Transaction-Specific Tax Reform in Three Steps: The Case of Constructive Ownership

Thomas J. Brennan[*] and David M. Schizer[**]

Abstract

Similar investments are often taxed differently, rendering our system less efficient and fair. In principle, fundamental reforms could solve this problem, but they face familiar obstacles. So instead of major surgery, Congress usually responds with a Band-Aid, denying favorable treatment to some transactions, while preserving it for others. These loophole-plugging rules have become a staple of tax reform in recent years. But unfortunately, they often are ineffective or even counterproductive. How can Congress do better? As a case study, we analyze Section 1260, which targets a tax-advantaged way to invest in hedge funds. This analysis is especially timely because a multi-billion dollar litigation is pending about this rule.

This Article proposes a three-step approach. First, when faced with a new type of tax planning, policymakers should decide whether a response is really necessary. How harmful is the transaction? How feasible is it to target this transaction without also burdening "good" transactions, which don't involve the same abuse? This first phase determines what we call "the normative presumption" about the transaction.

Second, Congress should define which transactions are potentially problematic. An "initial filter" should exempt transactions that clearly don't pose the relevant concern.

Third, once a transaction is deemed to be potentially problematic, a sophisticated test is needed to check whether it actually is. Admittedly, a sophisticated test is costly to administer. This is why initial filters are needed to limit how often it is used.

Along with proposing this three-part framework, this Article offers a novel critique of a sophisticated test the government has begun using: a "delta" test, which measures how closely investments track each other. Although delta is often considered the gold standard, we show how easy it is to manipulate. The trick is to add contingencies (e.g., so the investment terminates when the price reaches a specified level). To head off this gaming, we recommend an alternative test that focuses on value instead of on changes in value–and, more generally, on enduring features instead of temporary quirks.[***]

I. INTRODUCTION

As any tax planner will tell you, form matters in the tax law. With a few tweaks, taxpayers often can get almost the same economic return with very different tax treatment.[1] Unfortunately, as well-advised (wealthy) taxpayers jockey for better treatment, the tax system becomes less efficient and fair. In principle, fundamental tax reforms like a consumption tax or mark-to-market accounting could eliminate

---

[1] For example, different tax rules apply to forward contracts, swaps, contingent bonds, life insurance contracts, and annuities. *See* I.R.C. § 1234A; Treas. Reg. § 1.446-3; Treas. Reg. § 1.1275-4; I.R.C. § 101; I.R.C. § 72. All references to Sections are to the Internal Revenue Code of 1986, as amended, and to treasury regulations interpreting the Code.

these inconsistencies,[2] but these reforms face familiar challenges of administrability and politics.[3]

So instead of major surgery, Congress typically responds with a Band-Aid, denying favorable treatment to some transactions, while preserving it for others. These loophole-plugging rules, which we call "transaction-specific reforms," are supposed to block a particular planning strategy, without disrupting similar transactions that aren't tax-motivated. Yet these measures are often ineffective and even counterproductive. While they complicate the law and create traps for the unwary, these measures don't stop well-advised taxpayers, who simply adjust their planning to avoid them.

To show how transaction-specific reforms can be more effective, this Article uses a case study: the tax on investors[4] in hedge funds.[5] Because these funds trade frequently, their profits usually are short-term capital gain.[6] To qualify for (lower) long-term rates, investors began investing–not in the funds themselves–but in contracts based on their value.[7] By holding these hedge fund derivatives[8] for at least a year, investors used to be eligible for the long-term rate. In response, Congress enacted a transaction-specific reform, Section 1260, to treat these gains as short-term. Yet taxpayers found ways to plan around this statute, which prompted pending litigation with billions of dollars at stake.[9] What is the right tax

---

[2] Under a consumption tax, investment returns generally are not taxed. *See generally* William D. Andrews, *A Consumption-Type or Cash Flow Personal Income Tax*, 87 HARV. L. REV. 1113, 1113-23 (1974) (arguing that a consumption tax would be more simple, fair, and efficient than an accretion-type tax); David A. Weisbach & Joseph Bankman, *The Superiority of an Ideal Consumption Tax over an Ideal Income Tax*, 58 STAN. L. REV. 1413 (2006). Under mark-to-market accounting, investment returns are taxed annually based on changes in the asset's fair market value, regardless of whether the property has been sold. *See generally* David J. Shakow, *Taxation Without Realization: A Proposal for Accrual Taxation*, 134 U. PA. L. REV. 1111, 1111-18 (1986) (proposing broader use of mark-to-market taxation).

[3] Political and other obstacles to either reform are thoroughly explored elsewhere. *See generally* UNEASY COMPROMISE: PROBLEMS OF A HYBRID INCOME-CONSUMPTION TAX (Henry J. Aaron et al. eds., 1988).

[4] A hedge fund usually is taxed as a partnership, which means that its gains and losses are computed by the partnership, but tax is paid by the individual partners. *See* David M. Schizer, *Frictions as a Constraint on Tax Planning*, 101 COLUM. L. REV. 1312, 1368-69 (2001) [hereinafter Schizer, *Frictions*].

[5] Hedge funds engage in sophisticated trading strategies that are supposed to deliver superior returns on a risk-adjusted basis. They usually are open only to sophisticated investors who satisfy income and asset tests under the securities law. *See* The Investopedia Team, *What Are Hedge Funds? Examples, Types, and Strategies,* INVESTOPEDIA (Aug. 11, 2022), https://www.investopedia.com/terms/h/hedgefund.asp [https://perma.cc/M9KC-NXS7].

[6] When taxpayers hold an investment for more than one year, their capital gains qualify as "long term," and thus are eligible for a reduced rate. In contrast, short-term gains are taxed at the (higher) rate for ordinary income. *See* I.R.C. § 1(h).

[7] *See* Schizer, *Frictions, supra* note 4, at 1368-69.

[8] A derivative is a contract whose value derives from some financial fact. *See generally* Global Derivatives Study Group, *Derivatives: Practices and Principles* 26, THE GROUP OF THIRTY (July 1993), https://group30.org/images/uploads/publications/G30_Derivatives-PracticesandPrinciples.pdf [https://perma.cc/J7EE-XUV3] ("In the most general terms, a derivatives transaction is a bilateral contract or payments exchange agreement whose value derives, as its name implies, from the value of an underlying asset or underlying reference rate or index.").

[9] GWA v. Comm'r, No. 6981-19 (U.S.T.C. filed May 10, 2019).

treatment for hedge fund derivatives? More generally, how can transaction-specific reforms be more successful in enhancing the tax system's efficiency and fairness?

This Article recommends a three-part approach. First, when faced with a new type of tax planning, policymakers should decide whether the strategy is harmful enough to warrant a response. This first phase determines what we call "the normative presumption" about the transaction. Second, Congress should use an "initial filter" to exempt transactions that clearly don't pose the relevant concern. Third, once a transaction is deemed to be potentially problematic, a sophisticated test is needed to check whether it actually is. For example, is a hedge fund derivative enough like the underlying fund that it should be recharacterized? With a nuanced and sophisticated rule, the government can catch transactions that actually are economically similar, regardless of what form the taxpayer uses. Admittedly, a sophisticated test can be costly to administer. This is why initial filters are needed to limit how often it is used.

Along with proposing this three-part framework, this Article offers a novel critique of a sophisticated test the government has begun using: a "delta" test,[10] which measures how closely investments track each other.[11] Although delta is often considered the gold standard,[12] we show how easy it is to manipulate. The trick is to add contingencies, for instance, so the investment terminates when the price reaches a specified level. As far as we know, this critique of delta is new to the literature. To head off this gaming, we recommend an alternative test that focuses on value instead of on changes in value—and, more generally, on enduring features instead of temporary quirks.

As should be clear by now, this Article focuses on incremental reform, not fundamental reform. If only modest measures are politically and administratively feasible, policymakers need to know how to choose the right ones. To give them guidance, we take the current highly imperfect system as given and explore how to implement better constructive ownership rules and, more generally, more effective transaction-specific reforms. Although our approach will not achieve optimality, it can still improve the system in meaningful ways.

Part II introduces our case study, Section 1260. Part III lays out the normative criteria for evaluating transaction-specific reforms: distribution and efficiency. Part IV explains the first step in our proposed approach, which we call "the normative presumption" about a tax planning strategy. Part V explains the second of our three steps, "the preliminary filter," which exempts transactions that do not warrant careful vetting. Part VI covers the third stage, which uses sophisticated tools to specify which transactions should be recharacterized. Part VII applies our approach to a pending case, *GWA, LLC v. Commissioner*. Part VIII is

---

[10] *See, e.g.,* Treas. Reg. § 1.871-15 (using delta to determine whether withholding is required on dividend equivalent payments on equity derivatives).

