0.01 is 1% of price, or 100 basis points).

Translation to long-end magnitudes. At 26–30 years, the indicator coefficients in columns (1), (2) and (4) are about 8–15 bp in absolute value, and the continuous column (3) coefficient of 27.262 implies 27.3 bp for a 0.01 spread contrast. At these maturities, more liquid issuers have lower offering yields on all four measures.

C.3 Paired issues: Build America Bonds (BABs) and tax-exempt bonds

This comparison uses taxable and tax-exempt bonds from the same municipal issuer to examine the maturity pattern without an external taxable benchmark. Multiplying taxable Build America Bond (BAB) offering yields by 0.65, one minus the period’s top federal household rate, puts them on the household after-tax scale used in Section 3.8.

Sample and construction. The issue-level Mergent sample, one observation per bond identifier (CUSIP), includes bonds issued in 2009Q2–2010Q4 and not subject to the alternative minimum tax (AMT). Table 28 uses AA-and-above issues (agency score ≤ 3) and A-or-above issues (≤ 7). Figure 3 uses the former and averages yields within cells of integer maturity $n \in \{1, \dots, 30\}$ and bond type.

Regression specification. In the notation of equation (4), with n_b the remaining maturity at issuance, I estimate

$$y_b = \alpha_1 EXMP_b + \beta (EXMP_b \times n_b) + \gamma (0.65 y_{n_b, t_b}^C) + \lambda_{j(b) \times n_b \times q_b} + \varepsilon_b. \quad (141)$$

Table 27: Issuer liquidity and offering yield, by maturity bin.

	Trading frequency		Transaction-cost spread	
	(1) Above corp median	(2) Top 25% trade	(3) Issuer-avg spread (cont.)	(4) Lowest-spread quartile
Maturity 1–5y	−0.046*** (0.009)	−0.028*** (0.007)	13.446*** (0.489)	−0.057*** (0.001)
Maturity 6–10y	−0.007** (0.003)	0.021*** (0.005)	11.624*** (0.216)	−0.048*** (0.002)
Maturity 11–15y	−0.011*** (0.004)	0.041*** (0.001)	11.860*** (0.355)	−0.040*** (0.001)
Maturity 16–20y	−0.035*** (0.003)	0.024*** (0.006)	13.447*** (0.557)	−0.043*** (0.003)
Maturity 21–25y	−0.091*** (0.009)	−0.041** (0.016)	20.427*** (1.225)	−0.104*** (0.011)
Maturity 26–30y	−0.093*** (0.030)	−0.082*** (0.018)	27.262*** (4.815)	−0.146*** (0.037)
Fixed effects	Yes	Yes	Yes	Yes
Maturity cluster	Yes	Yes	Yes	Yes
Observations	785,256	785,256	779,988	785,256

Notes. AA-and-above federally exempt issues, 2000–2021. The outcome is the offering yield in percent, so −0.046 is −4.6 bp. Each row reports the coefficient on an issuer-average liquidity measure in that maturity bin: (1) an indicator that daily MSRB trades per bond exceed the TRACE median for plain-vanilla corporate debentures and medium-term notes; (2) an indicator for the top quartile of municipal trade frequency; (3) the spread, the imputed round-trip cost in price units, so 13.4 is 13.4 bp per 0.01 of spread; (4) an indicator for the lowest-spread quartile. Fixed effects interact rounded maturity, rating rounded down, state, issue year–month and call status. Standard errors are clustered by rounded maturity. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The corporate control $0.65 y_{n_b, t_b}^C$ varies across issue months within matched quarters, identifying γ . Equation (4) uses the unscaled control, with $\gamma_C = 0.65\gamma$, which changes only the control’s coefficient and standard error.

Results. Table 28 shows a positive maturity gradient under both rating cuts and call matching, with all maturity interactions significant at one percent and Panel B bin coefficients rising with maturity. Call matching lowers the longest-bin coefficient by about fifteen percent in the AA-and-above sample but leaves the [5, 10) coefficient unchanged.

Call features. Call flags occur on 49 percent of BABs, versus 39 percent of tax-exempt bonds. Call-matched columns match these flags only, so exercise terms, coupons, liquidity and the treatment of original-issue discount (OID) or premium can still differ.

Original-issue premium and discount. Exact after-tax internal rates of return would treat BAB and exempt OID, acquisition premium, interest income and basis separately, so I check the 0.65 yield approximation on bonds issued near par. Restricting the call-matched AA-and-above sample to offering prices of 99–101 leaves a maturity gradient of 2.1 bp per year (s.e. 0.6), against 3.1 in the full call-matched sample. The 98–102 restriction gives 2.7 bp (s.e. 0.4). Both remain significant at one percent.

Interpretation. The paired comparison of Section 3.8 holds issuer and credit fixed, with the BAB yield replacing the corporate benchmark. The widening gap supports the participation mechanism under the 0.65 approximation and comparable security design and issuance premia. The 35% top federal rate behind the

scaling is below the 41.1% average household wedge of the AA-and-above curve (Appendix Figure 28), and a lower after-tax share would widen the gap. The BAB gaps are closest to that curve's gaps at the shortest and longest maturities and farthest at [15, 20] years.

Table 28: Paired issues: the tax-exempt-versus-BAB gap widens with maturity.

	AA and above (score ≤ 3)		A or above (score ≤ 7)	
	(1) base	(2) call-matched	(3) base	(4) call-matched
<i>Panel A. Continuous maturity</i>				
Tax-exempt	−0.102*** (0.034)	−0.091** (0.036)	−0.062** (0.027)	−0.043 (0.030)
Tax-exempt $\times n_b$	0.034*** (0.003)	0.031*** (0.003)	0.033*** (0.002)	0.030*** (0.003)
After-tax corporate yield	0.379*** (0.138)	0.376*** (0.135)	0.264* (0.134)	0.261* (0.130)
<i>Panel B. Five-year maturity bins</i>				
Tax-exempt	−0.048 (0.033)	−0.050 (0.033)	−0.007 (0.024)	−0.008 (0.024)
Tax-exempt $\times [5, 10)$ y	0.192*** (0.022)	0.191*** (0.022)	0.192*** (0.017)	0.190*** (0.017)
Tax-exempt $\times [10, 15)$ y	0.396*** (0.033)	0.364*** (0.033)	0.368*** (0.026)	0.342*** (0.030)
Tax-exempt $\times [15, 20)$ y	0.450*** (0.041)	0.390*** (0.039)	0.419*** (0.029)	0.403*** (0.039)
Tax-exempt $\times [20, 25)$ y	0.576*** (0.061)	0.540*** (0.066)	0.593*** (0.071)	0.510*** (0.066)
Tax-exempt $\times [25, 30]$ y	0.755*** (0.079)	0.639*** (0.039)	0.996*** (0.074)	0.804*** (0.060)
After-tax corporate yield	0.380*** (0.138)	0.378*** (0.135)	0.264* (0.134)	0.262** (0.130)
Issuer $\times$ maturity $\times$ quarter FE	Yes	Yes	Yes	Yes
$\times$ call status	No	Yes	No	Yes
State cluster	Yes	Yes	Yes	Yes
Observations	29,839	28,890	61,816	59,886

Notes. Offering yields in percent, 2009Q2–2010Q4, excluding AMT bonds. BAB yields are multiplied by 0.65. The corporate-yield row reports γ , the coefficient on the after-tax control in equation (141). The omitted bin is [1, 5] years. Standard errors are clustered by issuing jurisdiction (51 in columns (1)–(2), 52 in columns (3)–(4)). Call-matched columns add the optional-call flag to the absorbed issuer-by-maturity-by-quarter cell. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4 The municipal puzzle on unlimited-tax general-obligation bonds

Section 3.6 states that the declining implied tax also appears among unlimited-tax general-obligation (GO) bonds.

Sample. Unlimited-tax GO bonds pledge the issuer's full faith and credit and unlimited taxing power, while revenue bonds pledge project revenues. Figure 21 restricts the comparison to AA-and-above unlimited-

tax GO bonds²⁶ passing the main call-pricing and calendar screens (Appendix C.1.2). No default adjustment is applied, because local obligations' credit differs from the state-government CDS proxy (Appendix C.1.6).

Result. After call and issuance-premium adjustments, GO implied tax falls from 43.1% at two years to 26.0% at thirty, matching the 17.1-point decline of the A-or-above full sample and exceeding the 14.6-point decline of the AA-and-above sample. The screens exclude 4,029 GO bonds (0.86%). Relative to the unscreened GO call-adjusted series, with the issuance premium held fixed, they change the decline by 0.49 points, almost entirely at the short end. The result agrees with earlier evidence that the long-end residual persists after default and liquidity adjustments and under high-grade or GO restrictions (Longstaff, 2011; Fleckenstein and Longstaff, 2025; Chalmers, 1998).

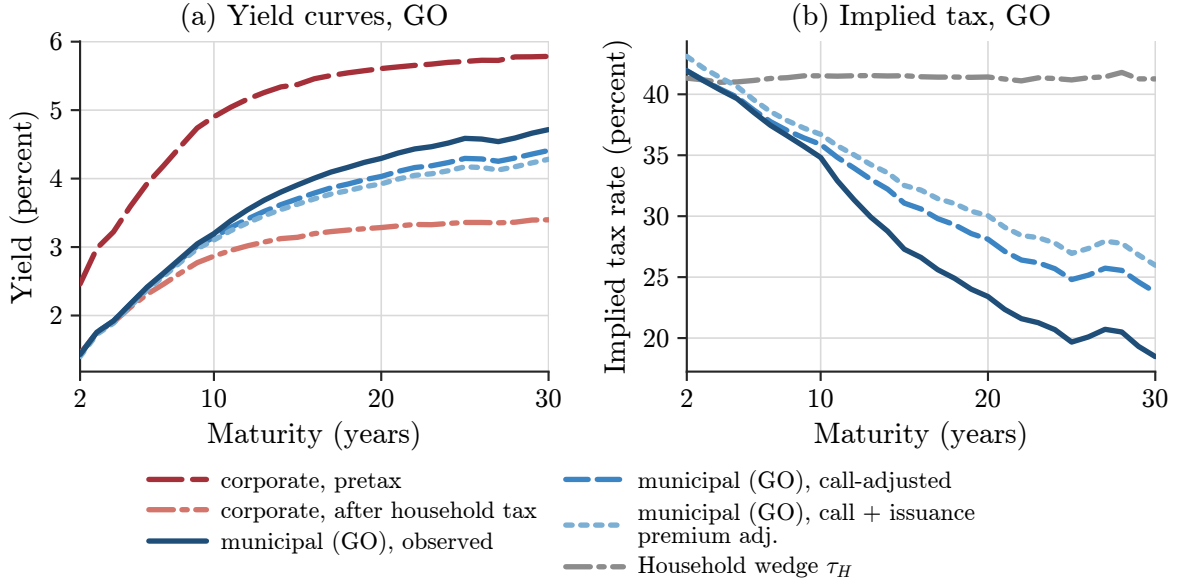


Figure 21: AA-and-above unlimited-tax GO bonds: 465,258 bonds pass the main call-pricing and calendar screens, and curves use the sample's own weights and household benchmark. Calls use the calibrated model. Because GO offering-week matches are scarce, the issuance-premium adjustment uses the all-bond institutional estimate (Appendix C.1.3). Adjusted implied tax falls 17.1 points from two to thirty years, matching the A-or-above full sample.

C.5 The single-tax-rate benchmark of Green (1993)

Sections 2 and 3.7 refer to the single-tax-rate explanation of Green (1993).

Green's mechanism. Green explains rising ratios of municipal to taxable yields through taxable-bond tax timing. Buying two five-percent market-discount bonds and shorting one ten-percent par bond of equal maturity produces no net cash flow before maturity and one dollar of principal then, deferring tax. At rate τ , the package costs P_T and pays $1 - \tau(1 - P_T)$ after tax at maturity. With d_T the price of a dollar paid at T , the package is worth $P_T = (1 - \tau)d_T / (1 - \tau d_T)$. The limiting friction is the asymmetric taxation of long and short positions, because borrowing and short-sale deductions cannot exceed the investment income they support. Green's high-tax investor prices both municipal bonds and the package, and dealers, who face no such limit, link taxable par prices to it, lowering long taxable yields below fully taxed returns. Thus one investor can price municipal bonds at every maturity as the yield ratio rises, without the different price-setting investors across maturities that Fama (1977) requires.

²⁶Mergent security code "K".

Relation to participation. Deep market-discount bonds may be scarce, annual OID accrual removes the deferral and shorts consume deduction capacity. Households forgoing the package can accept the household benchmark ξ_{HYT}^C , below Green’s relation. With sufficient capacity, these higher-benefit households set municipal prices, and investors using the strategy do not. Raising municipal yields to Green’s relation would require shorting municipal bonds, subject to the frictions in Appendix D.6. His relation therefore need not bind. Nor does it account for observed long yields: his own 1987–1990 estimates leave thirty-year municipal yields about twenty taxable-equivalent basis points above his relation, a gap that increases with maturity.

Green’s critique of clientele explanations. Green rejects earlier clientele explanations because term premia persisted after mid-1980s reforms disrupted bank holdings. Here participation depends on exemption benefits and the duration benefit (Section 4.3), so the clientele that prices each maturity can change. Households alone hold short maturities and set their yields. Institutions with smaller exemption benefits also hold long maturities, whose yields reflect their tax treatment too, and implied taxes respond to supply and risk capacity. Section 3.4 documents ownership across benefits. I retain ξ_{HYT}^C , measure tax timing separately in the capital-gains adjustment (Appendix C.1.4) and test household state-tax pass-through in P1 and bank and insurer benefits in P2 (Sections 6.1 and 6.2), distinguishing participation from tax timing alone.

C.6 State-tax pass-through: full regression tables

These tables check whether the declining state-tax pass-through in P1 depends on the yield construction or state credit risk. They retain the specification of Section 6.1 and first compare offering and zero-coupon yields, then add CDS controls on a common sample.

Sample. The main P1 regressions use federally exempt AA-and-above issuance (agency score ≤ 3) from 2000–2017. They exclude zero-maturity bonds and group maturities into $[0, 5)$, $[5, 10)$, $[10, 15)$, $[15, 20)$ and $[20, 30]$ years. The zero-yield columns have 681,838 complete-control observations. Six missing offering yields leave 681,832 in the offering-yield columns.

Call-adjusted outcome. The main outcome uses call-adjusted zero-coupon yields. The P1 call adjustment prices each bond on the fifteenth of its issue quarter’s final month, while fixed effects use actual issue months. The option-adjusted spread is fitted on the calibrated model’s supply-weighted curve for the issue date, with a regular semiannual cash-flow grid. As in the main adjustment, each bond starts at the observed rate state of its issue quarter, and the short-rate tree’s rate loadings reproduce the curve’s local rate sensitivity at every maturity. Observed call dates map to the nearest grid date within 0.27 years, and the call is exercisable on every grid date from the first mapped date, unlike the contractual calendar of the main adjustment (Appendix C.1.2). Callable bonds are excluded when the first call date lies more than 0.27 years before the first semiannual grid date or the call schedule lacks a usable date and price.

Controls and fixed effects. All P1 tables retain equation (25)’s six lagged state controls and the rounded average rating, each interacted with all five maturity bins, plus log offering par and security-type indicators. Stock and issuance controls use all observed federally exempt bonds, regardless of rating. Outstanding par equals the offering amount until the reported outstanding date and the reported outstanding amount afterward. Fixed effects are cells defined by issue year, issue month, rounded maturity and callability. Observations with missing controls are excluded without imputation, and singleton cells are then removed.

Estimation and inference. Unweighted ordinary least squares (OLS) is the main estimator. For the instrumental-variables (IV) comparison, the 1990 rate interacted with bins instruments the current-rate interactions, giving a Kleibergen–Paap (KP) rk Wald F of 10.4 and Sanderson–Windmeijer (SW) conditional statistics of 76.1, 115.9, 104.4, 155.7 and 230.4. Standard errors are state-clustered, with $t(G - 1)$ critical values for G states ($t(46)$ in the main sample). The endpoint test of $\beta_1 = \beta_5$ gives each table’s decline p -value.

C.6.1 Comparison across yield constructions

Table 29 compares Mergent offering yields, which may be yields to worst for premium callable bonds, with zero yields from offering prices, with and without call adjustment. Both OLS and IV reject endpoint equality at one percent for every construction.

Table 29: P1 across yield constructions, 2000–2017.

	Offering yield		Call-adjusted zero		Unadjusted zero	
	OLS	IV	OLS	IV	OLS	IV
β_1 : [0, 5) years	0.985*** (0.238)	1.084*** (0.249)	1.069*** (0.195)	1.088*** (0.225)	1.080*** (0.196)	1.102*** (0.227)
β_2 : [5, 10) years	0.635*** (0.075)	0.648*** (0.093)	0.677*** (0.069)	0.694*** (0.086)	0.683*** (0.069)	0.693*** (0.087)
β_3 : [10, 15) years	0.414*** (0.053)	0.410*** (0.075)	0.484*** (0.096)	0.447*** (0.118)	0.515*** (0.100)	0.479*** (0.117)
β_4 : [15, 20) years	0.268*** (0.060)	0.257*** (0.067)	0.390*** (0.102)	0.326*** (0.116)	0.428*** (0.108)	0.360*** (0.121)
β_5 : [20, 30] years	0.193*** (0.052)	0.152** (0.057)	0.305*** (0.075)	0.229*** (0.081)	0.335*** (0.078)	0.255*** (0.085)
Decline $\beta_1 - \beta_5$	0.791***	0.933***	0.764***	0.859***	0.745***	0.847***
p -value of decline	0.004	< 0.001	0.002	< 0.001	0.003	0.001
Issue month $\times$ maturity $\times$ call FE	Yes	Yes	Yes	Yes	Yes	Yes
State cluster	Yes	Yes	Yes	Yes	Yes	Yes
Observations	681,832	681,832	681,838	681,838	681,838	681,838
First-stage F (Kleibergen–Paap)	—	10.4	—	10.4	—	10.4

Notes. Outcomes are implied state taxes (24) from Mergent offering yields or from zero-coupon yields with and without the call adjustment. The call-adjusted zero is the main outcome. Controls, fixed effects and state-clustered inference (47 states) follow Appendix C.6, 2000–2017. IV instruments each bin’s current rate with its 1990 rate. The decline row reports $\beta_1 - \beta_5$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.6.2 State credit-risk controls

Table 30 restricts the pre-2018 sample to observations with a preceding-month five-year state-government CDS quote and adds that quote interacted with the maturity bins. On this common sample of 345,877 bonds from 25 states, the CDS controls raise the longest-bin OLS estimate from 0.335 to 0.416 and IV from 0.322 to 0.368. Both estimators retain a declining profile and reject endpoint equality at five percent in every row.

Table 30: P1 sample and credit controls, 2000–2017.

Sample	Obs.	States	β_1 , OLS	β_5 , OLS	Decline p	β_1 , IV	β_5 , IV	Decline p
Main sample	681,838	47	1.069***	0.305*** (0.075)	0.002	1.088***	0.229*** (0.081)	< 0.001
Observed CDS sample	345,877	25	1.210***	0.335*** (0.081)	0.009	1.349***	0.322*** (0.098)	0.009
CDS sample + CDS $\times$ bin	345,877	25	1.400***	0.416*** (0.068)	0.011	1.391***	0.368*** (0.087)	0.025

Notes. Controls, fixed effects and state-clustered inference follow Appendix C.6, 2000–2017. The observed-CDS sample keeps bonds with a prior-month five-year state-government CDS quote, without imputation, and the last row adds that quote interacted with the maturity bins. IV uses bin-interacted 1990 rates. Parentheses give standard errors of β_5 , and decline p -values test $\beta_1 = \beta_5$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.7 State-tax pass-through: instrument vintages, weighting and further checks

These checks ask whether P1 depends on the historical tax-rate instrument, the weight assigned to each state or any single state. They retain the controls, fixed effects and state-clustered inference of Appendix C.6, then examine state fixed effects and the post-2018 tax regime separately.

Instruments. The IV comparisons use the state’s 1990 top income-tax rate as the baseline instrument. The maintained exclusion restriction is that the historical rate affects current implied taxes only through the current rate, conditional on the controls and fixed effects. Tables 31 and 32 compare 1945–2000 vintages on the same pre-2018 sample.

Continuous-maturity specification. Table 31 replaces the five current-rate bin interactions in (25) with the current rate and its interaction with maturity in years, n_b . The interaction coefficient is the pass-through gradient per year, and IV uses the historical rate and its maturity interaction. All instrument vintages give negative IV gradients, unweighted and with equal total weight per state, as does OLS.

Table 31: Continuous-maturity P1 checks, 2000–2017.