[11] In the parlance of derivatives traders, "delta" measures how much the value of a derivative changes when the price of the underlying asset changes by a dollar. James Chen, *What is Delta in Derivatives Trading, and How Does it Work?*, INVESTOPEDIA (July 26, 2023), https://www.investo pedia.com/terms/d/delta.asp [https://perma.cc/4YPD-3R9F].

[12] *See, e.g.*, N.Y. STATE BAR ASS'N TAX SECTION COMM. ON FIN. INSTRUMENTS, COMMENTS ON "SHORT-AGAINST-THE-BOX" PROPOSAL, NYSBA TAX SECTION REP. #868 18-22 (March 1, 1996) [hereinafter "NYSBA, REP. #868"] (discussing use of delta to test for constructive sales under early proposed version of Section 1259). One of us (David Schizer) helped prepare this report.

the conclusion. Parts IX and X are appendices, which elaborate on specific aspects of the analysis.

II. SIMULATING OWNERSHIP OF HEDGE FUNDS: THE TRANSACTION TARGETED BY SECTION 1260

Often, a high-profile transaction exposes a deeper problem in the tax system. But instead of addressing the underlying issues, Congress opts for a more targeted response. To illustrate this common pattern, this Article focuses on hedge fund derivatives.

A. The Targeted Transaction: Hedge Fund Derivatives

To implement complex trading strategies that are supposed to deliver superior returns, hedge funds trade constantly. So instead of qualifying as long-term capital gain, their profits usually are taxed at (higher) ordinary rates. The frequent trades also mean that gains are taxed currently, instead of being deferred.

This tax is paid by investors, not the fund, which typically is a partnership for tax purposes.[13] As a result, taxable investors often have mixed feelings about hedge funds. They want the pre-tax return, but not the steep tax bill.

In response, investment banks offered a way for investors to get hedge fund returns without the high tax: investing in a derivative based on the hedge fund's value, instead of in the hedge fund itself.

For example, an investor could enter into a forward contract with a securities dealer, committing to buy the fund interest for a preset price after a fixed term of years. This derivative would immediately transfer the hedge fund's economic return to the investor. After all, the investor would pay the same fixed price, whether the fund gained or lost value. But unlike the fund interest, the derivative would not be taxable until it matured or was terminated. At this point, if the transaction was structured properly, the gain (or loss) would be treated as long-term capital gain.[14]

How could a securities dealer offer this contract? The key was for the dealer to invest in the hedge fund. By purchasing an interest in the fund, the dealer could hedge its obligation to deliver an interest to its client in the future.

But didn't this just shift the heavy tax burden to the dealer? As an investor in the fund, didn't the dealer have to pay the very tax its client wanted to avoid? Actually, the answer is "no," since securities dealers are subject to different tax rules. In general, they mark their inventory to market and their gains and losses are ordinary income, not capital gain.[15] As a result, the derivative and the hedge fund were offsetting not just economically, but also in their tax treatment. In general, ordinary income on the hedge fund was offset by ordinary loss on the derivative, and vice versa.[16]

---

[13] *See* Schizer, *Frictions*, *supra* note 4, at 1368-69.

[14] *See* I.R.C. § 1234A (termination of a contract is treated as long-term capital gain or loss).

[15] I.R.C. § 475 (requiring securities dealers to mark inventory to market); I.R.C. § 1221(a)(1) (noting that inventory is not a capital asset).

[16] *See* David M. Schizer, *Sticks and Snakes: Derivatives and Curtailing Aggressive Tax Planning*, 73 S. CAL. L. REV. 1339, 1367-68 (2000); *see also* Schizer, *Frictions, supra* note 4, at 1369 n.206.

Even so, this transaction was not invulnerable. Invoking traditional principles of tax ownership, the government could argue that the real investor in the fund was the client, not the dealer. To defeat this argument, taxpayers and their advisors introduced various features to reinforce the claim that the dealer was the true owner. For example, the dealer needed to have discretion about whether (and how) to hedge its position in the contract.[17]

B. Section 1260

In response, Congress enacted Section 1260 in 1999. Under this provision, if a derivative tracks a hedge fund's return too closely, a portion of the long-term capital gain on the derivative is taxed as ordinary income.[18] In addition, to offset the tax deferral offered by the derivative, an interest charge is imposed once the tax comes due.[19]

Yet Section 1260 is not precise in specifying how closely the derivative's economic return has to track the hedge fund's return. How much of a gap is sufficient to avoid the statute? The key word in the statute is "substantial"; in general, a derivative is caught if it provides "substantially all" the risk of loss and opportunity for gain in the hedge fund.[20] But what does "substantial" mean in this context?[21] Almost twenty-five years later, no regulations have been issued to clarify this question.[22]

---

[17] *See* Schizer, *Frictions, supra* note 4, at 1369 n.206.

[18] I.R.C. § 1260(a).

[19] I.R.C. § 1260(b).

[20] For example, a notional principal contract is covered if the holder "**(A)** has the right to be paid (or receive credit for) all or *substantially all* of the investment yield (including appreciation) on such financial asset for a specified period, and **(B)** is obligated to reimburse (or provide credit for) all or *substantially all* of any decline in the value of such financial asset." I.R.C. § 1260(d)(3) (emphasis added). Likewise, the statute also applies to a taxpayer who "is the holder of a call option, and is the grantor of a put option, with respect to the financial asset and such options have *substantially equal* strike prices and substantially contemporaneous maturity dates." I.R.C. § 1260(d)(1)(C) (emphasis added). I.R.C. § 1260 also applies to forward contracts and, under regulations, to "other transactions (or . . . positions) that have *substantially the same effect*" as the transactions listed above. I.R.C. § 1260(d)(1)(D) (emphasis added).

[21] Guidance can be gleaned from another provision that uses similar language, Section 1259, which imposes tax when taxpayers hedge appreciated assets too perfectly. In 2003, the I.R.S. indicated that a forward contract does not deliver a substantially fixed number of shares if the number would vary between 80 and 100, depending on the stock price when the forward contract matures or is terminated. "According to the Agreement, delivery of a number of shares, which may vary between 80 and 100 shares, depends on the fair market value of the stock on the Exchange Date," the I.R.S. said. "Because this . . . is a significant variation, the Agreement is not a contract to deliver a substantially fixed amount of property for purposes of §1259(d)(1)." Rev. Rul. 2003-7, 2003-1 C.B. 363. Even so, the strategies for avoiding Section 1259 and Section 1260 are somewhat different. *See generally* Schizer, *Frictions, supra* note 4 (noting similarities in the statutory language of Section 1259 and Section 1260, but important differences in the nontax cost of avoiding these provisions, and thus in the strategies for planning around them).

[22] Likewise, there are no regulations to resolve other issues. For example, aside from the derivatives specified in the statute, Congress authorized the Treasury to recharacterize "other transactions . . . that have substantially the same effect" as the listed transactions. I.R.C. § 1260(d)(1)(D). Along with derivatives based on hedge funds, the statute also applied to derivatives based on mutual funds, REITs, and other pass-through entities. I.R.C. § 1260(c)(2). But what about derivatives based on debt instruments or common stock? Congress authorized the Treasury to cover these instruments

Meanwhile, the government has pursued a range of cases against hedge fund derivatives that might–or might not–be covered by the statute. The George Weiss case went to trial in 2022.[23] As of this writing, a decision is still pending.

III. NORMATIVE CRITERIA FOR EVALUATING TRANSACTION-SPECIFIC REFORMS: DISTRIBUTION AND EFFICIENCY

How do we know whether a rule like Section 1260 is a good idea? After all, there are a host of other ways to raise (and lower) taxes. In deciding which incremental reforms to pick, policymakers should consider two familiar criteria: distribution and efficiency.

A. Distribution

A key question is whose tax bill is changing. Do they have high incomes? Does the change make the system more or less progressive?

Policymakers should care about not only whose tax bill is cut, but also how it is cut. If the tax system is easy to game, sophisticated advice commands a premium, which is easier for wealthy taxpayers to afford. Yet we should not want a system in which, in the infamous words of a famous wealthy tax cheat, "only the little people pay taxes."[24]

This unequal access is not just unfair, but also demoralizing. If taxpayers believe that others are not paying their fair share, they are less likely to comply voluntarily. In response, the I.R.S. has to spend more on enforcement.[25]

B. Efficiency and the "Hit or Miss" Quality of Transaction-Specific Reforms

Along with distribution, policymakers also need to consider efficiency. Do Section 1260 and other transaction-specific reforms increase the efficiency of the tax system? In general, an efficient tax changes taxpayer behavior as little as possible.[26] This is a critical goal not only in fundamental tax reform,[27] but also in the transaction-specific reforms analyzed here.

---

under regulations. I.R.C. § 1260(c)(1)(B). But again, no regulations have been written on these issues as well.

[23] The taxpayer filed its original petition with the Tax Court in May 2019.

[24] *Top 10 Tax Dodges*, TIME, https://content.time.com/time/specials/packages/article/0,28804,1891 335_1891333_1891317,00.html [https://perma.cc/QZ8X-AJ5N] (quoting Leona Helmsley, a wealthy real estate developer who was imprisoned for tax fraud).

[25] *See* Benno Torgler, *Tax Morale, Rule-Governed Behaviour and Trust*, 14 CONST. POL. ECON. 119 (2003), https://doi.org/10.1023/A:1023643622283 (trust in public officials and the legal system has a significant positive effect on tax morale).

[26] To be more precise, an efficient tax changes behavior as little as possible as compared to a world with no taxes, except that taxpayers have less money. In other words, efficient taxes minimize substitution effects, but there is still an income effect. *See* David A. Weisbach, *Line-Drawing Doctrine and Efficiency in the Tax Law*, 84 CORNELL L. REV. 1627, 1652 (1999) [hereinafter Weisbach, *Line-Drawing*] ("tax efficiency is concerned with the difference between consumers' actual after-tax behavior and the behavior they would engage in merely because they have less revenue").

[27] *See, e.g.,* Alan Auerbach, *Retrospective Capital Gains Taxation,* 81 AM. ECON. REV. 167, 168 (1991) (arguing that charging interest upon realization would enhance efficiency of tax system);

It is well understood that the key in an incremental reform like Section 1260 is to reduce various types of social waste, including administrative costs, changes in work and savings decisions, and tax planning costs.[28] There usually are tradeoffs among these costs; for instance, targeting tax planning more effectively can increase administrative costs. Even if some costs increase, a reform can still enhance efficiency by cutting others, so there is a net reduction in the sum of these costs.[29]

1. Administrative Costs

Let's begin with the cost of enforcing and complying with the law. Transaction-specific reforms like Section 1260 usually increase these costs by making the tax code more complicated.[30] Congressional staffers and Treasury officials must draft the new rules, and then taxpayers have to pay advisors to help them comply. Meanwhile, I.R.S. auditors need to monitor their compliance, and there may be litigation on contested issues.[31]

Even so, some transaction-specific reforms increase administrative costs more than others. Obviously, policymakers should look for ways to economize on these costs–a theme that features prominently in our analysis.

Bankman & Weisbach, *supra* note 2 (arguing that an ideal consumption tax distorts behavior less than an ideal income tax); Daniel Halperin, Saving the Income Tax: An Agenda for Research, Woodworth Lecture (Nov. 7, 1997), *in* 77 TAX NOTES (TA) 967, 967 (Nov. 24, 1997), *reprinted in* 24 OHIO N. U. L. REV. 493 (1998) (arguing that mark to market taxation for publicly traded securities would enhance efficiency of tax system); Michael J. Graetz, 100 MILLION UNNECESSARY RETURNS: A SIMPLE, FAIR, AND COMPETITIVE TAX PLAN FOR THE UNITED STATES (2008) (arguing that replacing the income tax with a VAT for most Americans would enhance the efficiency of the tax system); Louis Kaplow, *The Optimal Supply of Public Goods and the Distortionary Cost of Taxation*, 49 NAT. TAX J. 513 (1996) (reexamining relationship between optimal supply of public goods and distortionary cost of income taxation).

[28] *See, e.g.,* Daniel Hemel, Jennifer Nou & David A. Weisbach, *The Marginal Revenue Rule in Cost-Benefit Analysis*, 160 TAX NOTES 1507 (2018) (proposing method of cost-benefit analysis for tax regulations, "the marginal revenue rule," which accounts for behavioral effects (measured indirectly through a regulation's impact on revenue), administrative and compliance costs, and non-revenue-related effects); Daniel Hemel & David A. Weisbach, *The Behavioral Elasticity of Tax Revenue*, 13 J. LEGAL ANALYSIS 381, 382 (2021) (proposing "Behavioral Elasticity of Tax Revenue (BETR)" as "a tool for analyzing problems of tax system design," which "allows tax laws and policies to be measured and compared using a unified metric"); Daniel N. Shaviro, *An Efficiency Analysis of Realization and Recognition Rules Under the Federal Income Tax*, 48 TAX L. REV. 1 (1992); Joel Slemrod & Shlomo Yitzhaki, *The Costs of Taxation and the Marginal Efficiency Cost of Funds*, 43 I.M.F. STAFF PAPERS 172, 175 (1996) ("We explore an alternative approach that is limited to identifying desirable marginal changes, without a claim of optimality."); Weisbach, *Line-Drawing, supra* note 26; David A. Weisbach, *Ten Truths About Tax Shelters*, 55 TAX L. REV. 215 (2002).

[29] *See* Slemrod & Yitzhaki, *supra* note 28, at 183 ("[W]e offer a tractable methodology that can evaluate marginal changes in tax systems and take account of all five components of the cost of tax systems. The methodology is based on the concept of the marginal cost of public funds.").

[30] *See id.* at 179-81 (discussing administrative and compliance costs).

[31] *See generally* Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L. J. 557, 557-629 (1992) (noting various costs of drafting and applying laws and tradeoffs among them).

2. *"Regular" Deadweight Loss From Real Effects*

Even as transaction-specific reforms increase administrative burdens, do they reduce other sources of waste? Another familiar one is when tax rules distort work and savings decisions.[32]

In general, there is less of this "regular deadweight loss" when taxpayers have no easy way to change their behavior. In this situation (i.e., when their demand is inelastic), they are still likely to engage in the relevant transaction even if it is taxed more heavily.[33] As a result, this flexibility (or "elasticity") is a key variable.

Of course, the premise here is that taxpayers' choices are socially optimal, but this isn't the case when markets fail. For example, if an activity imposes social costs that aren't impounded in the price (as do greenhouse gas emissions, addictive drugs, and other activities that harm third parties), discouraging this activity actually is a good thing.[34] So the activity's social value is relevant when policymakers consider changing its tax treatment.

So, in deciding whether to target a particular planning strategy, policymakers should consider what the effective tax rate on the underlying activity actually should be. Obviously, tax planning reduces the effective rate (i.e., what taxpayers actually pay) below the nominal rate (i.e., the rate on the books), but is this a good or a bad thing?

For example, imagine that the nominal rate is 40%, but a planning strategy reduces the effective rate to 25%. If the rate really should be 40%–based on elasticity, externalities, and the like–the transaction is problematic. But if it should be 25% (or even 22% or 28%), the planning actually is bringing the effective rate closer to where it should be.[35]

In other words, is the relevant planning causing this activity to be undertaxed? If so, shutting down this planning can make the system more efficient. But if not–that is, if the nominal tax rate is too high–there is less reason to target the planning strategy.

In the case of Section 1260, for example, how bad is it to raise the tax on indirect investments in hedge funds? Will investors invest less in these funds? If so, how much less and what would they do instead? In addition to this sort of analysis of elasticity, policymakers should also consider externalities. For example, do hedge funds improve price discovery? Or are they just a form of zero-sum gambling? Even if the social contribution of hedge funds is limited, is Section 1260

---

[32] *See* Slemrod & Yitzhaki, *supra* note 28, at 181 (discussing "regular deadweight loss").

[33] *See* Weisbach, *Line-Drawing, supra* note 26*,* at 1658 ("If the elasticity is high, the tax should be low, and if the elasticity is low, the tax should be high."). Yet the inquiry is nuanced. While taxing inelastic activity generally is more efficient, this isn't always the case. For instance, if demand is inelastic but taxpayers change their behavior anyway, the welfare losses can be especially large. *See id.* at 1656 ("the ability to raise taxes on low-elasticity items is limited because as the tax on a commodity increases, the marginal deadweight loss increases").