Vintage	Corr.	Gradient	Unweighted			Equal per state	
			s.e.	KP F	SW F	Gradient	KP F
OLS	—	−0.0392***	(0.0131)	—	—	−0.0326**	—
1945	0.40	−0.0535**	(0.0251)	8.0	19.2	−0.0440	2.8
1950	0.31	−0.0607**	(0.0273)	7.8	18.4	−0.0878	1.8
1960	0.38	−0.0620**	(0.0254)	9.0	28.8	−0.0997*	2.5
1970	0.51	−0.0574***	(0.0201)	11.5	45.2	−0.0795***	7.3
1980	0.76	−0.0424**	(0.0164)	10.6	103.4	−0.0385**	26.9
1990	0.82	−0.0440***	(0.0127)	30.1	404.1	−0.0307**	25.9
2000	0.88	−0.0399***	(0.0124)	65.1	998.5	−0.0258*	36.5

Notes. Controls, fixed effects and state-clustered inference follow Appendix C.6, 2000–2017. The gradient is the coefficient on the current rate times maturity n_b , the change in pass-through per year. IV instruments the current rate and its maturity interaction with the vintage’s rate and its maturity interaction. KP F is the Kleibergen–Paap statistic and SW F the Sanderson–Windmeijer conditional statistic for the interaction. Corr. is the cross-state correlation between each state’s mean current top rate over the estimation sample and its vintage-year top rate. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Binned profiles. Table 32 compares the five-bin IV profiles with OLS, and Figure 22 plots them. Every vintage rejects endpoint equality at five percent.

Weighting. Every weighting in Table 33 retains a declining OLS profile, with endpoint differences significant at one percent.

State fixed effects. Adding state fixed effects to the pre-2018 OLS specification gives short and long coefficients of 0.665 (s.e. 0.201) and −0.112 (0.212), with a decline of 0.777 (0.220, $p = 0.001$), on 681,838 bonds. State effects absorb fixed tax levels. The five excluded instruments, the 1990 rate interacted with each maturity bin, sum to the state’s 1990 rate, which is collinear with the state effects, so IV is not identified.

Leave-one-state-out. All 47 single-state omissions retain a positive short-to-long decline under both OLS and IV (Figure 23). OLS long-bin estimates span 0.232–0.482 and declines 0.452–0.889. IV long-bin estimates span 0.116–0.315 and declines 0.669–1.004. The largest endpoint-test p -values are 0.062 for OLS and 0.036 for IV.

Table 32: P1 instrument-vintage checks, 2000–2017.

Estimator/vintage	β_1	β_2	β_3	β_4	β_5	Decline p	KP F
OLS (main)	1.069*** (0.195)	0.677*** (0.069)	0.484*** (0.096)	0.390*** (0.102)	0.305*** (0.075)	0.002	—
1945	1.257*** (0.421)	0.868*** (0.158)	0.551*** (0.158)	0.363** (0.143)	0.245** (0.103)	0.027	3.1
1950	1.453*** (0.473)	0.942*** (0.181)	0.622*** (0.153)	0.428*** (0.143)	0.313*** (0.116)	0.024	3.1
1960	1.483*** (0.439)	0.942*** (0.167)	0.627*** (0.147)	0.432*** (0.140)	0.332*** (0.113)	0.014	3.6
1970	1.394*** (0.350)	0.858*** (0.125)	0.566*** (0.143)	0.400*** (0.141)	0.325*** (0.114)	0.006	4.4
1980	1.084*** (0.273)	0.756*** (0.083)	0.528*** (0.121)	0.374*** (0.120)	0.280*** (0.084)	0.010	3.8
1990	1.088*** (0.225)	0.694*** (0.086)	0.447*** (0.118)	0.326*** (0.116)	0.229*** (0.081)	< 0.001	10.4
2000	1.072*** (0.215)	0.678*** (0.079)	0.480*** (0.095)	0.378*** (0.096)	0.280*** (0.069)	0.001	23.5

Notes. Controls, fixed effects and state-clustered inference follow Appendix C.6, 2000–2017. IV instruments each bin’s current rate with the vintage’s top rate in that bin. KP F is the Kleibergen–Paap statistic, and decline p -values test $\beta_1 = \beta_5$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

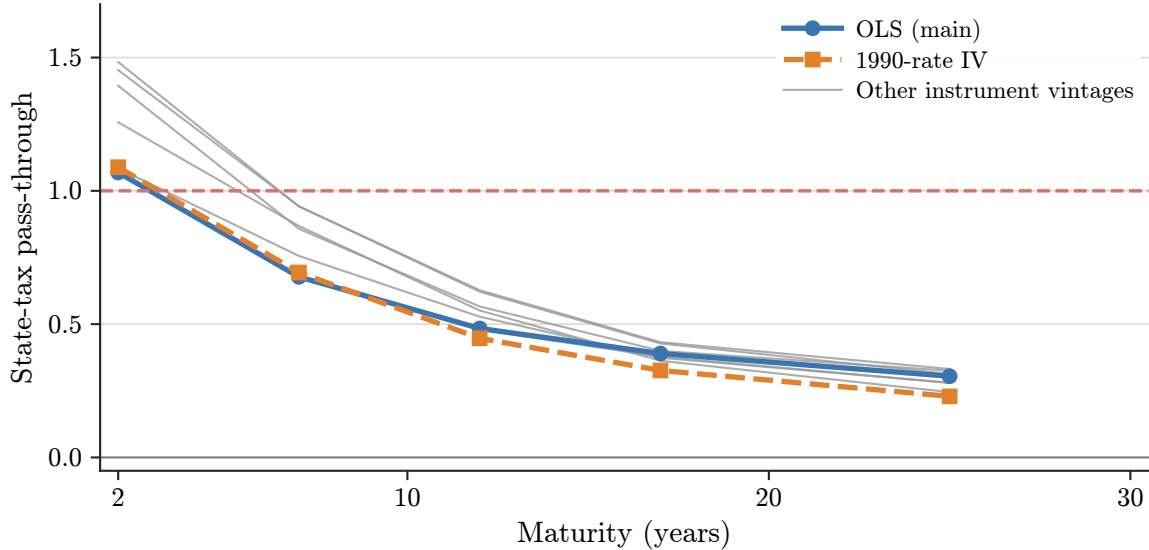


Figure 22: State-tax pass-through using historical tax-rate instruments, 2000–2017. Blue shows the main OLS profile. Orange shows IV using the 1990 rate. Gray lines use the 1945, 1950, 1960, 1970, 1980 and 2000 rates. The dashed red line is the representative-tax benchmark. Table 32 reports state-clustered uncertainty.

Table 33: P1 OLS estimates under four weighting schemes, 2000–2017.

Weighting	β_1	β_2	β_3	β_4	β_5	Decline p
Unweighted (main)	1.069*** (0.195)	0.677*** (0.069)	0.484*** (0.096)	0.390*** (0.102)	0.305*** (0.075)	0.002
Equal per state	0.863*** (0.191)	0.543*** (0.082)	0.374*** (0.101)	0.308*** (0.111)	0.234** (0.088)	0.006
Equal per state-year	0.911*** (0.190)	0.509*** (0.069)	0.317*** (0.080)	0.261*** (0.090)	0.211*** (0.074)	0.002
$1/\sqrt{n_{\text{state}}}$	0.990*** (0.202)	0.617*** (0.080)	0.420*** (0.109)	0.344*** (0.121)	0.282*** (0.091)	0.006

Notes. OLS with the controls, fixed effects and state-clustered inference of Appendix C.6, 2000–2017. The equal-weight rows give each state or state-year the same total weight, and n_{state} is the bond’s state observation count, all measured after the complete-case and singleton restrictions. Decline p -values test $\beta_1 = \beta_5$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

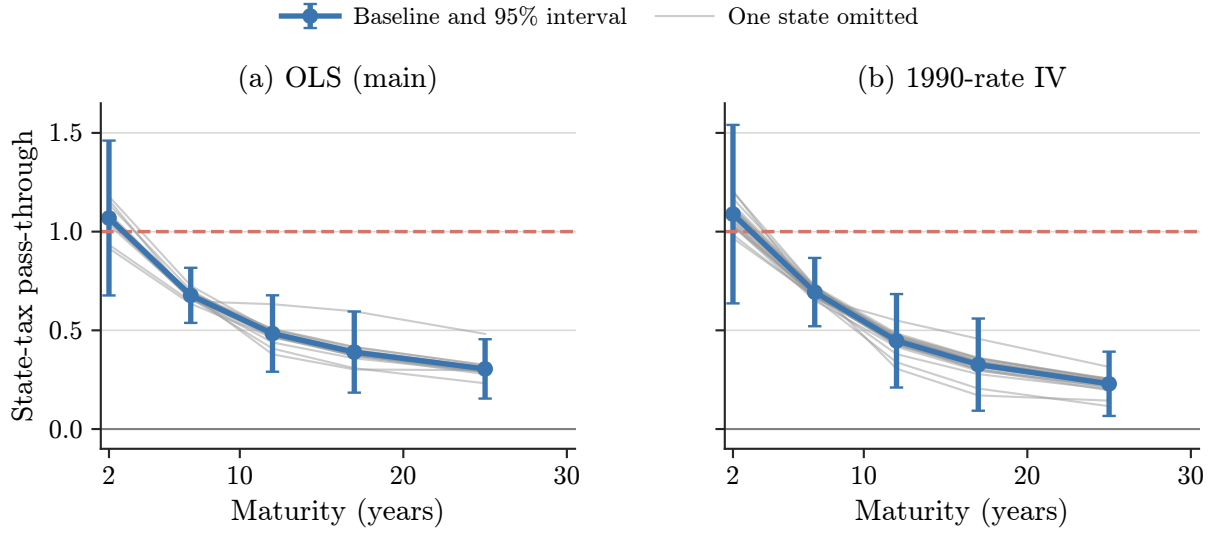


Figure 23: Leave-one-state-out P1 profiles for 2000–2017. Gray lines show each of the 47 state omissions, and blue shows the baseline estimates with state-clustered 95% intervals. Panel (a) uses OLS and panel (b) uses the 1990-rate instruments. The dashed line is the representative-tax benchmark.

The two tax regimes. The main P1 sample ends in 2017, so its product of federal and state after-tax shares matches the six pre-2018 model economies in Figure 9. From 2018, the Tax Cuts and Jobs Act capped state and local tax (SALT) deductions at \$10,000 per return. When the cap binds, an additional dollar of state tax is no longer federally deductible, and the combined after-tax share on taxable interest is $\xi^{\text{fed}} - \tau_s^{\text{st}}$ (notation of Section 6.1). The implied state tax for this regime is

$$\text{ImpliedStateTax}_{s,t}^{\text{cap}}(T) \equiv \xi^{\text{fed}} - \frac{y_{s,t}^M(T)}{y_t^C(T)}. \quad (142)$$

It recovers the statutory state rate under representative-tax indifference for a household with a binding cap.²⁷ Table 34 estimates each regime separately. Both estimators reject endpoint equality at one percent in each period.

²⁷See the IRS, [2017 Schedule A instructions](#), and its [summary of the 2018 changes](#).

Table 34: P1 under the two federal/state tax regimes.

	2000–2017 (main)		2018–2020	
	OLS	IV	OLS	IV
β_1 : [0, 5) years	1.069*** (0.195)	1.088*** (0.225)	0.948*** (0.182)	0.909*** (0.246)
β_2 : [5, 10) years	0.677*** (0.069)	0.694*** (0.086)	0.834*** (0.121)	0.879*** (0.143)
β_3 : [10, 15) years	0.484*** (0.096)	0.447*** (0.118)	0.446*** (0.061)	0.466*** (0.074)
β_4 : [15, 20) years	0.390*** (0.102)	0.326*** (0.116)	0.312*** (0.078)	0.323*** (0.105)
β_5 : [20, 30] years	0.305*** (0.075)	0.229*** (0.081)	0.088 (0.091)	0.082 (0.109)
Decline $\beta_1 - \beta_5$	0.764***	0.859***	0.860***	0.827***
p -value of decline	0.002	< 0.001	< 0.001	0.008
Issue month $\times$ maturity $\times$ call FE	Yes	Yes	Yes	Yes
State cluster	Yes	Yes	Yes	Yes
Observations	681,838	681,838	146,174	146,174
First-stage F (Kleibergen–Paap)	—	10.4	—	7.9

Notes. Controls, fixed effects and state-clustered inference (47 states) follow Appendix C.6. IV uses bin-interacted 1990 rates. The outcome is (24) before 2018 and, for a household whose SALT deduction cap binds, (142) from 2018. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.8 The bank-qualified discontinuity: construction, robustness and placebos

For the P2 test of Section 6.2, Table 35 reports the full five-bin two-stage least squares (2SLS) estimates of (27), and Figure 24 extends Figure 10 to all five maturity bins.

C.8.1 Benchmarks and the designation premium

The two benchmarks. Both benchmarks of (26) compare annual after-tax carry per dollar of basis on matched bonds, in the notation of Section 6.2. I suppress dates, and $y_k^C > 0$ is the taxable yield at maturity k .

The representative-tax benchmark has $y_k^N = (1 - \tau_H)y_k^C$, where τ_H is the top household rate. At this yield a designated bond earns the bank $y_k^N - 0.2\tau_{BC}$ after the tax cost of its disallowed interest deduction, against $(1 - \tau_B)y_k^C$ on a taxable bond, giving the taxable bond an advantage of

$$\underbrace{(1 - \tau_B)y_k^C}_{\text{taxable alternative}} - \underbrace{[(1 - \tau_H)y_k^C - 0.2\tau_{BC}]}_{\text{qualified bond at the household yield}} = (\tau_H - \tau_B)y_k^C + 0.2\tau_{BC} > 0. \quad (143)$$

It is positive whenever $\tau_H > \tau_B - 0.2\tau_{BC}/y_k^C$, in particular when household rates are at least banks' and funding costs are positive. Banks still prefer taxable bonds and designation leaves household taxes unchanged, so $y_k^Q = y_k^N$ and $\varpi_k = 0$ at every maturity.

In the bank-only benchmark, both municipal bonds satisfy the bank's after-tax indifference conditions with taxable bonds:

$$\begin{aligned} y_k^N - \underbrace{\tau_{BC}}_{\text{full disallowance}} &= (1 - \tau_B)y_k^C, \\ y_k^Q - \underbrace{0.2\tau_{BC}}_{\text{qualified disallowance}} &= (1 - \tau_B)y_k^C. \end{aligned} \quad (144)$$

Table 35: Bank-qualified designation and yield premia by maturity.

Dependent variable	(1) BQ designation	(2) Offering yield (bp)	(3) Yield / expected benefit
Estimator	OLS	2SLS	2SLS
Regressor, interacted with each maturity bin	Below-cap indicator	BQ designation	BQ designation
Reported effect	Designation increase	Yield reduction	Scaled yield reduction
1–2 years	0.42*** (0.015)	1.8 (2.2)	0.17 (0.11)
2–5 years	0.42*** (0.015)	8.7*** (2.2)	0.31*** (0.08)
5–10 years	0.44*** (0.015)	13.8*** (2.2)	0.41*** (0.06)
10–20 years	0.45*** (0.017)	18.4*** (2.3)	0.52*** (0.07)
Over 20 years	0.43*** (0.027)	18.2*** (6.3)	0.47*** (0.19)
10–20 minus 1–2 years		16.5*** (2.7)	0.34*** (0.11)
Fixed effects	Yes	Yes	Yes
Issuer cluster	Yes	Yes	Yes
Observations	355,372	355,372	355,372

Notes. Each column is a separate pooled regression containing all five maturity interactions. Column (1) reports the designation increase associated with below-cap status, estimated by ordinary least squares (OLS). Columns (2)–(3) instrument designation by below-cap status within each bin, as in (27), and report $-\beta_g$, so positive numbers mean lower yields on designated bonds. Column (3) divides each bond’s yield by its expected annual benefit $\bar{X}_b = 0.8\tau_{B,t}\hat{c}_b$ before estimation (Appendix C.8.3). Contrasts use unrounded estimates. Parentheses give standard errors clustered by issuer (15,889 issuers). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Fixed effects in every column. Maturity bucket $\times$ year, state $\times$ year, state $\times$ offering year–month, rating category $\times$ year, security type $\times$ year, use of proceeds, coupon bin $\times$ maturity bucket and callability $\times$ maturity bucket $\times$ year. Other controls are log tranche par, maturity-bin indicators and separate linear issuance slopes on each side of the cap within each bin.

Sample. Tax-exempt new issues in 2000–2021 within \$0.5–\$4 million of the statutory cap (\$30 million in 2009–10, \$10 million otherwise). Maturity is measured at issuance. Appendix C.8.2 defines the fixed-effect buckets.

Subtracting cancels the taxable yields within each maturity pair and gives $\varpi_k = 0.8\tau_B c = X$, so constant funding costs imply a premium flat in maturity in basis points. The expected annual benefit $\bar{X}_b$ (Appendix C.8.3) evaluates this benefit at expected annual funding costs.

Derivation of the designation premium. The derivation compares matched bonds of maturity k in supply $q_k > 0$, with outside opportunities and non-tax characteristics fixed. The designation benefit u is zero for a non-designated bond and X for a designated one. At yield y , household demand is $d_k^H(y)$ and bank demand $d_k^B(y + u)$, both reflecting baseline funding costs, taxes and risk compensation. For each $u \in [0, X]$, assume that a clearing yield exists, that demands are continuous and nondecreasing with finitely many continuously differentiable regimes and that total demand has a strictly positive slope at clearing. Then

$$d_k^H(y_k(u)) + d_k^B(y_k(u) + u) = q_k, \quad y_k^N = y_k(0), \quad y_k^Q = y_k(X). \quad (145)$$

Differentiating (145) on a smooth segment gives

$$-\frac{dy_k(u)}{du} = \frac{\partial d_k^B / \partial y}{\partial d_k^H / \partial y + \partial d_k^B / \partial y} \equiv w_{Bk}(u) \in [0, 1], \quad (146)$$

with demand slopes at the clearing net returns, so $w_{Bk}(u)$ is the banks' share of the total demand response. A larger benefit raises bank demand, and the yield falls to offset it. Integrating continuously across participation boundaries gives

$$\varpi_k = y_k(0) - y_k(X) = \int_0^X w_{Bk}(u) du = \theta_k X, \quad \theta_k = \frac{1}{X} \int_0^X w_{Bk}(u) du. \quad (147)$$

If banks never participate, $\partial d_k^B / \partial y = 0$ and $\varpi_k = 0$, the representative-tax benchmark. If only banks hold the bond, $\partial d_k^H / \partial y = 0$ and $\varpi_k = X$, the bank-only benchmark. When both respond, capitalization lies between them. The capitalization share θ_k , the yield reduction as a fraction of the benefit, is distinct from banks' share of holdings.

The flat- X prediction concerns annual yield-equivalent carry and makes no claim about a fixed price premium. The benefit applies per dollar of tax basis, so mapping it into offering yields uses comparable cash flows and the ratio of basis to price.

C.8.2 Design

Running variable. The running variable R_{jt} of Section 6.2 measures issuance as all tax-exempt par within each six-digit CUSIP issuer prefix and offering calendar year, before sample restrictions. Statutory issuance aggregates subordinate entities, includes bank loans and leases (mostly absent from Mergent) and excludes refundings of designated bonds (deemed designated under Section 265(b)(3)(D)) and private-activity bonds other than 501(c)(3) bonds. This measurement noise motivates the fuzzy design.

Among bonds \$0.5–4 million above the cap, 21% are designated. In that window, 4% of single-sale issuer-years designate, and issuer-years with multiple sales designate 44% of their issues. The statute tests reasonably anticipated issuance at each sale, so a March designation stands even when an unanticipated October sale takes the year over the cap. Refundings are 46% of the bonds in that window, and measuring issuance net of refundings reduces designation in the corresponding window to 9%. Designation falls from 24% at \$0.5–1 million above the cap to 15–16% at \$4–8 million and 5% beyond \$10 million, consistent with measurement noise. Designation, refunding status, security type, bond insurance and credit enhancement come from the Mergent bond master files.

Controls. In (27), W_b contains a linear function of R_{jt} on each side of the cap, separately by maturity bin, the log par amount of the maturity tranche (floored at \$1,000) and the fixed effects listed under Table 35. The seven maturity buckets, used only in these fixed effects and distinct from the five maturity bins g of the coefficients, have upper limits of 2, 5, 10, 15, 20 and 30 years, with a seventh bucket beyond thirty.

The window. The baseline window keeps issuer-years with $0.5 \leq |R_{jt}| \leq 4$, and its half-million donut excludes managed bunching (2,081 issuer-years within \$0.25 million below the cap versus 572 above) and 1,034 issuer-years with 12,076 bonds exactly at the \$10 million cap (\$30 million in 2009–10). At-cap issuers are eligible (Section 6.2) and designate 79.6% of their bonds, the below-cap rate, against 27% at \$1–50 thousand above the cap. The below-cap indicator is always $E_{jt} = \mathbf{1}\{R_{jt} \leq 0\}$; a strict inequality would misclassify at-cap issuers, lowering the no-donut first stage from 0.39 to 0.30 but barely changing 2SLS, because dividing by the first stage offsets the attenuation.

Balance at the cap. Local linear covariate regressions in the baseline window, with maturity-bucket-by-year and state-by-year effects, find callability and the number of maturities per issue balanced in every bin. Ratings, observed for 62%, differ only at 1–2 years (0.17 notch, $t = 2.1$). Coupons, set at pricing and covered by coupon-bin-by-maturity-bucket fixed effects, are 0.09–0.14 percentage points lower below the cap

in the four bins beyond two years. Log par, always controlled, is lower below the cap at 1–2 years (-0.26) and higher at 10–20 years ($+0.12$).

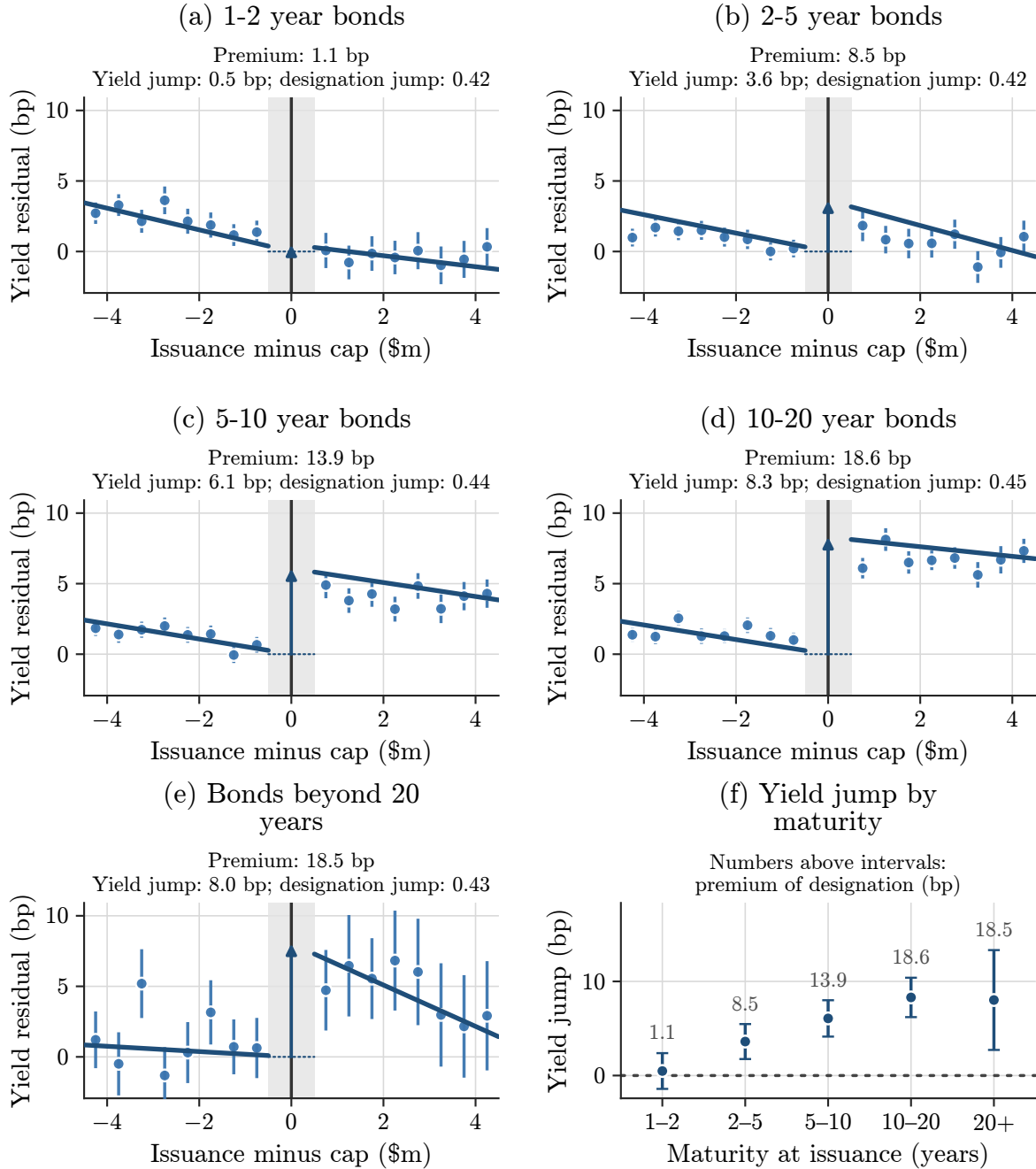


Figure 24: Offering-yield discontinuities by maturity. Panels (a)–(e) show residualized \$0.5 million issuance-bin means relative to the below-cap fitted value at the cap, with local linear fits, donut shading and 95 percent bars. Arrows show above-cap-minus-below-cap jumps, and headers divide each yield jump by the designation jump to obtain the premium. Panel (f) shows jumps and premia, the binwise counterparts to Table 35. The common scale covers the central 96 percent of binned means.

C.8.3 Funding costs and the benefit measures

Funding costs. Section 265(b) allocates a bank’s interest expense across its assets in proportion to their average adjusted bases, without matching each bond to borrowing of the same maturity. With bank n ’s annual interest expense I_{nt} and average asset basis $\bar{A}_{nt}$, a dollar of municipal basis held throughout the year is allocated $c_{nt} = I_{nt}/\bar{A}_{nt}$ of interest expense. Designation reduces the nondeductible fraction from one to 0.2, saving $X_{nt} = 0.8\tau_{B,t}c_{nt}$ at the bank’s federal tax rate $\tau_{B,t}$, 35% before 2018 and 21% thereafter. Here I_{nt} is the year-to-date interest expense in banks’ December reports in the BankFind database of the Federal Deposit Insurance Corporation (FDIC), already an annual total, and $\bar{A}_{nt}$ is average assets. This expense-to-assets ratio is neither a deposit rate nor a maturity-matched funding cost.

Banks are weighted by their municipal-security holdings: state and local securities held to maturity at amortized cost plus those available for sale at fair value, covering all municipal securities, bank-qualified or not, excluding loans and with no bank-size cutoff.²⁸ With M_{nt} these holdings and $\mathcal{B}_t$ the retained banks with positive holdings, the holdings-weighted funding cost and the current-year benefit are

$$c_t = \frac{\sum_{n \in \mathcal{B}_t} M_{nt}(I_{nt}/\bar{A}_{nt})}{\sum_{n \in \mathcal{B}_t} M_{nt}}, \quad X_t = 0.8\tau_{B,t}c_t. \quad (148)$$

The retained banks have positive average assets, nonmissing interest expense and an expense-to-assets ratio from zero to below 15%. Mean c_t is 2.34% in 2000–04, 2.31% in 2005–08, 1.09% in 2009–10, 0.51% in 2011–13, 0.31% in 2014–17 and 0.66% in 2018–19. The corresponding current-year benefits X_t are approximately 66, 65, 30, 14, 9 and 11 bp, with the 2009–10 value under the ordinary rule.

Scope of the measure. The Call Report measure approximates the tax allocation with book assets and expense, while the statute uses adjusted tax bases and otherwise-deductible interest, which are not observed bank by bank, and the IRS permits specified book-asset averaging conventions. X_t is an observable C-corporation scale of the ordinary-rule benefit and does not replicate any bank’s tax-return calculation.²⁹ The weights include Subchapter S banks’ municipal holdings. In 2009–10, the Section 265(b)(7) exception for holdings up to two percent of assets could reduce the disallowance on non-designated bonds and with it designation’s incremental benefit. Treating 2009–10 through the \$30 million cap alone omits this provision, and those years are reported separately.

Bank holdings of municipal securities rose from \$97 billion in 2000 to \$370 billion in 2017 and fell after the 2018 rate cut.

The expected annual benefit. Because a bond’s benefit accrues over its life, the expected annual benefit is $\bar{X}_b = 0.8\tau_{B,t}\hat{c}_b$ for a bond b issued in month t , where $\hat{c}_b$ is the expected annual funding cost over the bond’s life. Fitting annual holdings-weighted funding costs to one-year expected rates over 2000–2021 ($R^2 = 0.75$) gives $\hat{c}_b = 0.38 + 0.55r_b$. Here r_b is the issue-month Kim–Wright (Kim and Wright, 2005) expected-average-short-rate component at the rounded whole-year maturity, capped at ten years. Both $\hat{c}_b$ and r_b are expressed in percent per year. The expected rate is the fitted yield less its term premium, which keeps a term premium rising with maturity out of the denominator of the scaled outcome. Bin means of $\bar{X}_b$ are 33.9, 41.5, 50.5, 56.8 and 53.3 bp.

C.8.4 Scaled and interacted specifications

Premia and the current-year benefit. The current-year benefit X_t and the expected annual benefit over a bond’s life, $\bar{X}_b$, answer different questions. The first falls from 104 bp in 2000 to an average of 9 bp in 2013–17, but premia do not track it. Year terciles with average benefits of 8.5, 26.7 and 70.2 bp give 5–20

²⁸Interest expense is BankFind series EINTEXP (Schedule RI, item 2.e, RIAD4073), average assets ASSET5 and municipal holdings SCMUNI (Schedule RC-B, item 3, columns A/D, 8496/8499 with applicable prefixes), as defined in the FDIC’s [Financial API definitions](#) and the FFIEC [December 2017 Call Report](#), Schedules RI and RC-B. BankFind harmonizes the historical regulatory-report series.

²⁹See the IRS, [Bank Qualified Bonds: Section 265](#), pp. 13-11–13-12, for the allocation and 20% disallowance. The annual, per-dollar-of-basis benefit is distinct from the present value of deductions over a bond’s remaining life.

year premia of 20.8–21.1, 12.3–23.2 and 6.5–13.2 bp, each with a significant long-minus-short contrast. In two- to four-year blocks, ratios of the 5–20 year premium to X_t exceed two in 2013–17 and 2020–21, while the pre-2009 estimates (Appendix C.8.6) are at most about 0.3 of X_t . Ratios to this current-year benefit therefore depend on how years are weighted. The next specification instead uses expected benefits over each bond’s life.

Outcome scaled by the expected benefit. Dividing each offering yield by its own $\bar{X}_b$ before 2SLS estimation of (27), bond by bond and before any pooling, gives the premia in expected-benefit units in column (3) of Table 35: 0.17, 0.31, 0.41, 0.52 and 0.47, with a 10–20 minus 1–2 year contrast of 0.34 (s.e. 0.11). The Adrian–Crump–Moench (Adrian et al., 2013) expected-short-rate component subtracts larger term premia and gives 0.17, 0.35, 0.46, 0.55 and 0.48 (contrast 0.39).

Long-minus-short contrasts remain positive within benefit regimes: 0.57 in low-benefit years, 0.37 in middle years and 0.27 in high-benefit years. In the high-benefit years 2000–02 and 2005–08, expected benefits were nearly flat in maturity (62.8–71.2 bp) and yield curves often inverted, so the benefit’s maturity structure does not drive the gradient. Excluding the strong bank-demand years 2020–21 gives 0.18, 0.25, 0.35, 0.39 and 0.36 (contrast 0.21).

Yield-level benefit interactions. A second specification keeps offering yields in basis points and interacts designation with the expected annual benefit:

$$y_b = \sum_{g=1}^5 \beta_g B_b \bar{X}_b \mathbf{1}\{n_b \in g\} + \widetilde{W}_b' \delta + u_b, \quad (149)$$

$$B_b \bar{X}_b \mathbf{1}\{n_b \in g\} \text{ instrumented by } E_{jt} \bar{X}_b \mathbf{1}\{n_b \in g\}, \quad g = 1, \dots, 5.$$

Here $\widetilde{W}_b$ adds $\bar{X}_b$ by bin to the fixed effects and running-variable controls of (27), and u_b is the regression error. Table 36 reports the yield reductions $-\beta_g$ and the joint first-stage strength. Table 35 divides the outcome by $\bar{X}_b$, whereas this specification multiplies designation and its instruments by $\bar{X}_b$ and adds benefit controls, so their levels need not coincide. Both give rising profiles. The statement of Section 6.2 that about half of the benefit is priced at 10–20 years refers to the outcome-scaled specification. Excluding 2009–10 leaves the interacted 10–20 minus 1–2 year contrast at 0.113 (s.e. 0.032).

Table 36: Bank designation interacted with the expected annual benefit.

Maturity (years)	1–2	2–5	5–10	10–20	> 20
Yield reduction per benefit unit	0.058** (0.024)	0.096*** (0.025)	0.154*** (0.027)	0.175*** (0.031)	0.233*** (0.089)
10–20 minus 1–2	0.117*** (0.032)				

Notes. Entries are $-\beta_g$ in equation (149), the offering-yield reduction in basis points per basis point of expected annual benefit $\bar{X}_b$. Each designation–benefit interaction is instrumented by the below-cap indicator times the benefit in that bin. Controls add the benefit by bin to the fixed effects, log tranche par and running-variable slopes of Table 35, on its 2000–2021 sample within \$0.5–\$4 million of each year’s cap. Observations: 355,372, issuer clusters: 15,889, joint Kleibergen–Paap F : 122.0. Issuer-clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.8.5 Robustness

Donut, bandwidth and window. Donuts from zero to \$1.5 million and bandwidths from \$2 million to \$8 million (Table 37) give 10–20 minus 1–2 year contrasts of 13.5–18.4 bp, and the 1–2 year premia are never significant. In the \$5–\$15 million window of Dagostino (2025), without a donut, the contrast is 15.4 (1.8) bp, one of the most precise estimates.

The last row of Table 37 drops the region her bunching estimation treats as contaminated, \$8.5–\$13.5 million at the \$10 million cap, so within $|R| \leq 8$ it compares issuers at least \$1.5 million below the cap with issuers at least \$3.5 million above it. First stages rise to 0.47–0.54, and the more distant extrapolation gives premia of 7.6 and 10.0 bp at 5–10 and 10–20 years and a contrast of 9.8 (4.2).

Table 37: Bank-qualified discontinuity: sensitivity to the donut, the bandwidth and the window.

	first stage	1–2y	2–5y	5–10y	10–20y	20y+	10–20 minus 1–2
<i>Donut half-width, bandwidth \$4 million</i>							
none	0.39***	1.9 (1.8)	7.5*** (1.8)	13.9*** (1.8)	17.7*** (1.8)	17.7*** (4.5)	15.7*** (2.0)
\$0.25 million	0.40***	1.6 (2.0)	7.6*** (2.0)	13.4*** (2.0)	16.7*** (2.2)	16.3*** (5.5)	15.1*** (2.4)
\$0.5 million (baseline)	0.42***	1.8 (2.2)	8.7*** (2.2)	13.8*** (2.2)	18.4*** (2.3)	18.2*** (6.3)	16.5*** (2.7)
\$1 million	0.44***	1.2 (2.9)	8.2*** (2.7)	13.3*** (2.8)	19.7*** (3.1)	20.6*** (8.0)	18.4*** (3.5)
\$1.5 million	0.44***	2.0 (4.4)	7.3* (3.8)	11.8*** (4.0)	15.5*** (4.2)	19.5* (10.2)	13.5*** (5.2)
<i>Bandwidth, donut \$0.5 million</i>							
\$2 million	0.38***	0.8 (4.7)	7.9 (4.9)	13.3*** (4.8)	16.9*** (5.0)	21.0 (14.1)	16.1*** (4.8)
\$3 million	0.40***	2.7 (3.0)	8.9*** (3.0)	15.2*** (2.9)	19.2*** (3.1)	19.9*** (7.6)	16.5*** (3.3)
\$5 million	0.42***	−0.9 (1.9)	5.7*** (1.8)	11.4*** (1.8)	14.9*** (1.9)	15.9*** (4.9)	15.8*** (2.3)
\$6 million	0.43***	−0.5 (1.7)	5.4*** (1.5)	10.8*** (1.6)	13.2*** (1.6)	11.4*** (4.1)	13.6*** (2.1)
\$8 million	0.45***	0.0 (1.5)	6.3*** (1.2)	10.1*** (1.3)	13.5*** (1.3)	12.3*** (3.5)	13.5*** (1.8)
<i>Other windows</i>							
\$5–15 million, no donut	0.40***	0.1 (1.6)	5.6*** (1.5)	12.1*** (1.6)	15.5*** (1.5)	16.7*** (3.8)	15.4*** (1.8)
\$8.5–13.5 million dropped, $ R \leq 8$	0.47***	0.2 (3.5)	6.0** (3.0)	7.6** (3.1)	10.0*** (3.4)	−1.8 (9.5)	9.8** (4.2)

Notes. Cells report the 2SLS premium of designation in basis points by maturity bin, pooled over 2000–2021 at each year’s cap, with the baseline fixed effects and issuer-clustered standard errors in parentheses. The bandwidth and the donut half-width are the largest and smallest $|R_{jt}|$ retained, and the first-stage column refers to the 1–2 year bin. The \$5–15 million window is the one used in the spread comparison of [Dagostino \(2025\)](#). The last row drops the region her bunching estimation excludes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Construction and the relation to [Dagostino \(2025\)](#). [Dagostino \(2025\)](#) measures issuer-year issuance with Ipreo MuniIC data and documents the same bunching and first stage at the \$30 million cap in 2009–10. Her regressions on \$1 million issuance bins control for size, maturity, insurance, rating and use of proceeds, with year and issuer effects and standard errors clustered by year. In that study, tax-adjusted spreads over treasuries jump about 10 bp at the cap for general-obligation (GO) bonds issued before 2008. In her Call Report data for 2000–2010, designated bonds are about 4% of the average bank’s assets, rising to 5% by 2010, about nine times the non-designated share. Table 38 compares her construction choices with mine and nets refunding par from the running variable. Every row retains a larger long-maturity premium: issuer effects steepen it, while netting refundings shifts issuer-years and widens confidence intervals.

Applying her \$1 million issuance-bin estimator to her window, with maturity-bucket, state and rating effects by year and use-of-proceeds controls, gives flat below-cap bins and a step above the cap, matching her Figure 3. Pooled across maturities, the step is 4.2 bp in the offering yield and 5.6 bp in the tax-adjusted spread, 1.1 and 1.4 bp at 1–5 years and 7.8 and 10.2 bp at 10–20 years. In her pooled-spread setting for GO bonds issued before 2008, with her controls, state fixed effects and issuer clustering, the reduced-form jump at the cap is 10.4 bp (s.e. 2.5), against her approximately 10. With her issuer fixed effects and clustering by year, which leaves eight clusters, the jump is 6.3 (3.8), because within-issuer variation across the cap is thin before 2008.

C.8.6 Placebos and periods

Placebo cutoffs and the ARRA cap. At a placebo cutoff of \$5 million, first stages are 0.00–0.05 and all yield effects are within two bp of zero, with a 10–20 minus 1–2 year contrast of −0.4 (1.2). At \$15 million, first stages are −0.01 to 0.02, and yield effects are 0.1, −0.6, −1.7, −2.4 and −8.6 bp and are not jointly significant ($p = 0.17$). The 10–20 minus 1–2 year contrast is −2.5 (1.6), against −7.8 (1.2) in the reduced form at the cap. Around the \$30 million cap of the American Recovery and Reinvestment Act (ARRA), with the baseline fixed effects and a \$0.5–\$8 million window, the 2009–10 first stages are 0.19, 0.20, 0.19, 0.16 and 0.26 on 16 thousand bonds from issuers with \$22–\$38 million of issuance, jointly significant

Table 38: Bank-qualified discontinuity: construction checks.