[34] *See id*. at 1654 ("To the extent that there is a market failure, the definition of an efficient tax changes. In particular, so-called Pigouvian taxes are taxes, or subsidies, that attempt to cure market failures."); David M. Schizer, *Red White and Blue–and also Green: How Energy Policy Can Protect Both National Security and the Environment*, 96 S. CALIF. L. REV. (forthcoming 2024) (noting that Pigouvian taxes should be set equal to the marginal harm from the relevant activity).

[35] *See* Weisbach, *Line Drawing, supra* note 26, at 1669 ("Taxing a low-elasticity item too high is not optimal.").

overbroad? Does it inadvertently discourage other real activity that is more valuable, implicating other elasticities and externalities? In short, in crafting Section 1260 and other transaction-specific reforms, policymakers should consider whether the measure moves the effective tax rate closer to or further from the optimal level.[36]

3. *Planning Costs*

Even if the effective rate is too high, tax planning usually is an inefficient way to reduce it. Taxpayers have to invest time and resources to plan around the relevant rule.[37] Along with hiring the right advisors, they have to modify their transactions in ways that might be unappealing. If the effective rate needs to be cut, policymakers generally should just do that directly, instead of relying on taxpayer self-help.

All else being equal, then, reducing the waste from tax planning improves efficiency. Indeed, these efficiency gains can offset–and in some cases can exceed–increases in administrative costs and "regular" deadweight loss.[38]

But obviously, this can happen only if reforms actually discourage planning, instead of merely prompting a more elaborate variation of it. This brings us to a critically important question: Does the targeted planning actually stop? Or does it simply metastasize into a different (more costly) variation?[39]

For example, assume that the tax law distinguishes between two transactions: Y (which is taxed unfavorably) and Z (which is taxed favorably). For instance, Y could be a hedge fund and Z could be a growth stock, or Y could be equity and Z could be debt, or Y could be subject to withholding while Z is not, etc.

In response to this inconsistent treatment, taxpayers develop a new transaction, X, which is economically like (unfavorable) Y but is taxed like (favorable) Z. For example, X could be a hedge fund derivative, which offers the economics of a hedge fund but the tax treatment of a growth stock.

As taxpayers begin engaging in X, Congress changes its tax treatment, so X is taxed like (unfavorable) Y, instead of (favorable) Z. Notably, the response is usually a narrow one. Congress does not eliminate the distinction between Y and Z. Instead, Congress just redefines the boundary between Y and Z, ensuring that X is taxed like Y.

But what if the new rule covers X, but not a similar but costlier variation of X that we call XX? In this case, some taxpayers may start doing (favorable) XX to

---

[36] *Id.* at 1679 (noting that one of the two most important factors in line-drawing is "whether transactions are taxed appropriately when considered by themselves (i.e., without regard to line drawing)").

[37] Slemrod & Yitzhaki, *supra* note 28, at 181.

[38] As Joel Slemrod and Shlomo Yitzhaki have emphasized, the efficiency of an incremental reform depends on its effect on the *sum* of various efficiency costs, a concept they call the marginal efficiency cost of funds. *Id.* at 183. ("[W]e offer a tractable methodology that can evaluate marginal changes in tax systems and take account of all five components of the cost of tax systems. The methodology is based on the concept of the marginal cost of public funds.").

[39] Weisbach, *Line Drawing, supra* note 26, at 1670 ("We cannot simply look at how many taxpayers avoid a line. We must also look at the costs of doing so.").

avoid the rule.[40] Meanwhile, others who would have done X, but consider XX an inadequate substitute, will simply do (unfavorable) Y.

The distribution of these reactions is important. On the one hand, if everyone who was going to do X "doubles down" on tax planning to do XX, the reform is not successful.[41] The government does not collect more revenue, and distributional goals are not advanced. Instead, the relevant tax planning continues but in a more elaborate way, making the tax system less efficient.

On the other hand, if everyone who was going to do X "gives up" and simply does Y–since XX is too hard or unappealing–the reform is successful. The government collects more revenue, distributional goals are advanced, and the planning costs of X (and of XX) are avoided.

Of course, with any reform there is unlikely to be a uniform reaction. Some taxpayers will double down with XX (increasing planning costs), while others will give up and do Y (reducing planning costs), and the net of these effects determines whether the reform actually reduces planning costs. The distribution of these reactions depends, in turn, on how well-crafted and rigorous the reform is.

At the end of the day, the efficiency analysis turns on the net of all these effects–on administrative costs, real behavior, and planning. Some transaction-specific reforms increase social waste, while others reduce it.

4.   Building on the MECF Framework

While others have highlighted key issues and tradeoffs in incremental tax reform, this Article adds to the literature in three ways. First, we help policymakers operationalize this literature by recommending three steps, which focus on different aspects of the analysis. In asking what the right tax burden should be on the relevant transaction, the first step ("the normative presumption") focuses on regular deadweight loss and distribution. Likewise, in urging policymakers to identify the relevant transaction in a way that is both precise and economical, the second step ("preliminary filters") and third step ("rigorous analysis") manage the tradeoff between administrative and planning costs.

Second, and relatedly, we recommend using preliminary filters as a way to economize on administrative costs, enabling more precise (and administratively costly) tests to be deployed only when they are needed. This insight has broad application not just to our case study of Section 1260, but to any transaction-specific anti-abuse rule.

Third, we identify problems with a sophisticated test that commentators have recommended and the government has used for financial transactions: comparisons based on "delta." As far as we know, these insights are new to the literature.

---

[40] *See* Shaviro, *supra* note 28.

[41] Daniel N. Shaviro, *Economic Substance, Corporate Tax Shelters, and the* Compaq *Case*, 88 TAX NOTES 221, 223 (2000) (noting that the desirability of an anti-abuse rule depends on how effective it is in "generating such deterrence rather than simply inducing taxpayers to jump through a few extra hoops before getting the desired tax consequences anyway").

IV. STEP 1: THE NORMATIVE PRESUMPTION

Let's turn to the first of the three steps we propose: when confronted with a new planning strategy, policymakers should decide how urgently a response is needed, as well as how broad and tough it should be. In making these judgments, policymakers establish what we call the "normative presumption" about a transaction. To do so, they should consider three sets of issues, which this Section explores in turn:

- First, how unfair is the targeted transaction? How does it affect the distribution of tax burdens, as well as public trust in the tax system?

- Second, what is the right tax burden on the targeted transaction (in light of elasticity, externalities, and other efficiency considerations)?

- Third, how costly would it be for the transaction-specific response to be overbroad, so it reaches activity other than the targeted transaction?

While policymakers should consider these issues in crafting any transaction-specific reform, this Part focuses on Section 1260 as a case study.

A. Fairness: Distribution and Trust

In deciding how tough to be, policymakers need to consider whether the targeted planning makes the system less fair. As this Section shows, the relevant issues are distribution, horizontal equity, and trust in the system.

1. Vertical Equity

Since high-net worth taxpayers have better access to tax advice, they often are the main beneficiaries. When they are the main users of a planning strategy, it reduces the tax burdens of those with the greatest ability to pay.

This is likely the case with hedge fund derivatives. Under the securities law, hedge funds are open only to investors who satisfy income or asset tests.[42] The same is true of the customized derivatives used in these transactions. They are not available on an exchange, where a broader group of customers would have access to them, but in the "over-the-counter" market, where investors need to satisfy income or asset tests.

Admittedly, transactions that are first crafted just for sophisticated taxpayers can become more widespread over time. For example, a variation of tax shelters that became popular in the 1980s–leveraged investments in depreciable property–were mass marketed through partnerships, so upper-middle-class dentists and accountants could use them (until Congress shut them down).

In principle, hedge fund derivatives might also start attracting a broader clientele over time, for instance, as legal restrictions ease and transactions costs decline.[43] But as of now, hedge fund derivatives are available only to wealthy taxpayers.

---

[42] Of course, this includes pension funds and nonprofits, which may serve low-income people.

[43] For example, mutual funds appeal to a much broader class of investors. Some funds trade frequently, triggering current tax liability at a higher rate. To cut their investors' tax bills, these funds might find ways to make their investments indirectly through derivatives. Alternatively,

As a result, vertical equity is a plausible rationale for targeting them. Yet in our view, this is not a "slam dunk," if only because high-income taxpayers can get the same tax treatment–tax deferral and a reduced rate–with a host of other investments, including growth stocks, index funds, venture capital, private equity, and real estate. It's not clear why getting these tax advantages this way is worse than getting them other ways.

In principle, there might be a difference based on Congressional intent but, in our view, this difference isn't relevant. Perhaps Congress meant to reduce the tax burden on these other investments, but not on hedge fund derivatives. Yet even if true,[44] this is beside the point. After all, the issue explored in this Article is not what Congress has done, but what it should do. The relevant question is whether a transaction-specific reform like Section 1260 is a good idea. The answer should turn on fairness and efficiency, not Congressional intent.

2. *Horizontal Equity*

Arguably, policymakers should consider not just vertical equity, but also horizontal equity, which provides that similarly situated taxpayers should be taxed the same way.[45] Hedge fund derivatives violate this norm by allowing investors in derivatives to pay less tax than investors in the underlying fund. This inconsistency puts a premium on good advice. Taxpayers who don't know about the more tax-efficient strategy face a "trap for the unwary."

In principle, horizontal equity might justify a response like Section 1260, but we are skeptical, if only because both groups of taxpayers–investors in funds and in derivatives–are quite well off. How sympathetic should we be to wealthy people who have bad tax advisors? Are reforms really needed to protect them?

This is not to say that traps for the unwary are desirable. Yet in our view, they are a problem of efficiency, not fairness: the issue is not so much that they are unfair (at least in ensnaring wealthy people), but that they encourage an over-investment in tax advice.[46]

3. *Trust*

Along with vertical and horizontal equity, planning by high-income taxpayers can implicate another concern as well: when publicized in the media, this planning can erode confidence in the tax system, as noted above. If other taxpayers learn of these strategies, they might think "the game is rigged" and stop paying tax voluntarily.

---

derivatives exchanges might start offering derivatives based on the value of these mutual funds. *See generally* Andriy Blokhin, *Can Mutual Funds Invest in Options and Futures?*, INVESTOPEDIA (Sept. 29, 2022), https://www.investopedia.com/ask/answers/091115/can-mutual-funds-invest-options-and-futures.asp [https://perma.cc/B66X-JEN2].

[44] It is not obvious that the favorable treatment of hedge fund derivatives before Section 1260 was somehow inadvertent. After all, this treatment was simply an application of the general rules for derivatives that Congress had enacted. *See, e.g.,* I.R.C. § 1234 (treatment of options); I.R.C. § 1234A (treatment of the termination of derivative contracts).

[45] Arguably, horizontal equity is not a compelling norm, if only because it simply raises the further question of whether taxpayers *actually are* similarly situated. *See generally* Louis Kaplow, *Horizontal Equity: Measures in Search of a Principle*, 42 NAT'L TAX J. 139 (1989).

[46] We focus on planning and administrative costs in Parts V & VI, *infra*.

In our view, this is a plausible, but incomplete, rationale for Section 1260 and other transaction-specific reforms. Confidence in the tax system undoubtedly is important, but how vigorous must the response be to preserve it? After all, if the problem is one of perception, is the mere appearance of a response sufficient? In principle, the government might achieve its goal by seeming to stop the planning, while (quietly) letting it continue. Presumably, though, this symbolic response would no longer suffice–and might even be counterproductive–if the media exposes its ineffectiveness.

To sum up, the most persuasive equity-based argument for targeting hedge fund derivatives is that they are available only to wealthy taxpayers. A transaction-specific reform also may be useful in reinforcing confidence in the system.

B. Efficiency: The Right Tax Burden

In deciding whether to target a planning strategy–and, if so, how vigorously– policymakers also need to consider efficiency. As detailed in Part III, this requires policymakers to consider three issues: the right tax burden on an activity (based on elasticity, externalities, etc.); administrative costs; and planning costs.

In setting the normative presumption for a reform–which, again, is the first stage in our three-step analysis–we urge policymakers to focus on the first of these issues: the efficient tax burden. Specifically, they should consider two questions, which this Part discusses in turn:

- First, what is the right tax burden for the targeted transaction?

- Second, how great is the risk of overtaxing "good" transactions that don't present the relevant abuse?

1. The Right Tax Burden: The Gravitational Pull of Fundamental Reforms

In addressing the first of these questions–the right tax burden–policymakers face a fundamental issue when the planning strategy is for investments, as often is the case: How should investments be taxed? Should our system use an income tax (which taxes investments) or a consumption tax (which does not)? The debate over these tax bases has raged for decades, and we do not seek to contribute to it here.[47]

Instead, we emphasize two ways in which this debate influences transaction-specific reforms. First, if more ambitious reforms are enacted, Section 1260 (and many other transaction-specific reforms) would become irrelevant.

Second, policymakers' priors on fundamental reform can influence their view of a particular transaction-specific reform. After all, to assess how abusive a transaction is, they need to determine what the tax on it is supposed to be. Their answer is likely to vary, for instance, depending on whether they would prefer a

---

[47] *See* Bankman & Weisbach, *supra* note 2, at 1414 ("Perhaps the single most important tax policy decision is the choice between an income tax and a consumption tax. The topic has been discussed and argued over since at least the time of Hobbes and Mill, without apparent resolution."); *See also, e.g.*, Andrews, *supra* note 2; Barbara H. Fried, *Fairness and the Consumption Tax*, 44 STAN. L. REV. 961 (1992); Michael Graetz, *Implementing a Progressive Consumption Tax*, 92 HARV. L. REV. 1575 (1979); Alvin Warren, *Would a Consumption Tax Be Fairer than an Income Tax?*, 89 YALE L. J. 1081 (1980).

consumption tax, on the one hand, or mark-to-market accounting, on the other. In a sense, their take on fundamental reform can exert something like a gravitational pull on their judgments about transaction-specific reforms.

2.   *The Right Tax Burden: A Consumption Tax?*

So if policymakers prefer not to tax any investments, they are less motivated to backstop the tax on specific investments, such as hedge fund derivatives. For example, assume that a policymaker wants to let taxpayers deduct the cost of their investments.[48] Under this fundamental reform (a cash-flow consumption tax), the form and timing of investments no longer matter. Hedge funds and hedge fund derivatives are taxed the same way–as are taxpayers who trade frequently and those who "buy and hold." Their tax depends on *how much* they save, not *how* they save.

For a policymaker who favors this approach, allowing deferral and long-term capital gains for hedge fund derivatives is not particularly objectionable; on the contrary, it can be a step in the right direction.[49]

3.   *The Right Tax Burden: Should Holding Period Matter?*

Alternatively, even if policymakers want to tax investments, they still might not object to hedge fund derivatives. In part, their view might turn on how committed they are to a rule this strategy games: the holding period rule for long-term capital gains.

Arguably, this rule is misguided. After all, although there are reasons to encourage taxpayers to save and invest, why encourage them to stick with the same investment? Is "buying and holding" really more socially valuable than monitoring markets and redeploying capital?[50] On the contrary, a tax incentive not to sell–a phenomenon known as lock-in"–has familiar efficiency costs. When taxpayers keep positions they want to sell, not only are they less happy with their portfolios, but market prices don't reflect new information as effectively.

Policymakers who worry about lock-in (as we do) might want to reward taxpayers for investing for over a year not just in a single position, but in a series of positions. In this spirit, they might favor a "rollover" rule that defers tax (and continues the holding period) when taxpayers sell one investment and immediately

---

[48] As Cary Brown showed years ago, allowing taxpayers to deduct the cost of investments generally is equivalent to exempting the yield. *See* E. Cary Brown, *Business-Income Taxation and Investment Incentives*, *in* INCOME, EMPLOYMENT AND PUBLIC POLICY: ESSAYS IN HONOR OF ALVIN H. HANSEN 300, 300-16 (1948), *reprinted in* Am. Econ. Ass'n, READINGS IN THE ECONOMICS OF TAXATION 525, 525-37 (Richard A. Musgrave & Carl S. Shoup eds., 1959). To implement this reform, Congress could offer deductible IRAs with no limit on the amounts invested and freedom to withdraw proceeds at any time. With these "mega-IRAs," the treatment of all investments would be consistent: Amounts contributed would be deductible, transactions inside the IRA would have no tax consequences, and withdrawals would be taxed as ordinary income.

[49] Admittedly, a partial step can sometimes turn out to be less appealing than a fundamental reform– and even counterproductive–by introducing inefficiencies and inequities of its own. This is a key reason why incremental reforms are so challenging–and, indeed, why the normative presumption is just one step in our three-step analysis.