	1–2y	2–5y	5–10y	10–20y	20y+	10–20 minus 1–2
Baseline (Table 35)	1.8 (2.2)	8.7*** (2.2)	13.8*** (2.2)	18.4*** (2.3)	18.2*** (6.3)	16.5*** (2.7)
Bond insurance and enhancement $\times$ year	0.9 (2.2)	8.0*** (2.2)	13.0*** (2.2)	17.7*** (2.3)	17.3*** (6.2)	16.8*** (2.6)
General-obligation bonds only	1.7 (2.3)	7.6*** (2.2)	13.3*** (2.3)	15.2*** (2.5)	15.8** (7.3)	13.5*** (2.9)
Issuer fixed effects	–5.0* (2.6)	1.8 (2.5)	7.5*** (2.6)	17.3*** (2.8)	16.8*** (4.1)	22.3*** (2.4)
Tax-adjusted spread over treasuries	5.3 (3.3)	12.6*** (3.2)	21.4*** (3.2)	26.9*** (3.5)	22.8*** (8.6)	21.6*** (3.8)
Insurance controls, year clusters (22)	0.9 (2.9)	8.0*** (2.6)	13.0*** (3.6)	17.7*** (3.4)	17.3** (7.6)	16.8*** (3.7)
Running variable net of refundings	0.8 (3.1)	4.9 (3.1)	8.1*** (3.1)	15.7*** (3.6)	11.0 (8.2)	14.8*** (3.7)
and refunding bonds excluded	0.9 (3.1)	6.3* (3.2)	8.9*** (3.2)	16.0*** (3.6)	10.9 (7.5)	15.1*** (3.6)

Notes. Cells report the 2SLS premium of designation in basis points, pooled over 2000–2021 at the year’s cap in the window \$0.5 to \$4 million, with the baseline fixed effects and issuer-clustered standard errors unless stated. First stages are 0.40 to 0.51 in every bin of every row. Insurance and enhancement: adds bond insurance $\times$ year and credit enhancement $\times$ year (47% of window bonds are insured). The general-obligation, issuer-fixed-effect, tax-adjusted-spread and year-clustered rows also include these controls. General obligation: unlimited-tax, limited and double-barreled general-obligation bonds, 70% of the window. Issuer fixed effects: identified from issuers that cross the cap across years. Tax-adjusted spread: $100 [y/(1 - \tau_{B,t}) - y^{\text{TSY}}]$, the outcome of [Dagostino \(2025\)](#). Net of refundings: the running variable excludes refunding par, and the last row also drops refunding bonds from the estimation sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

at five percent. At that cutoff, first stages are 0.01–0.04 in 2004–08 and –0.05 to 0.01 in 2011–15. At \$10 million in 2009–10, they are 0.01–0.06 and not jointly significant. The 2009–10 reduced-form yield effects at the \$30 million cap are –0.5, –3.2, –1.6, –10.9 and –21.5 bp, and the average beyond ten years minus the 1–2 year effect is –15.7 (6.7), significant at five percent.

By period. Period estimates for 2000–04, 2005–08, 2009–10, 2011–13, 2014–17 and 2018–19 use the baseline fixed effects and issuer clustering, with the \$30 million cap and a \$0.5–8 million window in the ARRA period and the \$10 million cap and \$0.5–4 million window otherwise. Before 2009, with funding costs c_t between 2.3 and 2.4% and X_t of 65–66 bp, premia are at most about 0.3 of X_t at every maturity: the deduction is weakly capitalized when the carrying penalty on non-designated bonds approaches a percentage point.

The 2009–10 \$30 million cap has a weaker first stage than the \$10 million cap’s 0.42–0.45, reflecting limited use. With an ordinary-rule X_t of 30 bp, the 2009–10 premia are 11 (24), 22 (23), 17 (23), 50 (28) and 60 (29) bp. The contrast between the average beyond ten years and the 1–2 year premium is 43 (20), and its reduced-form counterpart above is significant at five percent.

In 2011–17, with funding costs of 0.31–0.51% and X_t of 9–14 bp, premia are 2.4–2.5 bp at 1–2 years and 14–28 bp at 5–20 years, at or above X_t .

After the 2018 cut in the bank tax rate, premia are 6.9 (4.3), 6.4 (5.0), 7.7 (5.7), 14.9 (7.7) and 3.7 (24.2) bp. A difference-in-discontinuities estimate, which pools 2014–2019 and interacts designation with a post-2018 indicator, gives post-2018 changes of 3.1 (6.5), –9.1 (6.7), –19.2 (7.6), 3.6 (10.0) and –1.3 (28.7) bp, with unchanged first stages. Subchapter S banks, whose income is taxed to their shareholders and which the 2017 Tax Cuts and Jobs Act (TCJA) affected less, hold 6–14% of banks’ municipal holdings over 2000–2021.

Annual reduced-form yield reductions at the \$10 million cap, pooled across maturities with the baseline fixed effects, are largest in 2016 and 2019–21, at 11–18 bp, are 6–8 bp in 2011–13 and are within about 3 bp of zero in 2000–01, 2004–07 and 2018. First stages are 0.30–0.62 in every year except 2009–10, when the cap was \$30 million. Premia do not simply track the current-year benefit, and the pattern is consistent with changing bank participation.

C.9 Investor holdings by tax status

Portfolio shares vary with exemption benefits (Section 3.4 and Table 2). In 2000–2017, exempt municipal bonds account for 34–48% of the average P&C insurer’s bond portfolio, compared with 6–9% for life insurers (Figure 25). Public pensions’ exempt and taxable municipal shares never exceed 0.13% and 0.78%, respectively (Figure 26). After the corporate-rate cut, the P&C exempt share fell from 38.0% in 2017 to 28.6% in

2019, while the other-taxable share rose from 50.2% to 57.5%. This reallocation is the quantity counterpart of the repricing in the insurer test of P2 (Appendix C.10).

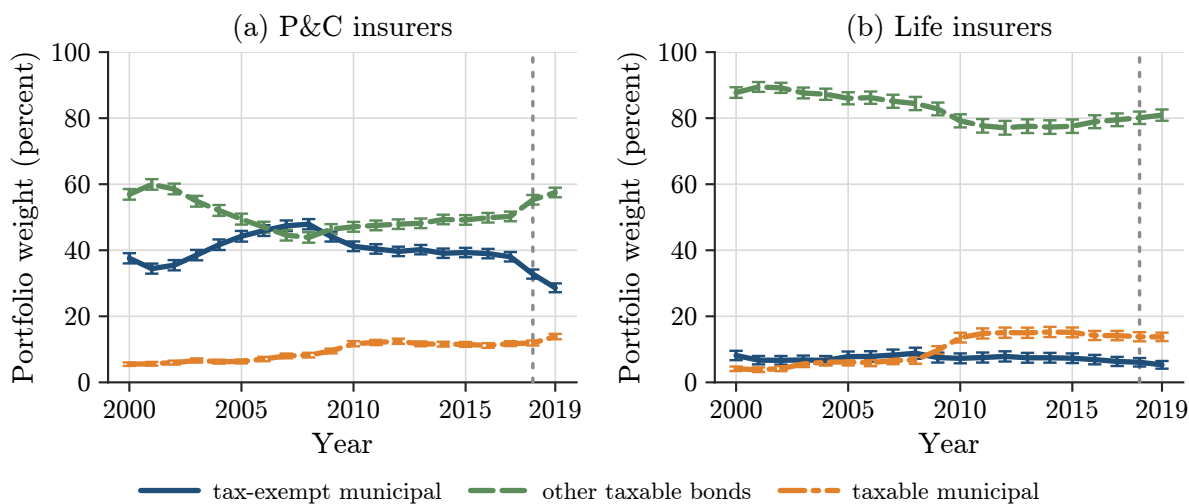


Figure 25: Tax-exempt municipals, taxable municipals and other taxable bonds as shares of insurers’ total bond holdings: P&C in panel (a) and life in panel (b), 2000–2019, NAIC Schedule D. The dotted vertical line marks the TCJA corporate-rate cut. Bars are 95 percent intervals for yearly mean insurer portfolio shares, with standard errors clustered by insurer.

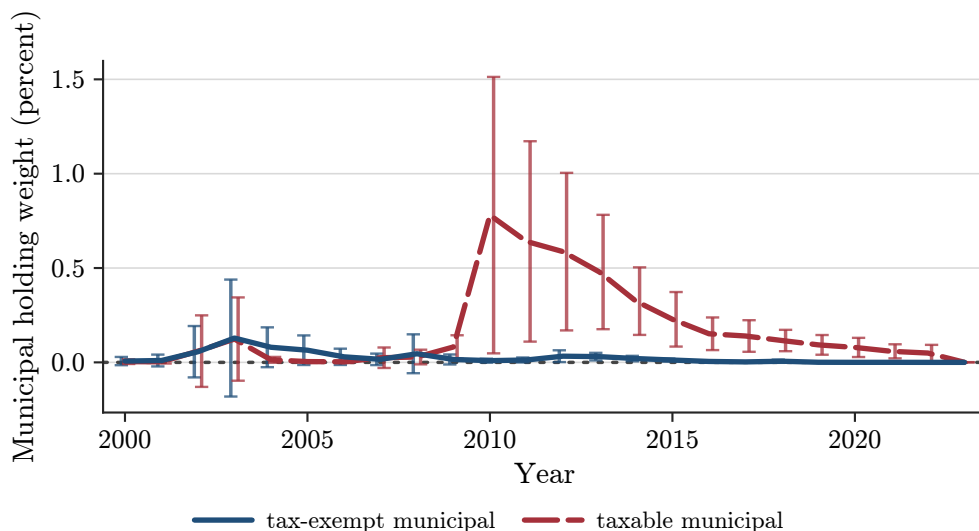


Figure 26: Municipal holdings by federal tax status as shares of total bond portfolios in the public-pension sample obtained by Freedom of Information Act (FOIA) requests (Section 3.1), 2000–2023. All thirty plans are covered from 2015. Bars are 95 percent intervals for yearly mean plan portfolio shares, with standard errors clustered by plan. Other taxable bonds are omitted. Taxable municipals include Build America Bonds.

C.10 The 2017 Tax Cuts and Jobs Act and insurer exposure: difference-in-differences and event study

This test asks whether the TCJA’s reduction in P&C insurers’ exemption benefit lowered yield-implied taxes more at long maturities in states where these insurers held more debt before the reform. The difference-in-

differences (DiD) design compares maturity profiles across states with different pre-reform insurer exposure. Section 6.2 summarizes the result. Appendices C.10.1 and C.10.2 discuss inference and the concurrent household tax change.

Prediction. Appendix Proposition A.9 gives zero pass-through of the lower-tax investor’s exemption benefit at the riskless short end, where the household’s after-tax cash return anchors the yield, and positive pass-through at the long end. Through portfolio adjustment, a change in that benefit also reaches risky maturities shorter than those the lower-tax investor holds. P&C insurers are lower-tax institutions in this sense. Since the TCJA lowered their exemption benefit, P2 predicts that after the reform the implied tax fell more at longer maturities in states whose municipal debt P&C insurers held more heavily.

The 2017 tax change. The TCJA cut the corporate rate from 35% to 21% effective 2018Q1, and the Section 832 proration rule amplified the resulting decline in P&C insurers’ exemption benefit. The disallowed fraction of their tax-exempt interest rose from 15 to 25 percent, so the effective tax on municipal interest stayed at 5.25% ($35\% \times 15\%$ before, $21\% \times 25\%$ after), while the after-tax share of taxable interest rose from 0.65 to 0.79. The P&C relative after-tax share ξ_P rose from 0.686 to 0.834, and the exemption benefit $1 - \xi_P$ narrowed from 0.31 to 0.17.

The top household ordinary rate also fell, from 39.6% to 37%, and the cap on state and local tax (SALT) deductions changed the value of in-state exemptions. Life insurers’ relative after-tax share rose from 0.65/0.825 $\simeq$ 0.788 to 0.79/0.937 $\simeq$ 0.843 (Appendix D.1), and banks’ taxable after-tax share and the tax cost of their financing changed. The design compares states with different P&C exposure, so tax changes common to all states drop out.

Treatment and sample. The treatment is a state’s exposure to P&C insurers, $PCShare_s$: the market value of P&C insurers’ holdings of state s ’s municipal bonds at the end of 2015 relative to the state’s par outstanding, before the 2016Q3 start of the estimation window and two years before the reform. Across the 50 states and the District of Columbia in the sample, it averages 0.103 with a standard deviation of 0.030, and each bond b takes the exposure of its issuing state, $PCShare_{s(b)}$. The sample is the 1,733,881 customer trades reported to the Municipal Securities Rulemaking Board (MSRB) over 2016Q3–2019Q1 in non-callable municipal bonds rated AA and above and not subject to the alternative minimum tax (AMT). Restricting to non-callable bonds rules out direct repricing of embedded options. The outcome is the implied tax of (2) for each trade, $ImpliedTax_{b,t} = 1 - y_{b,t}^M/y_t^C(n_b)$, computed from the trade yield and the corporate zero yield for the last month of the trade’s calendar quarter at a maturity equal to the bond’s maturity year minus the trade year.

Specification. I estimate

$$ImpliedTax_{b,t} = \beta_0(PostTCJA_t \times PCShare_{s(b)}) + \beta_1(PostTCJA_t \times PCShare_{s(b)} \times n_b) + \alpha_{[n_b] \times Purchase} + \alpha_{issuer \times Purchase} + \alpha_t + \Gamma' Z_{b,t} + \varepsilon_{b,t}, \quad (150)$$

where $PostTCJA_t = \mathbf{1}\{t \geq 2017Q4\}$ starts the post period at the enactment quarter and n_b is the bond’s remaining maturity in years. The fixed effects compare trades of the same issuer and of the same whole-year maturity, purchases with purchases and sales with sales ($Purchase_{b,t}$ is the MSRB dealer-side indicator for a customer purchase), and α_t absorbs each quarter’s level. The controls $Z_{b,t}$ are $PCShare_{s(b)} \times n_b$; the issuing state’s resident tax rate, the pre-period life-insurer holding share and the bond rating, each alone and interacted with maturity; and interactions of event time with maturity and with the state tax rate. The event-time-by-maturity interactions let the common maturity gradient change each quarter, so β_1 is identified from cross-state differences in $PCShare$ within a quarter.

The differential effect of exposure at maturity n_b is $\beta_0 + \beta_1 n_b$. The model predicts $\beta_1 < 0$, and the representative-tax benchmark predicts $\beta_0 = \beta_1 = 0$.

Instruments. Pre-reform P&C ownership may be correlated with state conditions that also move the maturity profile of implied taxes, so the instrumental-variables (IV) specifications instrument $PCShare_s$ with the municipal capacity of the P&C insurers domiciled in the state:

$$IV1_s = \log\left(\frac{\sum_{i \in D_s} \text{Municipal Assets}_i}{\text{Municipal Outstanding}_s}\right), \quad IV2_s = \log\left(\frac{\sum_{i \in D_s^{\text{act}}} \text{Bond Assets}_i}{\text{Municipal Outstanding}_s}\right), \quad (151)$$

where D_s is the set of P&C insurers domiciled in state s and D_s^{act} the subset holding any municipal bonds. $\text{Municipal Assets}_i$ and Bond Assets_i are insurer i 's par holdings of municipal bonds and of all bonds in its Schedule D filings with the National Association of Insurance Commissioners (NAIC), and $\text{Municipal Outstanding}_s$ is the state's municipal par outstanding. $IV2_s$ uses total bond holdings, so it does not depend on which municipal bonds the domiciled insurers chose to hold. For each instrument, its interactions with n_b , with $PostTCJA_t$ and with their product serve as instruments for the three corresponding exposure terms. Relevance comes from insurer home bias: locally domiciled insurers disproportionately hold the state's bonds. The Kleibergen–Paap first-stage F statistics are 25.7 for $IV1_s$ and 17.5 for $IV2_s$.

The exclusion restriction is that, conditional on the fixed effects and controls, the domiciled insurers' capacity affects the post-TCJA change in the maturity gradient of implied taxes only through P&C insurers' exposure to the state's bonds. Insurer domicile largely predates the reform. Quarter effects and event-time-by-maturity controls absorb the common component of the federal tax changes. Event-time-by-state-tax interactions absorb level changes associated with state tax rates, including SALT-induced ones, and state-tax-by-maturity controls capture a persistent maturity gradient, but neither absorbs a post-reform change in the gradient that varies with state tax rates. Appendix C.10.2 states the covariance conditions under which omitting such a SALT gradient biases the OLS estimate toward zero.

Results. Table 39 reports the static DiD estimates. The maturity gradient β_1 is -0.017 under OLS and -0.038 and -0.040 under the two IV specifications, with issuer-clustered standard errors of 0.005, 0.010 and 0.015. For two states whose pre-reform P&C shares differ by one standard deviation, 3.0 percentage points, the implied tax in the more exposed state falls after the reform by an additional 0.05 percentage point per year of maturity under OLS and 0.11 under the first IV specification. Between a five- and a twenty-five-year bond, the first IV estimate cumulates to about 2.3 percentage points of implied tax. At a corporate yield of four percent, that is about 9 basis points of municipal yield relative to the taxable yield.

Table 39: TCJA difference-in-differences: implied-tax response to insurer exposure. The sample is MSRB trades in non-callable, AA-and-above, non-AMT municipal bonds in 2016Q3–2019Q1. The outcome is the implied tax of equation (2), and $PCShare$ is the end-2015 P&C ownership share. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

	(1) OLS	(2) IV1	(3) IV2
$PostTCJA \times PCShare$	-0.122^{***} (0.029)	-0.129 (0.084)	-0.017 (0.100)
$PostTCJA \times PCShare \times n_b$	-0.017^{***} (0.005)	-0.038^{***} (0.010)	-0.040^{***} (0.015)
$[n_b] \times Purchase$ FE	Yes	Yes	Yes
$Issuer \times Purchase$ FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Issuer cluster	Yes	Yes	Yes
Observations	1,733,881	1,733,881	1,733,881
First-stage F (Kleibergen–Paap)		25.7	17.5

Notes. Columns (2) and (3) instrument $PCShare$ with $IV1_s$ and $IV2_s$ of equation (151). All columns include the controls of equation (150), among them the rating and its interaction with n_b . Standard errors are clustered by issuer (six-digit CUSIP, 9,903 issuers) and reported in parentheses.

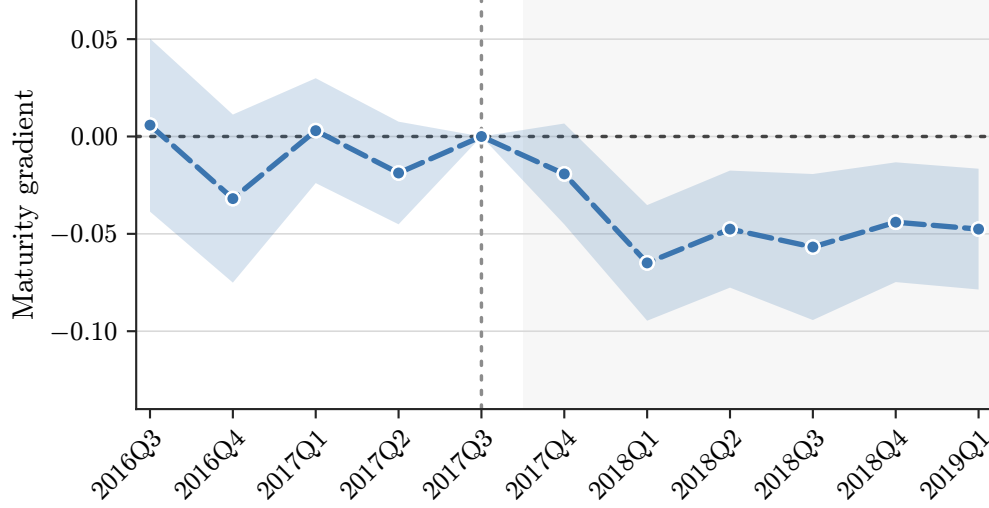


Figure 27: IV1 event-study estimates of the quarter-specific maturity gradient, the coefficient on $\mathbf{1}\{t = q\} \times PCShare \times n_b$, relative to 2017Q3 (vertical dotted line). Shading marks the post-reform quarters from 2017Q4. Bands are 95% intervals clustered by six-digit CUSIP issuer. Appendix C.10.1 reports further inference checks.

The maturity gradient persists on the wider A-or-above sample (Appendix C.12.2 and Table 46).

Event study. The event study replaces the post-reform indicator in (150) with quarter indicators, omitting 2017Q3, so $\hat{\beta}_{1,q}$ measures the maturity gradient in quarter q relative to 2017Q3. The IV versions use the corresponding quarter interactions of the instruments, with the fixed effects and clustering of Table 39. These estimates locate the change at the reform. Across all three estimators, eleven of the twelve pre-TCJA coefficients are individually insignificant at five percent, while the coefficients turn significantly negative at 2018Q1 in every column and stay negative. Five of the six post-reform quarters exclude zero under IV1 and four under IV2 (Figure 27; Table 40 reports each quarter).

C.10.1 State-clustered inference

As a check for shocks shared within states, I also cluster the insurer regressions above by issuing state (51 clusters). Point estimates are unchanged. State-clustered standard errors are 0.0129, 0.0245 and 0.0342 for the OLS, IV1 and IV2 maturity coefficients, and their 95% intervals include zero. The post-reform level coefficient under OLS remains negative and significant: -0.122 (s.e. 0.061), with $p = 0.044$ from a state-level wild-cluster bootstrap (9,999 replications).

C.10.2 SALT-induced household demand and the insurer-channel regression

A binding SALT cap can raise high-tax households' exemption benefit and lower municipal yields where they are active. The conditions below determine when this household channel attenuates the OLS insurer coefficient.