[50] Perhaps the concern is that taxpayers otherwise would trade too often, indulging speculative impulses and racking up transaction costs. But it is not obvious why tax policy is the right instrument to constrain these impulses.

reinvest the proceeds in another.[51] If enacted, this provision would extend the favorable treatment of derivatives–deferral and reduced rates–to the underlying fund.[52]

So if policymakers would favor this reform but can't get it enacted, would they be motivated to target hedge fund derivatives? Arguably, the answer is "no." The treatment of the derivative (a deferred tax at long-term rates) is more consistent with the rule they want (rollover) than the treatment of the underlying fund (annual taxes at short-term rates). As a result, these policymakers might view the derivative as self-help that moves the system in the right direction.

4. The Right Tax Burden: Should There Be a Capital Gains Preference?

In contrast, policymakers who are skeptical of the capital gains preference have the opposite view. They want ordinary rates to apply to all investment gains, which would eliminate a key advantage of derivatives: it would be taxed at the same rate as the fund (but would still offer tax deferral).

If these policymakers aren't able to enact this reform, they will be motivated–much more than, say, proponents of a reduced tax (or no tax at all) on investments–to police the holding period rule, if only to reduce the volume of investments taxed at rates they consider too low. Put another way, since they think current law taxes hedge funds appropriately, they are especially eager to keep taxpayers from avoiding this treatment.

5. The Right Tax Burden: Should We Have Mark-to-Market Accounting?

The same is true of policymakers who favor mark-to-market accounting. To measure income more accurately, they want to tax gains every year, regardless of whether there has been a sale. This reform would eliminate the timing advantage of hedge fund derivatives (taxing gains every year, as generally is the case with underlying funds that trade frequently).

Again, if this reform is unavailable,[53] these policymakers will be especially motivated–more than advocates of a reduced tax (or no tax) on investments–to target tax deferral on hedge fund derivatives. For example, Section 1260 pursues this goal by imposing an interest charge when the derivative terminates.

To sum up, this Section has shown that a number of more ambitious reforms would eliminate the need for Section 1260, as well as other transaction-specific

---

[51] *See, e.g.*, I.R.C. § 1033 (providing a rollover rule for involuntary conversions). Tacking holding periods is the norm in tax-free reorganizations. I.R.S. Pub. 544, *Sales and Other Dispositions of Assets*, 35 (Feb. 7, 2023), https://www.irs.gov/pub/irs-pdf/p544.pdf [https://perma.cc/TB7Q-URAR] ("If you acquire an asset in exchange for another asset and your basis for the new asset is figured, in whole or in part, by using your basis in the old property, the holding period of the new property includes the holding period of the old property. That is, it begins on the same day as your holding period for the old property.").

[52] The assumption here is that the hedge fund reinvests profits, instead of distributing them to investors, so these trading gains benefit from the rollover rule.

[53] This reform faces familiar political and administrative challenges. *See, e.g.,* David A. Weisbach, *A Partial Mark-to-Market Tax System*, 53 TAX L. REV. 95 (1999). Another issue is whether mark-to-market accounting is constitutional. The Supreme Court has granted certiorari on this issue. *See* Moore v. United States, *cert. granted*, No. 20-36122 (June 26, 2023).

reforms. In our view, although many of these would be advisable if crafted the right way, Congress is unlikely to enact them. Even so, policymakers should still be guided in part by the fundamental reform they prefer in deciding whether to target a planning strategy. Again, hedge fund derivatives are likely to be more troubling if one's ideal is mark-to-market accounting, instead of a consumption tax.

C. Efficiency: Risks of Overbreadth

In setting the normative presumption, policymakers should consider not just how to tax the targeted transaction, but also how to avoid "good" transactions that don't involve the same abuse. In other words, policymakers also need to worry about overbreadth.

1. *False Negatives Versus False Positives*

This means policymakers have to walk a fine line. On the one hand, if a rule is too narrow, it misses some abusive transactions that should be covered. These "false negatives" have familiar costs, discussed above, in cutting the taxes of wealthy households, putting a premium on sophisticated tax advice, and the like.[54]

On the other hand, if a rule is too broad, it sweeps in "good" transactions that should not be covered. These "false positives" can be costly in overtaxing–and, potentially, discouraging–socially useful activity.

In managing this tradeoff, policymakers have to tolerate some false positives to avoid false negatives, and vice versa. As they strike this balance, they need to make context-specific judgments about the relative costs of these different errors. Ultimately, the goal should be to minimize the total cost from both types of errors.

2. *Costs of False Positives*

The costs of false negatives are, in effect, the costs of allowing the relevant tax planning strategy to continue, which we already have surveyed above.[55] Yet what about false positives? How costly is it to cover (and, thus, overtax) transactions that do not pose the relevant abuse? To make this judgment, policymakers should analyze two issues, which we consider in turn: first, the likelihood of reaching the wrong transactions; and second, the cost of doing so.

a. Probability

So, starting with probability, which "good" transactions are at risk of being covered? The key question is how similar an abusive transaction is to others that are not abusive. For example, how similar is a hedge fund derivative to an S&P 500 index future? Or to derivatives on gold or oil, which give investors greater diversification?

Relatedly, are there notable differences between the abusive and "good" transactions, which a rule can use to distinguish them? For example, can a line be drawn based on who is engaging in the transaction (e.g., high-net worth individuals, corporations, etc.)? Or who is offering it (e.g., public exchanges, tax-indifferent counterparties, etc.)? This sort of distinguishing characteristic is more reliable in

---

[54] *See supra* Part IV.A & B.
[55] *See supra* Part IV.B.

some contexts than in others. Some tax planning strategies stand out, while others look a lot like conventional business transactions. Indeed, truly gifted tax advisors make tax planning look like "business as usual."

The risks of overbreadth turn not just on the similarity between the planning and other transactions, but also on the breadth of the government's response. Needless to say, the broader the rule, the greater the risk of burdening transactions far removed from the targeted planning. For example, if you rent out a room in your home, you don't have to worry about a (narrow) rule on hedge fund derivatives, but you probably are covered by a (broad) rule on losses from passive investments like mink farms.[56]

Even so, when policymakers seek to avoid overbreadth, they have a motivated ally: the tax bar. Tax advisors have a stake in protecting "good" transactions, so clients can keep doing them. The bar also expects (or, at least, hopes) that policymakers will sympathize with this goal.

In contrast, the dynamic is quite different for false negatives. If there is an easy end run around a rule, which policymakers have missed, tax advisors have an obvious reason to stay silent: until this hole is plugged, their clients can use it.

b. Magnitude

In assessing the risk of overbreadth, policymakers should consider not just the likelihood of false positives, but also their cost. How problematic would it be to tax these transactions more heavily? How socially valuable are they? Would taxpayers really stop doing them? Again, policymakers should consider elasticity and externalities in deciding what the tax burden on these "good" transactions should be.

Based on this analysis, if an over-broad rule would threaten a large volume of socially valuable activity, policymakers should go to greater lengths to avoid false positives, even at the cost of allowing more false negatives. In contrast, an overbroad rule is less of a concern if the costs of overbreadth are low. Maybe these transactions are of only marginal importance to the economy. Or maybe the relevant activity is important, but taxpayers could still pursue it without being caught by the new rule.

3. Managing the Tradeoff: False Negatives Versus False Positives

In short, in setting the normative presumption, policymakers need to consider not just how bad the targeted transaction is, but also how bad the collateral damage would be from targeting it imprecisely. If false negatives are the more daunting prospect, policymakers should err in the direction of overbreadth. But if false positives are the more troubling scenario, a narrower rule is needed.

---

[56] The passive activity loss rules of Section 469 are notoriously overbroad. *See* Justyana Mueller & Susanna Forbes, *I'm Not Your Friend—I'm a PAL! Passive Activity Loss Rules Explained*, James Moore (July 23, 2019), https://www.jmco.com/articles/tax/passive-activity-loss-rules-explained/ [https://perma.cc/5UNN-GPPX] ("Sounds like a tax nightmare, right? Unfortunately, the situation is not uncommon.").

What if false negatives and false positives are both very costly? When this is the case, there is more pressure to craft a precise rule, which is better at distinguishing the targeted planning from "good" transactions.

Yet this precision isn't free. Usually, a sophisticated rule is needed, which increases compliance and enforcement costs (e.g., in requiring complex calculations, challenging valuations, subtle distinctions, and the like).

As a result, policymakers face a tradeoff: On the one hand, a more sophisticated rule targets the abuse more precisely but is harder to administer. On the other hand, a blunter rule is cheaper, but it is likely to be either overbroad (with many false positives) or easy to avoid (with many false negatives). So is it better to increase administrative costs? Or to tolerate more false positives and false negatives? There is no one-size-fits-all answer. Instead, policymakers need to make context-specific judgments.

In principle, this tradeoff can be avoided (or at least mitigated) with a test that is both accurate and administrable. To help policymakers come up with this sort of test, we propose two more steps in our three-step process. In the second step, preliminary filters should be used to lower administrative costs. In the third step, sophisticated tests should be used to draw accurate distinctions. The next two Parts consider these steps in turn.

V. STEP 2: PRELIMINARY FILTERS

Part IV showed that when policymakers learn about a new type of tax planning, the first step is to decide whether it warrants a response. Assuming it does, policymakers should seek to cover all the relevant variations, without burdening "good" transactions. At the same time, the response also needs to be administrable.

Striking this balance between precision and administrability is the job of our second and third steps. In Part VI, we argue that a sophisticated test, which minimizes false positives and false negatives, should be used as the third step.

But first, this Part recommends a second step: to economize on administrative costs, the sophisticated test should apply to only a subset of transactions. To hone in on the right ones, policymakers should use what we call "preliminary filters." Instead of imposing tax liability, these initial tests merely determine whether a more rigorous test is warranted. As a result, they can be fairly crude, turning on indicators that correlate–sometimes only roughly–with the targeted abuse. The presence of these factors triggers further review, while their absence is a safe harbor.

This Part suggests two types of preliminary filters: one focuses on characteristics of the taxpayer, and the other on characteristics of the transaction. While these characteristics are likely to be relevant for various transaction-specific reforms, this Part focuses on Section 1260 as an illustrative example.

This Part concludes by highlighting challenges in administering preliminary filters, including: how to define the scope of the relevant transactions; how filters should evolve over time; and how they should interact with each other.

A. Taxpayer-Based Filters

Let's begin with qualities of the taxpayer. In focusing on who participates in the relevant transaction, filters can consider three qualities: income; assets; and the counterparty.

1. Taxpayer Income

Arguably, the best way to hone in on sophisticated planning is with an income test. For example, policymakers could provide that a provision like Section 1260 applies only to taxpayers with incomes above a specified level (e.g., the top 1%, whose adjusted gross incomes are above $550,000). The good news is that an income test is both easy to administer and well-targeted.

a. Administrability Advantages

After all, taxpayers already compute their income and the government already monitors it. As a result, an income test adds only modestly to compliance and enforcement costs.

b. Relevance of Income: Higher Stakes

Income is not just easy to track. It also is quite relevant for three reasons. First, in a progressive tax system, high-income taxpayers are supposed to pay more tax. Ensuring that they do advances distributional goals.

Second, these taxpayers also are an especially promising source of tax revenue. For the same reason that bank robbers target banks–"that's where the money is"–tax collection focuses especially on high-income taxpayers.[57] Since their income is taxed at a higher rate, their efforts to defer it or to convert it to capital gain cost the government more revenue.[58]

Third, for the same reason, these taxpayers are more motivated–and, therefore, more skilled–at tax planning. They can reap greater savings from a successful strategy, so they are willing to spend more on sophisticated tax advice and tailored transactions. These taxpayers also have access to planning tools that are not available to other taxpayers, such as the over-the-counter derivatives market.

Since their planning is more sophisticated, blocking it is more likely to require rigorous tests, which are costly to apply and enforce. These are precisely the types of rules that benefit from a preliminary filter.[59]

---

[57] When asked why he robbed banks, depression-era bank robber Willie Sutton supposedly replied, "that's where the money is." Thomas J. Bernard, *Willie Sutton*, ENCYCLOPEDIA BRITANNICA (June 26, 2023), https://www.britannica.com/biography/Willie-Sutton [https://perma.cc/7TY2-XKX7].

[58] Although the rates for both ordinary income and capital gains are progressive, ordinary rates are higher for very-high-income taxpayers, so the differential is wider. For taxpayers in the top bracket in 2023, converting ordinary income to capital gains changes a 37% tax into a 20% tax. In contrast, taxpayers in lower brackets might change a 24% tax into a 15% tax, so the differential is only 9%, instead of 17%.

[59] This is not to say that tax planning by other taxpayers is not worth policing. Since they are a larger group, they still owe a lot of tax in the aggregate. While this is true of taxpayers earning $150,000 per year, taxpayers earning $60,000 are unlikely to owe much (if any) income tax.

c.   Which Income?

While the tax system already tracks income, a few tweaks would still be needed for an income-based filter. For example, the test arguably should not turn on a single year, but on a rolling average. This way, taxpayers would not be "caught" just because their income spikes in a particular year (e.g., from the sale of their home or business). Likewise, taxpayers should not be able to avoid a filter by artificially reducing their income in one year (e.g., by accelerating or deferring it).

Taxpayers also should not be able to ignore the effect of the relevant transaction. For example, assume that Congress decides to target hedge fund derivatives, but to limit the statute to high-income taxpayers. How should this filter measure income? Specifically, should it count income taxpayers would have had if not for the derivative? If they had invested in the underlying fund instead, they would have had more (current) income. Should the filter count this deferred income? In our view, the answer is "yes." In determining whether a reform might apply, a filter should tentatively assume that it does apply. This approach helps the filter do its job, which is to identify taxpayers who require more vetting. Someone who would have significant income, as long as the rule applies, warrants a closer look.

2.   Taxpayer Assets

In some circumstances, policymakers may wish to look not just at income, but also at assets. By analogy, the securities law looks at both in defining an "accredited investor" who is allowed to invest in securities that are not available to the general public.[60] Indeed, a preliminary filter can simply use the securities law test.[61]

Admittedly, a test based solely on income often is adequate–and, indeed, preferable. Adding assets usually is redundant (since asset-rich taxpayers usually have high incomes), but costly (since the tax system doesn't already track assets).

Yet testing income isn't always enough. Some taxpayers actually do have significant assets but low incomes (e.g., when they have significant business losses or highly appreciated assets). Like high-income taxpayers, these asset-rich taxpayers have ample incentive and capacity to engage in sophisticated tax planning, and thus warrant more careful vetting.

3.   Counterparty

To hone in on the right transactions, a filter can focus not just on taxpayers, but also on their counterparties. The type of counterparty matters for four reasons.

First, when the counterparty is a financial intermediary, it has expertise to help taxpayers comply with a sophisticated rule. For example, if the counterparty on hedge fund derivatives is a securities dealer, it can help taxpayers do valuations and other calculations needed for rigorous tests (like those discussed below in Part

---

[60] U.S. Securities and Exchange Commission, *Accredited Investor*, www.sec.gov/education/cap italraising/building-blocks/accredited-investor [https://perma.cc/GCY5-C75R].

[61] An asset test might be relevant for taxpayers who have significant assets but have low taxable income (e.g., because they fund consumption by borrowing against appreciated assets instead of selling them, they have net operating losses, they are retired, etc.).

VI). But this usually won't be the case if the counterparty is the taxpayer's neighbor.

Second, the type of counterparty also can shed light on whether a deal is tax motivated. For example, compare the following two forward contracts. In the first, a taxpayer commits to buy a hedge fund interest from a securities dealer. In the second, one business partner commits to buy out the other and needs time to raise the cash. In principle, Section 1260 could apply to both,[62] but the identity of the counterparty shows that the second is probably motivated by conventional business reasons.

Third, and relatedly, a transaction is more likely to be tax motivated when the counterparty is subject to different tax rules. As noted above, this is how hedge fund derivatives became a viable tax planning strategy: although wealthy individuals did not like the tax costs of investing hedge funds, securities dealers were immune to these costs (because they mark their inventory to market).[63] As this example shows, some tax planning works only when the counterparty is subject to different rules. Like securities dealers, insurance companies, foreigners, and tax-exempt organizations also can play this role. As a result, deals involving these "tax indifferent counterparties" arguably should receive extra scrutiny.

Fourth, the same is true of related parties. It is well understood that in transactions between a subsidiary and parent or between a mother and daughter, the parties' economic interests usually are aligned enough that they can join forces to reduce their combined tax bill–for instance, by shifting income to the one with the lower tax rate. Instead of fretting about whether the pre-tax terms of their deal are fair, they are free to structure the deal to minimize their combined tax liability. Given this risk, the presence of related parties is another factor that can justify a closer look.