Write $Y_{b,t} = 1 - y_{b,t}^M/y_t^C(n_b)$ for the implied tax of equation (2), $X_{b,t} = PostTCJA_t PCShare_{s(b)} n_b$ for the insurer interaction and $W_{b,t} = PostTCJA_t Tax_{s(b)} n_b$ for a household-SALT interaction, where the state tax rate Tax_s proxies SALT exposure. Here $X_{b,t}$ is unrelated to the bank benefit X of Appendix C.8. Consider

$$Y_{b,t} = \beta_1 X_{b,t} + \beta_2 PostTCJA_t Tax_{s(b)} + \beta_3 W_{b,t} + \text{controls} + \text{FE} + \varepsilon_{b,t}. \quad (152)$$

The two hypotheses imply $\beta_1 < 0$ for the insurer mechanism, and $\beta_2 > 0$ and $\beta_3 < 0$ for stronger household demand at the short end.

Table 40: TCJA dynamic event study: full triple-interaction coefficients $\hat{\beta}_{1,q}$ on $\mathbf{1}\{t = q\} \times PCShare \times n_b$, by quarter. Standard errors in parentheses are clustered by issuer (six-digit CUSIP) as in Table 39, and 2017Q3 is the omitted base quarter. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Quarter q	(1) OLS	(2) IV1	(3) IV2
2016Q3	0.019** (0.010)	0.006 (0.023)	-0.020 (0.036)
2016Q4	-0.014 (0.009)	-0.032 (0.022)	-0.029 (0.036)
2017Q1	0.001 (0.007)	0.003 (0.014)	0.009 (0.020)
2017Q2	-0.005 (0.006)	-0.019 (0.013)	-0.021 (0.020)
2017Q4	-0.001 (0.006)	-0.019 (0.013)	-0.021 (0.020)
2018Q1	-0.028*** (0.007)	-0.065*** (0.015)	-0.063*** (0.023)
2018Q2	-0.027*** (0.007)	-0.048*** (0.015)	-0.039* (0.023)
2018Q3	-0.031*** (0.009)	-0.057*** (0.019)	-0.063** (0.027)
2018Q4	-0.011 (0.008)	-0.044*** (0.016)	-0.051** (0.023)
2019Q1	-0.014* (0.007)	-0.048*** (0.016)	-0.065*** (0.024)
$[n_b] \times Purchase$ FE	Yes	Yes	Yes
$Issuer \times Purchase$ FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Issuer cluster	Yes	Yes	Yes
Observations	1,733,881	1,733,881	1,733,881

Suppose that the SALT level term is controlled for, that W is omitted and that the remaining disturbance is orthogonal to residualized X , with tildes denoting residuals after projecting on all included controls and fixed effects. The omitted-variable formula is

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \beta_3 \frac{\text{Cov}(\tilde{X}, \tilde{W})}{\text{Var}(\tilde{X})}. \quad (153)$$

Consequently,

$$\beta_3 < 0, \quad \text{Cov}(\tilde{X}, \tilde{W}) < 0 \implies \text{plim } \tilde{\beta}_1 > \beta_1. \quad (154)$$

Equation (153) uses residualized variables, so it accounts for controls, fixed effects and weights. An upward bias attenuates a negative coefficient as long as the estimate keeps its sign.

IV identifies the insurer coefficient when the instruments are relevant and orthogonal to the remaining household-demand term, under the exclusion restriction stated in Appendix C.10.

Table 41: P3: observed coupon-yield response to treasury par yield by remaining maturity.

	(1) Continuous $\Delta y_{b,t}^M$	(2) Bin $\Delta y_{b,t}^M$
$\Delta y_{n_b,t}^{\text{Tsy}}$	0.346*** (0.016) [0.012]	0.351*** (0.013) [0.006]
$\Delta y_{n_b,t}^{\text{Tsy}} \times n_b$	0.007*** (0.002) [0.002]	
$\Delta y_{n_b,t}^{\text{Tsy}} \times Long_b$		0.075*** (0.012) [0.009]
$Long_b$		0.016*** (0.002) [0.002]
n_b	0.002*** (0.000) [0.000]	
Trade-date cluster	Yes	Yes
Observations	952,874	952,874
Weights	Inverse pair count	Inverse pair count

Notes. The sample is pairs of consecutive same-side MSRB customer trades within 90 days in federally tax-exempt, non-callable bonds rated AA and above (agency score ≤ 3), January 2005 to August 2020, over which the maturity-matched treasury yield moved at least 10 basis points in either direction. The treasury yield is interpolated on the daily GSW curve at the bond's exact remaining maturity on each transaction date. Observations are weighted by the inverse count of their CUSIP's base-sample trade pairs, computed before observations with missing or trimmed outcomes are dropped. The dependent variable is trimmed at its first and ninety-ninth percentiles. $Long_b = \mathbf{1}\{n_b \geq 5\}$. First differencing absorbs time-invariant bond characteristics. Parentheses give trade-date-clustered standard errors (3,864 clusters). Brackets give state-clustered standard errors (54 clusters). Stars are identical under either clustering. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.11 P3: municipal-yield response to treasury yields

This subsection explains how the trade data measure the relative yield response in Section 6.3. It first constructs changes in municipal and maturity-matched treasury yields, then checks whether the maturity pattern survives alternative trade windows and the conversion from zero-coupon yields to observed coupon yields.

Transaction sample. Each MSRB customer trade records the CUSIP, trade date, par, dollar price, implied yield to maturity and dealer side. Estimation trades, from January 2005 to August 2020, are in non-callable, federally exempt bonds rated AA and above (agency score ≤ 3 , Section 6). Noncallability removes option contamination. I exclude trades less than 92 days after issue, when offering yields still anchor prices, and trades less than 365 days before maturity, where the fitted treasury curve is erratic.

Matched treasury changes. At each trade, the daily treasury par-yield curve of Gürkaynak, Sack and Wright (GSW) is interpolated linearly between integer maturities at the bond's remaining maturity $n_b(t) = d_{b,t}/365$, where $d_{b,t}$ is its days to maturity, and clamped above thirty years. Within each CUSIP and dealer side, consecutive trades form the municipal yield change $\Delta y_{b,t}^M$ and the matched treasury change $\Delta y_{n_b,t}^{\text{Tsy}} = y_{n_b(t),t}^{\text{Tsy}} - y_{n_b(t-1),t-1}^{\text{Tsy}}$, where $t-1$ denotes the trade preceding t . Comparing purchases with purchases and sales with sales limits bid-ask and intraday contamination. Pairs with gaps of 1–90 days and treasury moves of at least 10 bp are kept, with the weights and trimming of Table 41. Over Δt elapsed days, aging contributes $-(\Delta t/365)\partial_m y_{m,t}^{\text{Tsy}} + O((\Delta t/365)^2)$ to the matched change, an effect measured below.

Specifications. Column (1) of Table 41 estimates the continuous interaction in (28), including maturity but no fixed effects. Column (2) replaces maturity by $Long_b = \mathbf{1}\{n_b \geq 5\}$, so its short and long responses are β_0 and $\beta_0 + \beta_1$.

Holding the benchmark maturity fixed. Part of the matched treasury change comes from the bond becoming shorter between trades. This aging component is $y_{n_b(t),t}^{Tsy} - y_{n_b(t-1),t}^{Tsy}$, holding the later date's curve fixed. Fixing the benchmark at the initial maturity, $\Delta y_{b,t}^{Tsy, \text{fixed}} = y_{n_b(t-1),t}^{Tsy} - y_{n_b(t-1),t-1}^{Tsy}$, removes it. On upward-sloping curves, aging lowers the reference yield by an average of 2.6 bp below five years and 1.6 bp beyond, against absolute changes near 26 bp. Removing this component lowers the short response more than the long response, increasing the gradient (Table 42). Municipal aging does not cancel it because the two curves have different slopes. The third row also controls for linear municipal aging through elapsed time and its maturity interaction, further increasing the gradient. The headline specification keeps the benchmark matched at each trade, the more conservative choice.

Table 42: P3 observed coupon-yield responses and treasury-curve aging.

Treasury benchmark	Bin specification		Interaction	
	< 5y	$\geq 5y$	Bin	Continuous
Benchmark matched at each trade's maturity (Table 41)	0.351***	0.426***	0.0749***	0.0070***
Benchmark held at the first trade's maturity	0.340***	0.423***	0.0825***	0.0084***
plus elapsed time and elapsed time $\times$ maturity	0.334***	0.419***	0.0850***	0.0086***

Notes. Each row estimates (28) on the 952,874 pairs of Table 41, with its weights and trade-date clustering. The first two columns are the bin specification's responses below and from five years, β_0 and $\beta_0 + \beta_1$, and the last two are the interactions with $Long_b$ and n_b . * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Sensitivity to the filter thresholds. Across the treasury-move cutoffs and trade windows of Table 43, all observed maturity-interaction estimates are positive. Every binned cell is significant at one percent, and every continuous cell is significant at five percent, with twenty-seven of the thirty significant at one percent. Appendix C.11.1 applies the coupon benchmark to each cell.

State clustering. Brackets in Table 41 give standard errors clustered by issuing state, which allow correlation among a state's bonds across dates, while the headline trade-date clustering allows common shocks across states on the same date. Both headline interactions remain significant at one percent, and binned-grid significance levels are unchanged. In the continuous grid of Table 43, the 0.05, 0.075 and 0.10 cutoffs give $p < 0.05$ at thirty days and $p < 0.10$ at sixty days, and the 0.15 cutoff gives $p < 0.05$ at sixty days. All other significance levels are unchanged, so twenty-seven of the thirty continuous cells and all binned cells remain significant at five percent.

Taxable-market comparison. I also compare municipal changes with monthly corporate-curve changes. Both reference curves are measured at monthly averages and the bond's remaining maturity in each trade month. Retaining cross-month pairs whose absolute monthly treasury change is at least 10 bp gives 517,864 pairs, with the original inverse-CUSIP weights. The municipal-to-corporate maturity interaction is 0.0132 (s.e. 0.0026), while the municipal-to-treasury interaction on these pairs is 0.0072 (s.e. 0.0025), both with trade-date clustering. The cash-flow comparison below translates each reference market's zero curve into municipal coupon yields on a common sample.

C.11.1 Coupon cash flows and the zero-yield benchmark

P3 concerns zero-coupon responses, while the data report coupon yields and treasury par yields. Even when zero yields satisfy $y^{M,0}(u, t) = \xi_H y^{Tsy,0}(u, t)$ at every horizon u , coupon yields need not have a constant

Table 43: P3 sensitivity: observed maturity interaction across the treasury-move cutoff and trade-gap window (weighted by inverse CUSIP base-pair counts).

<i>Panel A. Continuous interaction $\Delta y_{nb,t}^{\text{Tsy}} \times n_b$</i>					
Cutoff \ window	15d	30d	60d	90d	120d
≥ 0.05	0.0053***	0.0057***	0.0051**	0.0071***	0.0090***
≥ 0.075	0.0061***	0.0057***	0.0047**	0.0068***	0.0087***
≥ 0.10	0.0065***	0.0060***	0.0049**	0.0070***	0.0087***
≥ 0.15	0.0080***	0.0071***	0.0059***	0.0077***	0.0091***
≥ 0.20	0.0088***	0.0082***	0.0067***	0.0081***	0.0094***
≥ 0.25	0.0104***	0.0083***	0.0074***	0.0087***	0.0097***
<i>Panel B. Binned interaction $\Delta y_{nb,t}^{\text{Tsy}} \times Long_b$</i>					
Cutoff \ window	15d	30d	60d	90d	120d
≥ 0.05	0.049***	0.061***	0.070***	0.074***	0.088***
≥ 0.075	0.056***	0.063***	0.068***	0.074***	0.086***
≥ 0.10	0.062***	0.066***	0.069***	0.075***	0.086***
≥ 0.15	0.079***	0.077***	0.079***	0.083***	0.092***
≥ 0.20	0.094***	0.090***	0.085***	0.088***	0.097***
≥ 0.25	0.111***	0.100***	0.096***	0.096***	0.102***

Notes. Each cell re-estimates column (1) (Panel A) or (2) (Panel B) of Table 41 on pairs whose maturity-matched treasury move is at least the cutoff, in percentage points, and whose trades are at most the window apart. The headline sample uses 0.10 and 90 days. Standard errors are clustered by trade date as in Table 41. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

response ratio, because their cash flows span horizons with different yield movements. The coupon benchmark measures the response that coupon conventions alone produce.

I set $\xi_H = 0.65$ and price each bond's regular coupon and redemption payments at municipal zero yields of 0.65 times the GSW treasury zero yields:

$$P_{b,t}^{M,0} = \sum_k CF_{b,k} \exp[-0.65 y^{\text{Tsy},0}(u_{b,k,t}, t) u_{b,k,t}],$$

where $CF_{b,k}$ is payment k and $u_{b,k,t}$ its time to payment, and zero yields are continuously compounded fractions. I then recover the benchmark's nominal-semiannual yield to maturity from this price at each trade date, keeping the bond's coupon and maturity, both transaction dates and the original par-yield regressor.

Semiannual payments are reconstructed backward from maturity with US 30/360 day-count fractions. The comparison retains zero-coupon bonds and fixed-rate semiannual bonds whose first coupon precedes the first trade, and excludes other payment conventions, maturities over thirty years and pairs with incomplete curves. Zero yields are interpolated linearly, with the one-year value held below one year. The 918,350 common pairs retain 94.9% of the original weight, with the original trimming and inverse-CUSIP weights. Benchmark changes are not trimmed.

Table 44 applies the same regression to the observed changes, the benchmark changes and their difference, whose standard error retains their covariance. The coupon benchmark accounts for 30.0% of the observed daily gradient, leaving a difference of 0.0056 (s.e. 0.0021). With after-tax shares from 0.60 to 0.80, the benchmark accounts for 27.6%–37.0%, and the difference remains positive throughout. The binned difference is 0.0623 (s.e. 0.0123). Benchmarks that hold treasury maturity fixed, with or without elapsed-time controls, also leave a positive interaction.

The monthly comparisons keep the cross-month pairs and monthly reference regressors of the taxable-market comparison, restricted to the 500,006 of its 517,864 pairs with regular coupon schedules. The treasury benchmark prices at proportional daily GSW zero curves on trade dates, and the corporate benchmark at

proportional monthly corporate zero curves converted to continuous yields. Both differences remain positive and significant at five percent (Table 44).

Table 44: P3: maturity gradients of observed and coupon-benchmark yield responses.

	Observed	Coupon benchmark	Difference	Pairs
Daily, continuous	0.0080*** (0.0021)	0.0024*** (0.0001)	0.0056*** (0.0021)	918,350
Daily, long minus short	0.0738*** (0.0122)	0.0115*** (0.0003)	0.0623*** (0.0123)	918,350
Fixed initial maturity	0.0095*** (0.0021)	0.0034*** (0.0001)	0.0061*** (0.0021)	918,350
Fixed maturity + gap controls	0.0098*** (0.0021)	0.0036*** (0.0001)	0.0062*** (0.0021)	918,350
Monthly treasury	0.0082*** (0.0026)	0.0019** (0.0009)	0.0063** (0.0026)	500,006
Monthly corporate	0.0141*** (0.0027)	−0.0006*** (0.0001)	0.0147*** (0.0026)	500,006

Notes. The coupon benchmark prices each bond’s payments at municipal zero yields of 0.65 times the reference zero curve (Appendix C.11.1). Observed and benchmark coupon-yield changes use identical pairs, regressors and inverse-CUSIP-base-pair weights, and the difference is estimated jointly. Interactions are per year of remaining maturity, except in the long-minus-short row, which uses $Long_b$. Fixed-maturity rows retain the first trade’s treasury maturity. Gap controls are elapsed days and their maturity interaction. Monthly rows use cross-month pairs and monthly regressors. Parentheses give second-trade-date-clustered standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Applying the coupon benchmark separately in each cell of Table 43, on that cell’s common coupon pairs, leaves all thirty continuous differences positive, from 0.0034 to 0.0104, with twenty-four significant at five percent. All thirty benchmark-adjusted binned differences are positive and significant at five percent, ranging from 0.0393 to 0.1016.

C.12 Robustness to the rating cut

This section checks whether the yield pattern and empirical tests depend on the rating cutoff. Both samples use the unrounded mean agency score on the twenty-two-notch scale, on which 1 = AAA, 3 = AA and 7 = A−. The curves in Sections 3 and 5 use scores of seven or better, the A-or-above sample, matching the corporate benchmark’s credit band. The Build America Bond (BAB) comparison in Section 3 and the P1, Tax Cuts and Jobs Act (TCJA) and P3 tests in Section 6 use scores of three or better, the AA-and-above sample.

The curve pattern survives the narrower cut, and the BAB, P1, TCJA and P3 tests retain their gradient signs on the wider cut. P1’s pre-2018 state-tax pass-through declines across all five bins under OLS and IV (Table 45), with endpoint differences significant at five percent under OLS and one percent under IV. In the TCJA insurer test, the maturity gradients are negative and significant under all three estimators with issuer clustering, and Appendix C.10.1 reports state-clustered inference. The wider sample gives larger P3 interactions and a larger OLS TCJA gradient, with similar BAB, P1 and IV TCJA gradients.

C.12.1 Curve construction at AA and above

Figure 28 uses the AA-and-above sample and its own bond, rating and issue-quarter weights, from the three-stage averaging of Appendix C.1.1. Each implied tax is one minus the ratio of mean municipal to mean taxable yields, averaged with these weights. Net of calls and issuance premia, the implied tax falls from 40.1% at two years to 25.5% at thirty, a decline of 14.6 points against 17.1 on the A-or-above sample.

The model comparison repeats the accounting of Section 5.4, applying this sample’s model ratio of municipal to taxable mean yields to observed corporate yields (the corporate normalization of (22)). Call,

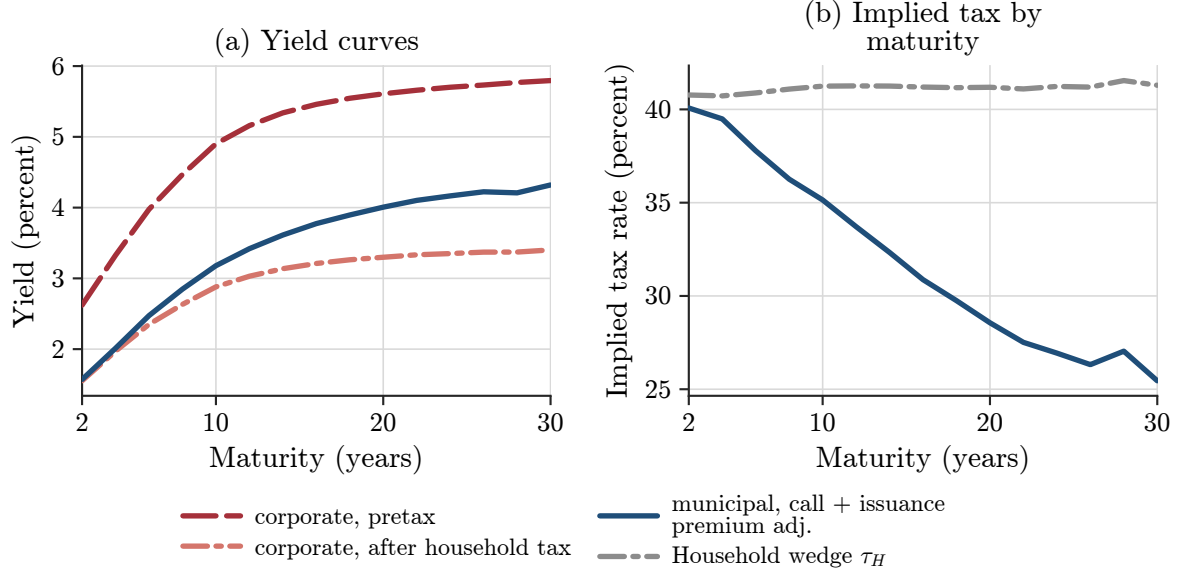


Figure 28: The AA-and-above municipal puzzle, 2000–2020, with this sample’s own weights and household benchmark. Panel (a) shows pretax and household after-tax corporate zero-coupon yields and the municipal yield after removing callability and the issuance premium, and panel (b) shows the implied tax $1 - y_T^M/y_T^C$ and the household wedge $1 - \bar{\xi}_H(T)$, labeled τ_H in the legend. Calls use dated calibrated prices, and the sample’s all-size issuance premium is scaled by the full sample’s institutional-to-all-size ratio (see text).

liquidity and default estimates are sample-specific, and capital gains use the common method. The issuance adjustment scales the sample’s all-size premium by the full-sample institutional-to-all-size ratio, giving 11.5 bp at thirty years. The normalized model gap exceeds the residual before the model at every maturity, 96.2 against 85.6 basis points at thirty years, so all unexplained residuals are negative. Seven of fifteen absolute residuals are below ten basis points, the thirty-year residual is -10.6 and the root-mean-square residual is 10.9 basis points.

C.12.2 Empirical tests on the A-or-above sample

Table 28 (Appendix C.3) compares the two BAB rating cuts in columns (1) and (3), with call matching in columns (2) and (4). Table 45 repeats the P1 specification of Appendix C.6, so its AA-and-above columns reproduce the headline P1 estimates. Table 46 repeats the TCJA difference-in-differences of Table 39 with issuer clustering but omits the headline rating and rating-by-maturity controls in both samples, so its AA-and-above columns differ slightly. Tables 47 and 48 repeat the P3 specification of Table 41 and the binned grid of Table 43 with state-clustered standard errors, where the headline P3 estimates cluster by trade date. Their AA-and-above coefficients equal the headline ones.

Table 45: P1 by rating cut, 2000–2017.

	OLS		IV	
	Score ≤ 3	Score ≤ 7	Score ≤ 3	Score ≤ 7
β_1 : [0, 5) years	1.069*** (0.195)	0.986*** (0.255)	1.088*** (0.225)	1.080*** (0.233)
β_2 : [5, 10) years	0.677*** (0.069)	0.607*** (0.100)	0.694*** (0.086)	0.665*** (0.117)
β_3 : [10, 15) years	0.484*** (0.096)	0.380*** (0.083)	0.447*** (0.118)	0.384*** (0.098)
β_4 : [15, 20) years	0.390*** (0.102)	0.285*** (0.086)	0.326*** (0.116)	0.264*** (0.095)
β_5 : [20, 30] years	0.305*** (0.075)	0.240*** (0.077)	0.229*** (0.081)	0.214*** (0.079)
Decline $\beta_1 - \beta_5$	0.764***	0.746**	0.859***	0.866***
p -value of decline	0.002	0.011	< 0.001	< 0.001
Issue month $\times$ maturity $\times$ call FE	Yes	Yes	Yes	Yes
State cluster	Yes	Yes	Yes	Yes
Observations	681,838	1,397,768	681,838	1,397,768
First-stage F (Kleibergen–Paap)	—	—	10.4	12.6

Notes. Controls, fixed effects and state-clustered inference (47 states) follow Appendix C.6, 2000–2017, with call-adjusted zero yields. Score ≤ 3 is the AA-and-above sample and score ≤ 7 the A-or-above sample. IV uses bin-interacted 1990 rates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 46: P2: TCJA insurer difference-in-differences, by rating cut

	OLS		IV1		IV2	
	≤ 3	≤ 7	≤ 3	≤ 7	≤ 3	≤ 7
$Post \times PCShare$	−0.118*** (0.029)	−0.161*** (0.035)	−0.124 (0.085)	−0.255** (0.110)	−0.015 (0.102)	−0.175 (0.131)
$Post \times PCShare \times n_b$	−0.018*** (0.005)	−0.036*** (0.005)	−0.037*** (0.011)	−0.039*** (0.009)	−0.038** (0.015)	−0.037*** (0.012)
$[n_b] \times Purchase$ FE	Yes	Yes	Yes	Yes	Yes	Yes
$Issuer \times Purchase$ FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Issuer cluster	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,733,881	3,043,035	1,733,881	3,043,035	1,733,881	3,043,035
First-stage F			25.6	45.5	17.5	34.6

Notes. $Post$ is the post-reform indicator $PostTCJA$ of equation (150). Standard errors are clustered by issuer (six-digit CUSIP), with 9,903 issuers at score ≤ 3 and 17,634 at score ≤ 7 . First-stage F is Kleibergen–Paap. Both samples omit Table 39’s rating and rating-by-maturity controls. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 47: P3: observed coupon-yield response to treasury par yields, by rating cut

	(1) continuous		(2) bin <5y vs ≥5y	
	score ≤ 3	score ≤ 7	score ≤ 3	score ≤ 7
$\Delta y_{nb,t}^{\text{Tsy}}$	0.346*** (0.012)	0.281*** (0.011)		
$\Delta y_{nb,t}^{\text{Tsy}} \times n_b$	0.007*** (0.002)	0.013*** (0.002)		
$\Delta y_{nb,t}^{\text{Tsy}} \times Long_b$			0.075*** (0.009)	0.113*** (0.006)
State cluster	Yes	Yes	Yes	Yes
Observations	952,874	2,023,637	952,874	2,023,637
Weights	Inverse pair count			

Notes. The estimates use Table 41’s first-differenced specification for coupon yields and par yields, weighted by inverse CUSIP pair counts and clustered by state, with $Long_b = \mathbf{1}\{n_b \geq 5\}$. Appendix C.11.1 maps the zero-coupon benchmark into the observed yield conventions on the main sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 48: P3 observed-response sensitivity grid, binned specification, by rating cut

cutoff	AA and above (score ≤ 3)					A– and above (score ≤ 7)				
	15d	30d	60d	90d	120d	15d	30d	60d	90d	120d
≥ 0.05	0.049***	0.061***	0.070***	0.074***	0.088***	0.078***	0.099***	0.098***	0.114***	0.121***
≥ 0.075	0.056***	0.063***	0.068***	0.074***	0.086***	0.080***	0.098***	0.096***	0.112***	0.118***
≥ 0.1	0.062***	0.066***	0.069***	0.075***	0.086***	0.088***	0.101***	0.097***	0.113***	0.119***
≥ 0.15	0.079***	0.077***	0.079***	0.083***	0.092***	0.095***	0.110***	0.105***	0.117***	0.123***
≥ 0.2	0.094***	0.090***	0.085***	0.088***	0.097***	0.108***	0.118***	0.111***	0.122***	0.128***
≥ 0.25	0.111***	0.100***	0.096***	0.096***	0.102***	0.120***	0.126***	0.119***	0.128***	0.134***

Notes. Cells report the observed interaction $\Delta y_{nb,t}^{\text{Tsy}} \times Long_b$ in the regression of coupon-yield changes on par-yield changes, with state-clustered standard errors. Rows give minimum absolute treasury moves in percentage points. Columns give maximum days between consecutive trades, as in Table 43. Every wider-sample coefficient is larger, with the largest proportional gains at the lowest cutoffs. Appendix C.11.1 reports the corresponding coupon-benchmark contrasts on the main sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Institutional detail and data

This appendix documents the tax rules and trading restrictions summarized in Section 3, separating statutory treatment from the assumptions of the quantitative model.

D.1 Investor-by-investor federal tax treatment since 2000

Investor i keeps the absolute after-tax shares $\xi_i^C = 1 - \tau_i^C$ of taxable and $\xi_i^M = 1 - \tau_i^M$ of municipal interest. Pricing depends on the relative after-tax share $\xi_i = \xi_i^C / \xi_i^M$ (Section 4.2). The exemption benefit $1 - \xi_i = (\tau_i^C - \tau_i^M) / (1 - \tau_i^M)$ equals τ_i^C only when municipal interest is fully exempt. Deduction rules for interest that finances bond holdings add a further margin, described for banks below and for other investors in Appendix D.6. Table 49 restates Table 2 for the 2018–2020 sample.

Households. Municipal interest is generally federally exempt and taxable bond interest is taxed as ordinary income, so a household’s federal after-tax shares are approximately $(\xi_i^M, \xi_i^C) = (1, 1 - \tau_{oi})$, with τ_{oi} its ordinary-income tax rate. Exempt interest is reported in both high- and lower-income brackets (Figure 1), so the exemption benefit varies within the household clientele.

Table 49: Investor heterogeneity in the municipal market after the 2017 Act

Investor type	Federal treatment	State treatment	Net exemption benefit
High-tax household	Interest exempt	Exempt on in-state bonds only	0.35–0.41
Middle-tax household	Interest exempt	Exempt on in-state bonds only	0.22–0.32
Low-tax household	Interest exempt	Exempt on in-state bonds only	0.10–0.12
P&C insurer	Exempt, less proration	Premium taxes, no distinction	0.17
Life insurer	Exempt, less proration	Premium taxes, no distinction	0.16
Bank	Exempt, funding cost disallowed	Varies by state	0.16–0.18
Other (pensions, etc.)	Varies	Varies	< 0.05

Notes. Household values are federal marginal tax rates in three groups (35% and above, 22% to 32%, below 22%), and the top value includes the 3.8% net investment income tax. The \$10,000 cap on the state and local tax deduction makes the home-state exemption more valuable to households whose cap binds. Institutions face the 21% corporate rate. P&C proration rises to 25% of exempt interest, giving $\xi = 0.79/0.9475$, and the life insurer’s policyholders’ share is set at 30%, giving $\xi = 0.79/0.937$. The bank value is the break-even implied tax of Table 2 at the 21% rate, with the 2018–2020 financing-tax cost of 0.117% a year and 2018–2020 taxable yields. The Act narrows the difference between the P&C and life-insurer exemption benefits.

Table 50 gives top federal rates. The 2001 phase-down, accelerated in 2003, cut the top ordinary rate to 35% in 2003. The rate rose to 39.6% in 2013 and fell to 37% in 2018. The top long-term capital-gains rate fell to 15% in 2003 and rose in 2013. Since 2013, the 3.8% net investment income tax (NIIT) has applied to taxable interest and many gains, but not to exempt interest. The calibration computes federal and state rates for each Survey of Consumer Finances (SCF) tax unit under each year’s law (Appendix B.7).

Table 50: Top federal tax rates relevant to municipal pricing, 2000–2020

Period	Top ordinary rate	+ NIIT (on taxable interest)	Top corporate rate
2000	39.6%	39.6%	35%
2001	39.1%	39.1%	35%
2002	38.6%	38.6%	35%
2003–2012	35.0%	35.0%	35%
2013–2017	39.6%	43.4%	35%
2018–2020	37.0%	40.8%	21%

Banks. With c and τ_B as in Section 6.2, interest expense that finances non-bank-qualified exempt bonds is generally nondeductible, so it costs c after tax, compared with $c(1 - \tau_B)$ if deductible, a penalty of $\tau_B c$. Section 265(b)(3) excepts bank-qualified bonds, but Section 291 cuts their deduction by 20%, leaving 80% and an after-tax cost of $c(1 - 0.8\tau_B)$. Interest expense is allocated to exempt bonds in proportion to adjusted tax basis.³⁰

The 15% disallowance under the Tax Equity and Fiscal Responsibility Act became 20% in 1984. The 1986 Act generally denied financial institutions’ interest deductions allocable to exempt obligations acquired after August 7, 1986. The ordinary bank-qualified exception covers designated small-issuer obligations under the \$10 million annual limit and other conditions, and temporary 2009–2010 rules expanded eligibility. In the model, each dollar of bank municipal holdings bears a financing-tax cost of τ_B times the measured ratio of interest expense to assets times $(1 - 0.8q_{BQ})$, where q_{BQ} is the bank-qualified share of those holdings. The aggregate calibration sets $q_{BQ} = 0$. Appendix C.8.1 derives how much of the deduction on designated bonds is capitalized in yields.

³⁰See the IRS, [Bank Qualified Bonds: Section 265](#), pp. 13-9 and 13-11. Special rules apply to certain 2009–2010 issues and grandfathered obligations.

Property and casualty insurers. P&C insurers pay the corporate rate τ_c on taxable bond interest. Proration cuts their losses-incurred deduction by a percentage of exempt interest, so that percentage of municipal interest is in effect taxed at τ_c . Before the 2017 Tax Cuts and Jobs Act, the percentage was 15%, giving an effective municipal rate of $0.15 \times 0.35 = 0.0525$. The Act replaced it with 5.25% divided by the highest corporate rate, or 25% at $\tau_c = 21\%$. The municipal after-tax share thus stays at 0.9475 while the taxable share rises from 0.65 to 0.79. The relative after-tax share rises from 0.686 to 0.834 and the exemption benefit falls from 31.4% to 16.6%. Cutting the corporate rate while keeping 15% proration would not reproduce the 2018 law.

The calibration applies the top corporate rate to tax-paying insurers. Before 2018, marginal corporate rates were 15–34% below \$100,000 and 34–39% from \$100,000 to \$18,333,333, including bubble rates that recapture the lower brackets’ benefit, and a flat 35% above. Since 2018 the rate is 21% at every taxable-income level.

Life insurers and other institutions. Section 812 divides a life insurer’s net investment income into company and policyholder shares, and exempt interest reduces its reserve deductions. Before 2018, a one-half policyholders’ share taxed at 35% gives an effective municipal rate of 17.5%, absolute after-tax shares of 0.825 on municipal and 0.65 on taxable interest and a relative after-tax share of 0.788. The post-2017 calibration’s 30% policyholders’ share and 21% corporate rate give an effective municipal rate of 6.3%, after-tax shares of 0.937 on municipal and 0.79 on taxable interest and a relative after-tax share of 0.843. Tax-exempt institutions and tax-deferred accounts gain little from the exemption.

D.2 Relative after-tax shares, absolute tax liabilities and fiscal scope

Relative after-tax shares govern the tax comparison in portfolio choice, while the receipts of Section 7 require absolute tax rates. For example, the pre-2018 P&C relative share of 0.6860 reflects a 5.25% tax on municipal interest and a 35% tax on taxable interest. Treating municipal interest as untaxed and taxing taxable interest at $1 - 0.6860$, or 31.4%, would preserve the relative share but misstate both receipts. Reforms change the absolute tax rates while keeping preferences and liability hedges in each investor’s baseline payoff units (Section 4.2).

Household tax assignments retain the dated federal marginal rates, state-tax approximations and Social Security taxation proxy of Appendix B.7. State receipts go to the investor’s state of residence. Before 2018, the federal tax per dollar of taxable interest equals total federal and state tax, one minus the combined after-tax share, less the state rate. This allocation is consistent with the multiplicative after-tax share and needs no new tax-unit calculation.

The fiscal ledger covers bond and taxable cash interest. Each investor’s cash balance is its funding less its bond holdings, and cash uses the same absolute tax rates as taxable bond interest. Fixed aggregate funding, debt supplies and the short rate imply unchanged aggregate cash interest. Changes in cash tax receipts therefore offset changes in investors’ after-tax cash income, preserving identity (29). Changes in the bank financing-tax cost are separate federal receipts, deducted from banks’ income without double-counting interest expense or issuers’ payments. Other liability flows remain outside the ledger. Financing costs sum yield times debt outstanding, $y_T D_T$, over maturities at $z = 0$, without resetting existing coupons, OID schedules or redemption payments. Averages for both reforms give each calendar year the same weight and divide mean dollar costs by mean receipts. Tables 52 and 53 report the individual periods.

For California, the \$100.8 million reallocation term in Table 52, Panel B counts both purchases and sales at post-repeal yields. Assigning every bond the pooled no-income-tax-state municipal yield at the same maturity, while retaining the taxable cash yield, gives a net portfolio effect of \$4.9 million. Actual taxable yields above this reference contribute \$96.8 million, and yield differences across municipal markets contribute $-\$0.9$ million. The increase therefore comes mainly from substitution into taxable bonds with higher gross yields. These are accounting components of the solved allocation, rather than separate policy experiments.

Portfolio substitution and interest receipts. After home-state repeal, both home- and out-of-state municipal interest are state-taxable, so switching issuers does not avoid tax. High-tax residents increase their riskless share by 0.6–2.7 percentage points across the seven states. Lower-tax residents increase their

Table 51: Portfolio allocation under the two tax reforms

Panel A. Home-state repeal: average resident allocations, 2000–2017			
	Home-state municipal	All risky	Riskless
California	20.33 → 4.35	30.02 → 31.56	69.98 → 68.44
New York	19.65 → 3.92	29.97 → 31.18	70.03 → 68.82
New Jersey	18.63 → 1.32	31.34 → 31.12	68.66 → 68.88
Massachusetts	18.96 → 1.55	29.95 → 30.99	70.05 → 69.01
Ohio	19.10 → 1.09	29.75 → 31.00	70.25 → 69.00
Pennsylvania	17.42 → 2.15	30.76 → 30.58	69.24 → 69.42
Minnesota	20.19 → 0.64	29.60 → 31.38	70.40 → 68.62
Panel B. Equal treatment: average investor allocations, 2000–2017			
	Risky maturity (years)	Share of long bonds (%)	Riskless
High-tax households	6.71 → 4.91	2.6 → 0.0	75.31 → 76.96
Lower-tax households	12.85 → 12.14	29.7 → 27.2	66.07 → 64.95
Banks	11.76 → 8.03	8.6 → 0.0	0.00 → 0.00
P&C insurers	9.93 → 6.83	5.3 → 1.7	21.70 → 0.65
Life insurers	11.81 → 12.07	14.6 → 14.0	0.38 → 0.68
Tax-neutral investors	7.73 → 8.53	39.2 → 57.1	27.90 → 29.56

Notes. Entries are baseline → reform values. Home-state, all-risky and riskless shares are percentages of each population’s modeled financial assets, including cash and excluding outside assets and liabilities. The share-of-long-bonds column instead reports each group’s percentage of the total long-bond market. Panel A aggregates resident households within each state and averages their portfolio shares across the six 2000–2017 economies. Each row is a separate repeal. Panel B averages each investor group’s allocations across the same economies under 99.9% equal treatment. Both panels weight types by their mass m_i and use rate state $z = 0$. Tax-neutral investors are the residual group of Table 3. Riskless holdings comprise cash and one-period municipal bonds (two years on the numerical grid). All longer bonds are risky. In Panel A, home-state municipal holdings overlap the other columns, and the last two columns sum to 100%. In Panel B, risky maturity is the value-weighted mean maturity in years of the group’s risky municipal and taxable bonds. The long-bond market comprises all municipal and taxable bonds maturing in more than ten years, and its shares sum to 100% across groups. Risky municipal and taxable bonds of the same maturity become close substitutes, so Panel B reports their combined maturity profile (Appendix B.2).

taxable-bond share by 5.1–14.6 points and reduce their riskless share. Aggregating the two groups, the riskless share rises in New Jersey and Pennsylvania and falls in the other five states (Table 51).

The high-tax group’s mean bond maturity rises by 0.4–0.9 years as it replaces riskless municipal bonds with cash, which is excluded from the bond-maturity average. Its risky bonds become slightly shorter. In California, mean bond maturity rises from 5.99 to 6.44 years, while the mean conditional on risky bonds falls from 7.05 to 6.52. The lower-tax group’s risky mean maturity there falls from 14.47 to 11.84 years. The added receipts thus accompany substitution across markets and tax groups, without a shift toward longer risky bonds.

Table 52: Home-state repeal: period financing ratios and the decomposition of costs and receipts

Panel A. Own-state / all-issuer financing costs per added state receipt dollar							
State	2000–02	2003–05	2006–08	2009–11	2012–14	2015–17	Average
CA	0.80 / 1.19	0.61 / 0.84	0.63 / 0.67	0.78 / 1.13	0.94 / 1.41	0.96 / 1.52	0.81 / 1.17
NY	0.74 / 1.15	0.64 / 0.89	0.59 / 0.49	0.78 / 1.11	0.94 / 1.38	0.92 / 1.41	0.78 / 1.08
NJ	0.82 / 1.35	0.79 / 1.23	0.80 / 1.11	0.83 / 1.18	1.00 / 1.49	0.98 / 1.46	0.88 / 1.30
MA	0.66 / 1.03	0.51 / 0.75	0.56 / 0.49	0.66 / 0.95	0.80 / 1.04	0.90 / 1.20	0.68 / 0.89
OH	0.73 / 1.17	0.57 / 0.83	0.60 / 0.57	0.70 / 1.04	0.88 / 1.19	0.99 / 1.39	0.73 / 0.99
PA	0.66 / 1.11	0.56 / 0.88	0.66 / 1.02	0.63 / 0.87	0.90 / 1.12	1.02 / 1.32	0.73 / 1.03
MN	0.69 / 1.07	0.52 / 0.74	0.53 / 0.50	0.68 / 0.95	0.86 / 1.22	0.83 / 1.16	0.68 / 0.93

Panel B. Average annual changes (\$million)							
State	Cost at fixed y^0	Change in y^0	Total cost	Tax at y^0	Yield changes	Reallo-cation	Total receipts
CA	798.7	17.2	815.9	902.5	1.0	100.8	1004.3
NY	455.4	6.6	462.0	526.0	0.1	65.2	591.4
NJ	138.4	0.2	138.6	141.8	-0.2	15.2	156.9
MA	76.1	0.1	76.2	98.9	-0.1	14.0	112.8
OH	75.1	0.1	75.1	90.1	-0.1	12.7	102.6
PA	59.8	-0.3	59.5	73.0	-0.1	9.0	81.9
MN	59.5	0.1	59.6	75.8	-0.1	11.3	87.0

Notes. Each state-period cell is a separate repeal from its dated baseline. All-issuer costs add other municipal and pooled taxable issuers to the state’s own issuers, and receipts are the repealing state’s added receipts. The average column divides six-period mean dollar costs by mean receipts. Panel B splits the average own-state costs and receipts around each economy’s pooled no-income-tax-state municipal yield y_T^0 . Cost at fixed y^0 is $\sum_T D_T y_T^0 (-\Delta \tau_T^{\text{impl}})$, with D_T the state’s debt and $\tau_T^{\text{impl}} = 1 - y_T^M / y_T^0$, and the remaining cost comes from the change in y^0 . Tax at y^0 applies each resident group’s state rate to its baseline home-state holdings at y_T^0 . Yield changes are the rest of the tax at new yields on those holdings. Reallocation is the net effect of residents’ portfolio changes, including cash. Components sum to totals, subject to rounding. All dollar calculations use continuous yields and holdings at $z = 0$ and include residents’ taxable cash interest.

Table 53: Equal treatment and a revenue-financed issuer subsidy across the six pre-2018 economies

Period	Thirty-year yield increase (bp)		Costs per receipt dollar		
	Municipal	Taxable	All issuers	Taxable issuers	Municipal after subsidy
2000–2002	199.8	38.2	1.434	0.282	0.152
2003–2005	136.9	9.3	1.125	0.089	0.036
2006–2008	146.5	11.4	1.077	0.073	0.004
2009–2011	152.7	11.2	1.113	0.043	0.070
2012–2014	178.9	33.9	1.245	0.170	0.075
2015–2017	208.6	34.7	1.242	0.174	0.068
Average 2000–2017	170.6	23.1	1.196	0.133	0.063

Notes. Each row removes 99.9% of the municipal–taxable tax distinction from its own calibrated baseline. Yield changes average each economy’s 36 monthly rate states. Fiscal ratios use continuous yields and holdings at $z = 0$, with bond and taxable cash interest in receipts. The last column subtracts a subsidy equal to all added federal and state receipts from municipal issuer costs. The other columns describe the reform before subsidy. The subsidy leaves gross yields and portfolios unchanged because supplies are fixed. The average row divides calendar-year-weighted mean costs by mean receipts.

Costs and receipts around the comparison yield. Panel B of Table 52 splits each state’s own-state cost and receipts into the fixed-yield comparison of Section 7.2 and adjustments for yield changes and reallocation. With D_T the state’s debt, y_T^0 the pooled no-income-tax-state municipal yield, $\tau_T^{\text{impl}} = 1 - y_T^M / y_T^0$ and primes marking post-repeal values, the cost increase is exactly

$$\Delta C = \sum_T D_T y_T^0 (\tau_T^{\text{impl}} - \tau_T^{\text{impl}'}) + \sum_T D_T (1 - \tau_T^{\text{impl}'}) (y_T^{0'} - y_T^0).$$

The first term holds y^0 fixed, and the second comes from the change in y^0 . The tax at y^0 is $\sum_i \tau_i^{\text{st}} \sum_T H_{iT} y_T^0$, where H_{iT} is resident group i ’s baseline home-state holding and τ_i^{st} its state rate. Beyond this fixed-yield tax calculation, almost all of the additional receipts come from residents’ reallocation. At the comparison yield, own-state cost per dollar of tax ranges from 0.770 in Massachusetts to 0.976 in New Jersey, and reallocation raises receipts by another 10.8% to 14.9% of that tax, so costs stay below receipts in every state. California’s additional annual receipts of \$1.004 billion comprise \$0.227 billion from home-state bonds and net increases of \$0.300 billion from other states’ bonds, \$0.368 billion from taxable bonds and \$0.110 billion from cash. Interest on these last three assets was already taxable. The increases are relative to baseline receipts and arise from changes in residents’ holdings and the interest they earn, not from a new tax on those assets.

Differences across states. At longer maturities, lower pass-through reduces the cost per resident-held dollar, but debt held by nonresidents and institutions adds cost without adding tax. New Jersey’s residents hold only 49% of its 12–20 year debt, where cost exceeds the tax at the comparison yield by a factor of 1.38, giving New Jersey the highest overall ratio. Minnesota’s residents hold 80% of its 22–30 year debt, against 67% in New Jersey, and its lower-tax residents face a higher rate, so cost on those bonds is 0.59 per dollar of tax, against 0.98 in New Jersey. Household composition, outside exposures and resident funding per local debt are common across states within each period, so these differences come from tax schedules and market sizes.

D.3 State tax exemption and who receives it

Resident households. For a household resident in state s , with state tax rate τ_s^{st} , the baseline normalizes the home-state municipal after-tax share to one and sets the out-of-state municipal share to $1 - \tau_s^{\text{st}}$, with no additional federal deduction credit (varied in Appendix D.4). Through 2017, the after-tax share on taxable interest multiplies the federal relative after-tax share ξ^{fed} by $1 - \tau_s^{\text{st}}$, tax unit by tax unit. The post-2017 mapping treats the cap on the state and local tax deduction as binding, so the combined share on taxable interest is $\xi^{\text{fed}} - \tau_s^{\text{st}}$, as in equation (142).

Funds and institutions. Municipal funds generally pass the federal exemption through to shareholders as exempt-interest dividends. State treatment depends on the shareholder’s residence and the share of interest from that state’s bonds. A single-state fund preserves the home-state exemption for resident shareholders, while a national fund passes it through only on the eligible share, attenuating the household state incentive. Among institutions, tax-deferred accounts and many pensions owe no current state tax on interest. Insurers commonly pay premium taxes, which do not distinguish municipal from taxable interest. Banks’ and other corporations’ state taxes depend on the tax base and business apportionment. P1 exploits the household exemption that depends jointly on issuer and holder residence (Section 6.1). The model assigns this home-state wedge to households and uses federal proration and financing deductions for insurers and banks.

D.4 Federal deduction credit for out-of-state municipal interest

Before 2018, itemizers could deduct state income taxes on federally exempt interest against ordinary federal income tax. The deduction’s full benefit is not available to every household: taking the standard deduction,

paying the alternative minimum tax or choosing a larger sales-tax deduction can eliminate the incremental saving.³¹

To test sensitivity, I grant every tax unit the full credit, ignoring deduction eligibility, at the ordinary-income component of its baseline federal marginal rate, which excludes NIIT, and form products tax unit by tax unit before aggregation. I re-solve the six pre-2018 economies, holding all other inputs and calibrated coefficients fixed. The post-2017 binding-cap convention is unchanged.

For the 2015–2017 California high-tax group, the effective after-tax shares with the credit are 0.498 for taxable interest and 0.919 for out-of-state municipal interest. At a riskless taxable yield of 5%, a municipal yield of 2.71% equates the two after-tax returns, compared with 2.87% for the high-tax group in states without an income tax. The taxable alternative already incorporates the baseline deduction treatment, and the added municipal credit lowers the yield needed to match it. Residents of high-tax states can consequently accept lower out-of-state yields than residents of states without an income tax, compressing cross-state yield differences, and so state-tax pass-through, even at the riskless maturity.

Using the main sample’s 2000–2020 weights and corporate normalization, the thirty-year model gap falls only 0.23 basis points, from 97.39 to 97.16. The model’s P1 pass-through, averaged over the six blocks with its original taxable-yield normalization, falls from 1.001 to 0.904 at the two-year riskless maturity and from 0.458 to 0.412 at twenty-five years (interpolated on the two-year grid). The credit lowers the level of pass-through but preserves its declining profile and the large long-maturity gap.

D.5 Accrual, price changes and secondary-market taxation

Equation (32) applies the investor’s interest-income after-tax share to the whole model return $\mathcal{R} = \mu + u$, approximating the federal rules below for accrual, basis and gains.³²

Gains and basis. For individuals, federal law separates interest accrual from gains realized at disposition (Sections 171, 1272–1278 and 1288). For a bond with face value 100, annual coupon payment c in dollars, original price A_0 and annual effective issue yield y_0 , the adjusted issue price satisfies

$$A_t = (1 + y_0)A_{t-1} - c, \quad t = 1, \dots, n, \quad A_n = 100. \quad (155)$$

Here A_t is a dollar price, unrelated to the rate loading A_t^r . A discount, $A_0 < 100$, is OID, and a premium, $A_0 > 100$, is original-issue premium (OIP). Annual interest $y_0 A_{t-1} = c + (A_t - A_{t-1})$, including the noncash accrual, is taxed at the ordinary-income rate τ_{oi} :

$$\text{tax}_t = \begin{cases} \tau_{oi} y_0 A_{t-1}, & \text{taxable bond,} \\ 0, & \text{tax-exempt bond.} \end{cases} \quad (156)$$

The basis change $A_t - A_{t-1}$ prevents taxing accrued interest again at redemption. OID raises basis, including for zero-coupon bonds, and OIP lowers it. Amortizing premium on a taxable bond requires a Section 171 election, which the capital-gains adjustment assumes, and amortization on an exempt bond gives no deduction. At maturity, basis equals face, so there is no capital gain. For taxable bonds, de minimis OID below a quarter point per full year can be treated as zero, with the discount recognized as capital gain at maturity. This OID exception generally does not apply to tax-exempt obligations.³³

³¹IRS Publication 550 (2017), “State income taxes”; 2017 Instructions for Form 6251, line 3; and 2017 Instructions for Schedule A, line 5. 2016 Instructions for Form 8960, line 9b, restrict NIIT deductions to taxes attributable to net investment income, which excludes exempt interest.

³²The after-tax share multiplies both the log of the expected gross return, $\log \mathbb{E}[1 + R]$, and the centered log-return innovation. Taxing a simple return R gives the after-tax gross return $1 + \xi R$, with $\log \log(1 + \xi R)$. For a deterministic two-year simple return of 5.4 percent and $\xi = 0.62$, this log differs from $\xi \log(1 + R)$ by 3.2 basis points per two-year period (1.6 a year) and zero for a fully exempt leg. Appendix A.1 reports the effect of taxing simple returns exactly in the solved model. The return convention matches the empirical implied tax and the after-tax yield mappings of Appendix C.1.4.

³³IRS Publication 550 (2017), “Original Issue Discount (OID)”; Publication 1212, “De minimis rule.” The market-discount threshold in (125) applies separately.

A sale at price P_h at date h realizes the gain $G_h = P_h - \text{basis}_h$, long-term after one year and then taxed at the capital-gains rate τ_{cg} . A secondary buyer's basis starts at the purchase price P . If P exceeds face, the premium amortizes at the buyer's yield, mandatorily for exempt bonds and by election for taxable bonds. On an OID bond bought at date s for $A_s < P \leq 100$, the acquisition premium $P - A_s$ reduces the remaining OID by the fraction $(P - A_s)/(100 - A_s)$, leaving basis 100 at maturity.

Market discount is the shortfall of the purchase price below A_s for OID bonds or below face for par and premium bonds. It accrues ratably, or at constant yield by election. Absent annual inclusion, accrued market discount D_h is taxed as ordinary income at sale or redemption, up to any positive gain. This rule also covers exempt bonds bought after April 30, 1993. Below the de minimis threshold (125), the discount is treated as zero and the gain is capital income. Thus

$$\text{tax at sale} = \tau_{oi} \min\{\max(G_h, 0), D_h\} + \tau_{cg}[G_h - \min\{\max(G_h, 0), D_h\}], \quad G_h = P_h - \text{basis}_h. \quad (157)$$

Individual capital losses offset gains and up to \$3,000 of ordinary income a year, with carryforwards (Sections 1211–1212). The capital-gains adjustment assumes full offsets at the capital-gains rate, as in (126). Corporations pay their ordinary rate on gains but offset capital losses only against gains. Banks treat gains and losses on debt as ordinary income (Section 582(c)). The model's institutional after-tax shares apply one rate to the full return. Appendix C.1.4 applies the household accrual and basis mappings above, with its stated federal tax rates and loss offsets.

D.6 Institutional foundations of the no-short-sale constraint

The model rules out short sales and borrowing (Sections 3.5 and 4.1). Tax asymmetries, financing rules and residual risk limit short arbitrage, and the nine cases below show how they sustain the different post-hedge duration valuations required by (167).

Household tax rules matter because high-income households hold much of the market. In custodial data, municipal bonds are wealthy households' second-largest liquid risky asset, and their portfolio share rises with wealth (Gabaix et al., 2026). In tax years 2008, 2016 and 2019, returns with adjusted gross income (AGI) of \$1 million or more, about 0.3 percent of all returns, reported 26–33 percent of tax-exempt interest.³⁴

Define the after-tax coupon gap $g = c_M - (1 - \tau)c_C$, where c_M and c_C are municipal and corporate par coupons of the same maturity and τ is the investor's marginal tax rate. This coupon measure differs from the 118-basis-point thirty-year puzzle in zero-coupon yields (Section 5.4) and from the model's one-period return gap g_T . Carry is computed for equal principal in each leg, before fees, margin and collateral costs, and the legs' duration and price risk need not match.

(i) Financing municipals with debt or a corporate short. Let c_F be the interest on a dollar borrowed to buy a dollar of municipal principal. If that interest were deductible, carry would be $c_M - (1 - \tau)c_F$. Section 265(a)(2) denies deductions for interest on debt incurred or continued to purchase or carry tax-exempt obligations, so carry becomes

$$c_M - c_F = \{c_M - (1 - \tau)c_F\} - \tau c_F. \quad (158)$$

At after-tax parity, $c_M = (1 - \tau)c_F$, the denial turns zero carry into a loss of τc_F . Financing the position with a corporate short does not restore the tax shield, because Section 265(a)(5) treats the short seller's compensatory coupon payments c_C as interest under Section 265(a)(2), leaving carry $c_M - c_C$. This carry is negative at every maturity because the sample-average municipal zero curve, and hence the implied par yield, lies below its corporate counterpart. At thirty years with annual compounding, the municipal and corporate par yields are 4.498% and 5.443%.³⁵ Municipal fund disclosures tell investors that interest on debt allocable to exempt-interest dividends is nondeductible.³⁶ While the deduction was available, depository institutions

³⁴IRS Statistics of Income, Individual Income Tax Returns, Table 1.4, all returns by size of adjusted gross income, tax years 2008, 2016 and 2019. Tax-exempt interest includes exempt-interest dividends from funds.

³⁵26 U.S.C. §265(a)(2), (5) and (b), 2018 U.S. Code. Section 265(a)(5)(B) excepts specified short-sale payments when the taxpayer posts cash collateral that earns no material return. That exception does not remove the opportunity cost of the collateral. See also Hines and Horton (2024) on banks' financing-tax restrictions.

³⁶BlackRock Municipal Bond Fund, Inc. (2025), Part II, "Dividends and Taxes," p. II-131; iShares Trust (2026), "Taxes," p. 210; Nuveen Municipal Trust (2026), "Tax Matters," p. S-80.

financed municipal holdings at scale, holding 51.8% of municipal bonds in 1972 and \$235.1 billion (27.3%) in 1985. After the 1986 denial (Appendix D.1), their holdings fell to \$120.4 billion in 1990 and \$95.4 billion (7.5%) in 1995, while municipal debt outstanding grew from \$862 billion in 1985 to \$1.27 trillion in 1995.³⁷

(ii) Pairing corporate positions. If the proceeds of a corporate short buy a corporate bond, the investor’s own capital finances the municipal bond, and the remaining trade is the hedge in (iii). If they finance the municipal bond, the denial in (i) applies. Revenue Procedure 72-18 tests the purpose of borrowing through tracing of the proceeds, exempt collateral or the facts and circumstances. Its presumption that business borrowing is not incurred to carry exempt bonds lapses when the borrowing exceeds business needs. If the positions form an offsetting straddle, Section 263(g) also capitalizes interest and carrying charges, net of the positions’ taxable income, deferring the deductions.³⁸

(iii) Hedging duration with a corporate short. An unlevered hedge of a municipal bond’s duration shorts a long-maturity corporate bond and buys a short-maturity corporate bond with the proceeds. Let c_ℓ and c_s be the long and short corporate coupons, so that $g = c_M - (1 - \tau)c_\ell$, and let $\delta \in [0, 1]$ be the currently deductible fraction of the short payment. With equal principal in each leg, annual carry relative to holding the short corporate bond alone is

$$\Delta = \{c_M + (1 - \tau)c_s - (1 - \delta\tau)c_\ell\} - (1 - \tau)c_s = g - (1 - \delta)\tau c_\ell. \quad (159)$$

Denial under Section 265 gives $\delta = 0$ and $\Delta = c_M - c_\ell$, negative at every maturity and a loss of 95 basis points a year at thirty years. Section 263(g) capitalization and Section 163(d) carryforwards defer deductions, so current carry approaches $c_M - c_\ell$ and later benefits carry a timing and liquidity cost.

Even where neither Section 265 nor Section 263(g) applies, Section 163(d) lets individuals deduct investment interest only if they itemize and up to net investment income, carrying the excess forward. Coupons on the short-maturity corporate bond add to this deduction capacity, while wages and exempt coupons do not. With principal q on both corporate legs, coupon difference $k = c_\ell - c_s > 0$, unused deduction capacity $C \geq 0$ and no prior carryforwards, the hedge’s current cash flow is

$$\Pi(q) = -(1 - \tau)qk - \tau[qk - C]_+, \quad (160)$$

where $[x]_+ = \max\{x, 0\}$, so $\delta < 1$ once qk exceeds C .³⁹

Even with $\delta = 1$, the hedge removes only common rate risk. Bond-specific risk (Appendix A.1) and credit, liquidity and call differences (Appendix C) remain. Shorting requires available bonds and incurs fees, recall risk and margin demands against an illiquid callable long position.⁴⁰ Fees for borrowing corporate bonds are 10–20 bp in the major-lender sample of Asquith et al. (2013).

(iv) Hedging with swaps. A pay-fixed, receive-floating swap hedges duration without borrowing or shorting a bond, but household access is restricted. A person that is not an eligible contract participant may trade a swap only on, or subject to the rules of, a designated contract market. An individual qualifies only with more than \$10 million invested on a discretionary basis, or more than \$5 million when the swap

³⁷Board of Governors of the Federal Reserve System, Financial Accounts of the United States (Z.1), release of September 8, 2023, Table L.212, year-end levels. Depository institutions comprise U.S.-chartered depository institutions, foreign banking offices in the U.S. and banks in U.S.-affiliated areas.

³⁸Rev. Proc. 72-18, 1972-1 C.B. 740, §§3.01–3.04 and 6.01, as summarized in [Revenue Ruling 2004-47](#) and in IRS [Field Service Advice 200129009](#), which is nonprecedential. On straddles, see 26 U.S.C. §263(g)(1)–(2) and §1092(c)(1)–(2). Positions are offsetting when holding one substantially diminishes the risk of loss from holding the other.

³⁹26 U.S.C. §163(d)(1)–(3), which includes otherwise deductible short-sale expenses in investment interest; IRS [Publication 550 \(2016\)](#), Chapter 3, “Interest Expenses” and “Interest and carrying charges on straddles”; IRS [Form 4952 and instructions \(2016\)](#). Net capital gain and qualified dividends count as investment income only through an election that relinquishes their preferential rates to that extent.

⁴⁰[FINRA Rule 4210](#), paragraph (e)(2), sets margin requirements for bond shorts.

manages the risk of an asset owned or a liability incurred.⁴¹ Ordinary swap payments are not automatically interest on indebtedness, so Section 265(a)(2) alone does not deny their deduction. A household's Section 212 investment expenses attributable to exempt interest fall under Section 265(a)(1).⁴² Independently of that classification, let X be a net taxable amount and $d \in [0, 1]$ the denied fraction of its loss tax shield, so that its after-tax value is

$$A_d(X) = (1 - \tau)X - d\tau X^-, \quad X^- := \max\{-X, 0\}. \quad (161)$$

Symmetric offsets give $d = 0$. Taxing gains and denying losses gives $d = 1$, which lowers expected proceeds by $\tau \mathbb{E}[X^-]$.

Netting gives the swap a tax advantage when the corporate short payment is disallowed. With matched periodic receipts $F \geq 0$ and payments $K \geq 0$, the bond trade's after-tax cash is $A_C = (1 - \tau)F - K$ when the corporate short payment is wholly nondeductible, and the swap's is $A_S = F - K - \tau[F - K]_+$, under the rule that taxes net gains and denies losses. Hence

$$A_S - A_C = \tau \min\{F, K\} \geq 0, \quad (162)$$

If the corporate payment is fully deductible, the two coincide for $F \geq K$. For $F < K$, the corporate trade benefits from a loss offset that the swap calculation denies. Additional deduction capacity therefore favors the corporate short. A pay-fixed swap has the rate exposure of a short fixed-rate bond and a long floating-rate bond, without the corporate short's credit and borrowing exposures but with its own collateral and counterparty exposure. It weakly dominates a corporate hedge in the mean-variance objective whenever the latter delivers the same exposure with only an additional nonnegative cost and uncompensated risk orthogonal to the existing portfolio and the swap.⁴³ An outside dealer can also absorb swap risk without trading domestic bonds, whereas physical shorts change the gross long positions required to clear those markets.

Because the wealthiest households can meet the eligibility thresholds, the equilibrium test of Appendix D.6.1 grants every household a thirty-year pay-fixed swap from an outside dealer with no fee, margin or notional limit and taxes swap gains while denying losses ($d = 1$). Physical bond shorts remain excluded, institutions' hedging is unchanged and both bond markets are re-solved. The model's thirty-year gap for 2015–2017 falls only from 104.3 to 94.8 basis points, retaining 90.9% of its baseline value.

(v) Treasury futures. Regulated futures avoid the swap eligibility rules but require cash margin and annual marking to market, with gains and losses treated as 60% long-term and 40% short-term. Individuals offset these losses fully against capital gains, deduct at most \$3,000 against ordinary income and carry the excess forward. Net losses on these Section 1256 contracts may be carried back three years against net Section 1256 gains. Section 1092 defers losses to the extent of unrecognized gains on offsetting bonds. Gains can be taxed before losses yield tax benefits, a timing and liquidity cost equivalent to partial loss denial in present value, $0 \leq d < 1$ in (161).⁴⁴

(vi) Shorting municipals. The substitute payment a short seller makes for the exempt coupon loses its exemption, so a lender with tax rate τ_L who requires an after-tax payment equal to the coupon c_M must receive L satisfying

$$(1 - \tau_L)L = c_M, \quad L - c_M = \frac{\tau_L}{1 - \tau_L}c_M. \quad (163)$$

⁴¹Commodity Exchange Act §2(e) and §1a(18)(A)(xi), 7 U.S.C. §§2(e), 1a(18)(A)(xi); see also the [CFTC's 2012 final-rule summary](#).

⁴²Treasury Regulations §1.446-3 and §1.265-1(a)(2), (c). IRS Field Service Advice 200037034 treats the ordinary swap payments before it as non-interest in the absence of a loan. The advice concerns a corporate taxpayer and is nonprecedential.

⁴³In common payoff units, let the corporate hedge be $B = \alpha S - \varsigma + \varepsilon$, where S is the swap payoff, $\alpha > 0$, $\varsigma \geq 0$, $\mathbb{E}\varepsilon = 0$, and $\text{Cov}(\varepsilon, W) = \text{Cov}(\varepsilon, S) = 0$ for the investor's existing risk index W . With $\text{CE}(Y) = \mathbb{E}Y - \gamma \text{Var}(Y)/2$, one has $\text{CE}(W + qB) - \text{CE}(W + q\alpha S) = -q\varsigma - \gamma q^2 \text{Var}(\varepsilon)/2 \leq 0$ for $q \geq 0$, provided the replacement is feasible at no greater cost of financing the position.

⁴⁴26 U.S.C. §1256(a)(1), (a)(3), (b) and (g)(1); §1211(b); §1212(b)–(c); §1092(a)(1)(A). Section 1256(a)(4) removes Sections 1092 and 263(g) only for straddles consisting entirely of Section 1256 contracts.

A short also requires a willing lender of the specific issue and satisfaction of the contract's collateral terms.⁴⁵ For the reverse trade, which shorts a municipal bond and buys an equal-principal corporate bond, let τ_S be the short seller's tax rate and δ the currently deductible fraction of the substitute payment, as in (iii). Its coupon carry under this lender compensation is

$$\pi_{M \rightarrow C} = (1 - \tau_S)c_C - (1 - \delta\tau_S)\frac{c_M}{1 - \tau_L}. \quad (164)$$

Even with full deduction, $\delta = 1$, borrower and lender rates equal to τ give $\pi_{M \rightarrow C} = (1 - \tau)c_C - c_M$. This carry is nonpositive when $c_M \geq (1 - \tau)c_C$, and the carry $c_M - c_C$ of (i) is nonpositive when $c_M \leq c_C$, so both coupon trades have nonpositive carry throughout the band

$$(1 - \tau)c_C \leq c_M \leq c_C, \quad (165)$$

and long-end municipal coupons lie in it: above the household after-tax benchmark and, on the sample-average curves, below the corporate coupon. A municipal curve trade, short long-maturity and long short-maturity municipal bonds, pays the gross-up in (163) on its short leg. To the extent its proceeds buy the short-maturity municipal bond, that payment is nondeductible as in (i).

(vii) Tax-exempt floaters. A tender-option-bond partnership divides a long municipal bond into variable-rate equity and an inverse interest that receives the residual. Exempt income passes through to both classes, so the structure needs no interest deduction. The inverse holder receives the municipal curve's slope and bears its duration, so capturing g still requires one of the hedges in (iii)–(v). Demand for the floaters limits the scale of the strategy.⁴⁶

(viii) The de minimis presumption. For insubstantial holdings, Revenue Procedure 72-18 ordinarily presumes, absent tracing or exempt collateral, that borrowing is not incurred to carry exempt bonds. A corporation qualifies only up to 2% of average business assets. Direct evidence overrides the presumption, and larger holdings face the facts-and-circumstances test. Sections 163(d), 263(g), 1092 and 1211(b) still constrain deductions and losses.⁴⁷

(ix) Other entities. Financial institutions outside the bank-qualified exception face the pro rata disallowance of Appendix D.1, without tracing. Revenue Procedure 72-18 applies a similar ratio to dealers' business borrowing that cannot be traced. Insurer proration reduces the exemption regardless of financing, through P&C insurers' loss deductions (Section 832(b)(5)(B)) and life insurers' reserve deductions (Sections 807 and 812). Partnerships pass the exempt character of municipal interest through to partners under Section 702(b), but their interest expense remains subject to Section 265(a)(2) under Section 703(a). Over 2000–2020, dealers held on average 0.85% of municipal bonds, with a maximum of 1.94%, and hedge funds, reported from 2012, held at most 0.36%.⁴⁸

The model constraint. With a concave objective and linear shorting costs, a zero short position is optimal when the risk-adjusted benefit of the first dollar shorted does not exceed the tax, borrowing and financing cost. The quantitative model imposes this corner, ruling out unconstrained after-tax arbitrage: bond positions are nonnegative and financed from funding W_i , and banks bear the measured financing-tax cost. The measured costs do not establish the corner for every investor, and some hedges remain profitable.

Residual risk and participation after hedging. The mechanism requires only that valuations of additional duration differ after available hedges, which allows continued derivative trading. To isolate the

⁴⁵IRS Publication 550 (2016), "Reporting Substitute Payments"; Treasury Regulation §1.6045-2(a)(4).

⁴⁶Revenue Procedure 2003-84, Section 2; 26 U.S.C. §702(b).

⁴⁷Rev. Proc. 72-18, §3.05, as summarized in Revenue Ruling 2004-47 and Field Service Advice 200129009.

⁴⁸Rev. Proc. 72-18, §§5.03 and 7, as summarized in Revenue Ruling 2004-47; 26 U.S.C. §265(b)(1)–(2), §832(b)(5)(B)–(C), §807(a)–(b) and §812. Holdings: Financial Accounts of the United States (Z.1), release of September 8, 2023, Tables L.212 and B.101.f, quarterly levels.

rate-hedging tradeoff, suppose the shorting constraint on cash is slack and bond-specific risk is zero. Let a hedge remove q_i units of rate exposure at an expected charge $f_i q_i$ and a differentiable convex additional cost $\psi_i(q_i)$, all in the investor's normalized payoff units. With residual exposure $\widehat{D}_i^{\text{hedge}} = \widehat{D}_i - q_i$, where $\widehat{D}_i$ is the net rate exposure (10) before hedging, the hedge-dependent part of the objective is $-f_i q_i - \psi_i(q_i) - \gamma_i \sigma_r^2 (\widehat{D}_i^{\text{hedge}})^2 / 2$, so an interior hedge choice satisfies

$$\lambda_i := \gamma_i \sigma_r^2 \widehat{D}_i^{\text{hedge}} = f_i + \psi'_i(q_i). \quad (166)$$

Here λ_i is the unit cost of rate risk at the post-hedge exposure. A binding hedge limit replaces the equality by the appropriate one-sided condition. At a household-only municipal allocation, with hedges and other holdings optimized, the lower-tax investor's municipal slack, its required return on the T -period municipal bond minus the offered return, is

$$\nu_{I,T}^{M,\text{hh}} = (\xi_I - \xi_H)r - A_{T-1}^M(\lambda_H^{\text{hh}} - \lambda_I^{\text{hh}}). \quad (167)$$

This is the participation comparison (59) applied to residual, post-hedge risk costs. Differences in liabilities sustain unequal duration valuations through unequal hedge access, binding limits or hedging costs that rise with the hedge, even when both investors hedge.

If unlimited hedging in either direction spans rate risk, $\psi_i = 0$ and both investors face a common normalized charge f , then $\lambda_H = \lambda_I = f$. With equal unit costs of rate risk, the slack reduces to the positive tax difference $(\xi_I - \xi_H)r$, which excludes lower-tax municipal holdings. This conclusion requires a common after-tax hedging opportunity; zero dealer fees alone do not suffice. Observed insurer holdings (Figure 25) therefore require a departure from this benchmark, which any of the documented frictions can sustain. The calibrated outside exposures are not measured directly net of derivative positions.

Insurers already hedge: at year-end 2015, interest-rate swaps accounted for 64% of the notional value of U.S. insurers' interest-rate hedges, and derivative users held 87% of life-insurer assets but 19% of P&C assets.⁴⁹

D.6.1 Household swaps: incentives and equilibrium

The model gives households no rate hedge, and off-exchange swaps face eligibility restrictions (Section 3.5). Using the pay-fixed swap of case (iv) above, I first evaluate the incentive to add a swap at fixed baseline bond portfolios, the 2015–2017 mean rate state and all 630 household types (14 types in each of 45 markets). I then let households choose swaps and bonds jointly and re-solve both bond markets.

Quotes, pricing and taxes. Swap quotes come from the 95 dates with a two-year rate within 10 bp of the model's mean rate, 3.0206% a year, all in 2001–2008, outside the 2015–2017 block. Fixed rates at ten, twenty and thirty years are 4.8217%, 5.0976% and 5.3736%, and the twenty-year rate is interpolated.⁵⁰

Floating payments are quarterly and fixed payments semiannual, and liquidation after two years uses an approximate calendar. One discount and forward curve with Gaussian pricing drifts fits quotes and liquidation values. Quarterly physical dynamics preserve the calibrated two-year persistence and variance. Annual net gains and the second-year liquidation payoff are taxed as ordinary income. Losses are denied in the baseline and offset symmetrically in a sensitivity check. Municipal funds likewise distribute swap gains as taxable ordinary income or capital gains and allocate net payments pro rata against taxable and exempt income.⁵¹ Segregated swap cash earns taxable interest when positive and pays nondeductible interest when negative. Existing cash is held fixed in this initial fixed-portfolio calculation.

⁴⁹NAIC, "Update on the Insurance Industry's Use of Derivatives and Exposure Trends," year-end 2015, Tables 1 and 8 and the accompanying discussion. Notional shares measure instrument use and are silent on how much liability risk is eliminated.

⁵⁰Federal Reserve H.15 [DSWP10](#) and [DSWP30](#) are mid-market fixed rates against three-month London Interbank Offered Rate (LIBOR) and differ from household quotes.

⁵¹[BlackRock Municipal Bond Fund, Inc. \(2025\)](#), Part II, pp. II-20, II-126 and II-132; [iShares Trust \(2026\)](#), "Taxes," p. 213.

Objective and tax units. Let $\widetilde{W}_i$ be the calibrated log-return risk index, including outside exposure $-h_i z'$, and M_i^0 its fixed baseline mean, where $h_i = -\ell_i$ is the outside duration of Appendix B.7 (positive values add rate risk). The absolute home-municipal after-tax share $\xi_{i,\text{abs}}^M$ is ξ_i^M of Appendix D.1 on home-state bonds; multiplied by the relative after-tax share ξ_i , it gives the swap's absolute after-tax share. High- and lower-tax home-municipal after-tax shares are 0.999905 and 0.990429, and swap tax rates span 21.49%–50.25%. After-tax cash A_i per unit notional (unrelated to the rate loading A_T^a) changes normalized wealth by $qA_i/\xi_{i,\text{abs}}^M$. The objective and the score b_i , its annualized slope at zero notional in basis points, are

$$\mathcal{W}_i(q) = \xi_{i,\text{abs}}^M \left[M_i^0 + \frac{q\mathbb{E}A_i}{\xi_{i,\text{abs}}^M} - \frac{\gamma_i}{2} \text{Var} \left(\widetilde{W}_i + \frac{qA_i}{\xi_{i,\text{abs}}^M} \right) \right], \quad (168)$$

$$b_i = \frac{10^4}{2} \mathcal{W}'_i(0+) = \frac{10^4}{2} \left[\mathbb{E}A_i - \gamma_i \text{Cov}(\widetilde{W}_i, A_i) \right]. \quad (169)$$

The score compares the swap's expected after-tax cash payoff with its contribution to portfolio risk. A positive score means that a small pay-fixed position improves the objective at the fixed bond portfolio. It uses the model's log-return approximation. Independent bond-specific shocks remain in baseline risk but have zero covariance with a rate-only swap.

Table 54: Fixed-portfolio swap benefits with gains taxed and losses denied.

Household group	Years	Cash	Risk benefit	Total	Median	Positive
High tax	10	−408.6	852.8	444.3	−14.4	135/315
High tax	20	−473.5	996.0	522.6	−13.1	135/315
High tax	30	−473.9	988.2	514.2	−17.2	135/315
Lower tax	10	−387.9	675.6	287.7	−28.8	135/315
Lower tax	20	−450.0	787.6	337.7	−31.5	135/315
Lower tax	30	−449.5	783.3	333.9	−33.3	135/315

Notes. Annualized scores b_i of equation (169), in basis points per unit notional. Years is the swap tenor. Cash and risk benefit are the score's two terms and total is their sum, each averaged within the group with household-mass weights, which equal portfolio-asset weights because budgets are identical. The median and the count of positive scores weight types equally. Entries are rounded separately.

Results. At every tenor (Table 54), 360 types have negative and 270 positive scores. The positive types, the top three outside-duration bins of both groups, hold 42.9% of household portfolio assets and 9.8% of household municipal holdings. Individual scores reach 3,100 bp, but a finite trade also bears the quadratic term $-\gamma_i q^2 \text{Var}(A_i)/(2\xi_{i,\text{abs}}^M)$. Most types have negative scores, giving a negative median. The larger gains among the positive-score types make the mean weighted by portfolio assets positive.

Relative to symmetric offsets, loss denial lowers expected swap cash for every type and, through the covariance term, lowers the score for 450 types and raises it for 180; on net it lowers 585 of the 630 scores. Doubling pricing resolution and draws preserves signs (maximum change 2.57 bp), as does using terminal dollar payoffs and their covariance. Widening the quote window to 25 bp gives 285, 302 and 275 positive scores at ten, twenty and thirty years, some near zero.

Equilibrium with household hedge access. With the 2015–2017 calibration fixed, households jointly choose bonds and nonnegative thirty-year pay-fixed swap notionals. Outside dealers quote swaps without trading the modeled bonds. A new swap has zero upfront value at the current rate state's par rate, and liquidation uses the same pricing rule and the historical quote family for the contract's age. Both markets' bond prices and all portfolios adjust across rate states, retaining nonnegative bonds and cash, gains taxation, loss denial and segregated interim financing.

Swap access reduces the thirty-year gap by 9.48 bp, or 9.1% (Table 55), with thirty-year municipal and taxable yields falling 15.95 and 12.04 bp. At the mean rate state, 270 types hedge, but over 99.99% of

their municipal holdings remain riskless two-year bonds, since they mainly hedge outside exposure. Portfolio responses at other rate states affect continuation prices.

Table 55: Household swap access and the equilibrium municipal gap.

Maturity	Baseline	Swap access	Change
2 years	0.00	0.00	0.00
10 years	23.68	22.24	−1.44
20 years	72.63	65.42	−7.20
30 years	104.27	94.79	−9.48

Notes. Entries are the supply-weighted municipal yield minus the household benchmark, in basis points, averaged over the 36 monthly rate states of 2015–2017 with continuous yields. After-tax shares, parameters and dollar supplies are fixed; both yields adjust. Changes are computed before rounding. Refining the simulation fourfold moves either yield curve or the gap by less than 0.003 bp, and price and optimality residuals are below 10^{-10} .

Hedging costs and interpretation. The test includes asymmetric taxes and nondeductible interest on negative swap cash but excludes dealer fees, bid–ask spreads, margin, collateral opportunity costs and notional limits. The largest mean-state notional is 6.88 times the type’s budget, so the test substantially expands household access.

At fixed bond prices and holdings, let $c_i \geq 0$ be any additional two-year certainty-equivalent cost per unit notional, in A_i ’s absolute units. Maximizing (168) net of this cost gives

$$q_i^*(x; c_i) = \frac{\xi_{i,\text{abs}}^M}{\gamma_i \text{Var}(A_i)} \left[\mathbb{E}A_i - \gamma_i \text{Cov}\left(\widetilde{W}_i(x), A_i\right) - c_i \right]_+. \quad (170)$$

Here $[u]_+ = \max\{u, 0\}$, as in (160), and $\widetilde{W}_i(x)$ is the risk index of bond portfolio x . Costs reduce hedge demand at given bond holdings, and margin and position limits restrict feasible hedges. These frictions oppose risk transfer and are expected to attenuate yield responses. The equilibrium gap depends on both yields, and the 9.1% reduction excludes these costs.

D.7 Data sources

Table 56 lists the data sources, variables and available coverage. The estimation windows used in each analysis are specified in the corresponding subsection.

Table 56: Data sources, variables and coverage.

Source	Variables	Coverage
Municipal issues: Mergent Municipal Bond Database (FTSE Russell)	CUSIP, issue & maturity dates, coupon, callability, tax status (federal/state), issuer state, purpose	1986–2023
Corporate issues: Mergent FISD (FTSE Russell)	CUSIP, issue & maturity dates, coupon, callability, rating, issuer industry, seniority	1986–2023
Insurers: NAIC Schedule D (annual, Q1–Q3 supplements)	CUSIP-level par, book, and market values for U.S. P&C and life insurers	1991–2023
Banks: FDIC BankFind (Call Reports RI, RC-B)	Annual interest expense, average assets and municipal securities holdings by bank	2000–2021
Public pensions: FOIA requests and annual financial reports	CUSIP-level municipal/taxable par and market values, 30 state/local plans (coverage broadens from 2015)	2000–2023
Sector holdings: Financial Accounts (Z.1, L.212, B.101.f)	Quarterly municipal holdings: households, funds, banks, P&C, life, state/local pensions, other	1972–2022
Households: SCF, Federal Reserve	Tax-unit income and filing status, financial assets and debt	2001–2019 waves
Household taxes: IRS SOI	Filer counts, taxable and tax-exempt interest, effective federal rates by AGI bracket	2000–2022
Municipal trades: MSRB RTRS	Trade-level price, par volume, customer-side flag, trade-condition codes	2005–2023
Corporate trades: FINRA TRACE Enhanced (Academic)	Trade-level price, par volume, customer-side flag, condition codes	2002–2023
Treasury curve: GSW, Federal Reserve	Daily par and zero yields, 1–30y. Par yields enter P3, and zeros price its coupon benchmark	2000–2023
Short rates: TB3MS, FRED	Monthly three-month treasury bill rates for the rate process	1952–2024
Treasury supply: Monthly Statement of the Public Debt	Security-level amounts and maturities for the treasury maturity profile	2001–2020
Corporate zeros: HQM (FRED), Mergent FISD robustness	Monthly A-or-above zero curve, 1–30y	2000–2023
Credit risk: Markit/IHS	State-government and AA corporate CDS at 1–30y tenors, coverage of the matched adjustment sample	2004–2020
Expected rates: Kim–Wright, Federal Reserve	Monthly fitted yields and term premia, expected-average-short-rate component capped at 10y	2000–2021
Swaps: Federal Reserve H.15 (FRED)	Par interest-rate swap rates for household swap pricing	2001–2008
State taxes: Tax Foundation	Top marginal rate, bracket thresholds, treatment of out-of-state municipal interest	2000–2022
State income: BEA	Quarterly state-level personal income	2000–2022
Historical state rates: Institute on Taxation and Economic Policy (ITEP)	Top statutory rate by state, the P1 instruments	1945–2000 vintages