B. Transaction-Based Filters

In determining when to apply a more sophisticated (and costly) test, the tax system should consider not only who the taxpayer and counterparty are, but also what they are doing. In other words, there also should be filters based on characteristics of the transaction.

These filters should define categories of transactions that lie well beyond a transaction-specific reform's scope. The goal is to exempt these transactions not just from the reform itself, but also from careful vetting.

Which transactions should be excluded? As noted above, there is no need to target a planning strategy that brings the tax burden closer to where it actually should be.[64] Alternatively, even if a planning strategy is problematic, stopping it

---

[62] If the business is organized as an S-corporation, the second contract arguably triggers Section 1260. *See* I.R.C. § 1260(d)(1) (defining forward contract as "contract to acquire in the future (or provide or receive credit for the future value of) any financial asset"); I.R.C. § 1260(c)(1)(A) & (c)(2) (defining "financial asset" to include not only hedge funds, but also equity interests in other pass-through entities, including S-corporations, partnerships, REITs, and trusts). The Treasury also has regulatory authority to treat stock in a C-corporation and debt as financial assets, but so far has not used this authority.

[63] *See supra* Part II.A.

[64] *See supra* Part IV.A & IV.B.

might pose too great a risk to "good" transactions or prove too costly to administer.[65]

To operationalize these insights, this Section suggests seven transaction-based filters. For any given reform, some will be a better fit than others. Policymakers should make context-specific judgments about which filters to use.

1.   *Tax Treatment is Not Objectionable*

First, when the tax burden on a transaction is inappropriately high, there is no need to block–or, in some cases, even to vet–planning strategies to reduce it. Arguably this is the case with short sales, which are bets that the value of an asset will decline.[66] Unlike other investments, short sales are never eligible for long-term capital gains rates, even when they last for more than a year. Yet discouraging short sales is a bad idea. Without them, market prices may not fully reflect negative information and pessimistic views. In other words, taxing pessimistic bets more than optimistic ones can have negative externalities.[67]

Yet short sales are not the only way to bet against the market; instead, taxpayers can use derivatives, which are eligible for long-term rates.[68] Should Congress target this use of derivatives, using a transaction-specific reform to tax them like short sales? The answer should be "no" if the treatment of derivatives is more appropriate. If this is the case, the better course is to fix the treatment of short sales.

2.   *Tax Treatment is Well Settled*

Second, vetting a tax planning strategy is unnecessary not just when it is socially useful, but also when it is well accepted. For example, if a business is taxed as a partnership, the partners are taxed on their share of its income. But if the business is taxed as a corporation, investors usually aren't taxed until they receive a dividend or sell the stock (though the corporation itself pays tax). This inconsistency is a feature, not a bug.[69] Taxpayers are largely free to make this

---

[65] *See supra* Part IV.C.

[66] In short sales, investors borrow an asset, such as a share of stock, and sell it. At some point in the future, they have to buy the asset so they can return it to the person who lent it to them. So in effect, short sales reverse the usual order of investing by selling first and buying later. Short sellers make a profit by selling high (initially) and buying low (later). *See* Michael R. Powers, David M. Schizer & Martin Shubik, *Market Bubbles and Wasteful Avoidance: Tax and Regulatory Constraints on Short Sales*, 57 TAX L. REV. 233 (2004).

[67] *Id.* (The fact that "long" bets are taxed more favorably than "short" bets has the potential to distort prices). This is not to say that short sales are *always* socially valuable. *See* Joshua Mitts, *Short and Distort*, 49 J. LEGAL STUD. 287 (2020) (short sellers can profit by spreading rumors, using assumed names).

[68] For example, if a taxpayer buys a put option (i.e., an option to sell), and cash settles it at a profit over a year later, this gain is long-term.

[69] The Treasury made this choice explicit in "check the box" regulations finalized on January 1, 1997. *See* T.D. 8697, 1997-1 C.B. 215. Some organizations are corporations. Treas. Reg. § 301.7701-2. With others, taxpayers can option for pass-through treatment. Treas. Reg. § 301.7701-3.

choice, so there usually is no need for rules to recharacterize partnerships as corporations and vice versa.[70]

3. Tax Treatment is Widely Available

Third, there is less need to target a planning strategy not only when the favorable treatment is appropriate or well settled, but also when it is widely available. For example, unlike a hedge fund derivative, a derivative on a growth stock just replicates–but generally does not improve upon–the tax treatment of the underlying asset: each offers tax deferral, as well as reduced rates for long-term investments.[71] So the rationale for a rule like Section 1260–preventing derivatives from offering better tax treatment–does not apply to derivatives based on growth stocks.[72]

4. Transaction is an Established Commercial Practice

Fourth, transactions also should be exempted if they are common and not tax-motivated. Otherwise, a tax increase on these "good" transactions could discourage economically valuable activity, as noted above.[73]

For example, a standard way to bet that an asset will appreciate is a call option, which entitles an investor to buy the asset (e.g., 100 shares of XYZ stock) for a set price (e.g., the current price of $10 per share). Yet as any derivatives trader will tell you, a call option can be replicated by buying a smaller position in the underlying (e.g., fewer than 100 shares), and constantly adjusting this position's size as market conditions change.[74] Indeed, when derivatives dealers sell a call to an investor, they usually hedge it with this sort of "dynamic hedging."

Notably, these two alternatives–a call option or a dynamic position in the underlying–offer different tax treatment. Since the dynamic position requires frequent trading, its tax bill is usually higher. So should the tax law recharacterize the call option, treating it as a dynamic position in disguise? After all, when an investor buys a call from a securities dealer, the dealer hedges it dynamically. Should these trades be attributed to the investor? This is what Section 1260 does with hedge fund derivatives. Should it do the same with a call option on common stock?[75]

Our answer is "no." The difference, again, is that call options are a well-established commercial practice. Investors buy and sell them all the time for familiar economic reasons. The investors' goal is not to avoid unfavorable tax

---

[70] Even in this context, though, the government has at times limited the use of the "check the box" regulations. For example, the government has focused on their use in cross-border tax planning. *See generally* Monica Gianni, *International Tax Planning After Check-the-Box*, 2 J. PASSTHROUGH ENTITIES 39 (1999).

[71] Notably, the analysis of dividend-paying stocks could be different, since dividends trigger a current tax on the stock, but not necessarily on a derivative based on this stock.

[72] Section 1260 provides regulatory authority to cover derivatives based on common stock, but the Treasury has not used it. *See* I.R.C. § 1260(c)(1)(B)(2).

[73] *See supra* Part IV.C.

[74] The precise number depends on how closely changes in the option's value track changes in the stock price, and this correlation varies with the stock price.

[75] Congress did grant regulatory authority to target derivatives based on common stock. But even if this authority is used, an at-the-money call option generally isn't covered by the provision.

treatment from dynamic hedging–a trading strategy that, frankly, most investors don't understand–but to pursue a longstanding commercial goal.

Yet even as we recommend sparing practices that are commercially common, we acknowledge a problem with this approach. Implicitly, it protects favorable tax treatment for old practices, but not necessarily for new ones that have not had time to become common.

Unfortunately, this bias against new practices can stifle innovation. After all, what is commercially common now was once new. Years ago, this practice was able to develop–and become commercially common today–because its tax treatment was not too unfavorable. Otherwise, important innovations like index funds and ETFs would not have succeeded. So even for novel practices, some restraint is warranted in considering whether to tax them more heavily.

Again, the normative presumption is relevant here. If policymakers think financial innovation (and, more generally, established commercial practices) are valuable, they should be careful not to overtax them. But if they doubt the value of these practices–and conclude that their main appeal is their tax advantage–a higher tax is warranted.

5. *Transaction is Commercially Impractical*

Fifth, just as less scrutiny is needed for transactions that are commercially well established, the same is true of transactions that are commercially impractical, but for a different reason: even if taxpayers want to do something tax-motivated, they can't do it anyway. After all, a tax strategy is not appealing if its non-tax costs (or "frictions") exceed its tax benefit, so a reform isn't needed to block it.[76]

For example, given the way Section 1260 is drafted, it can be avoided with a derivative that offers most, but not all, of the economic return from a hedge fund. But, although this sort of imperfect tracking is easy for derivatives based on publicly-traded assets, it is much harder for derivatives based on assets (like hedge funds) that aren't publicly-traded: dealers can't hedge it, so they won't be willing to offer it.[77] In this situation, a legal response arguably isn't necessary.

6. *Transaction is Administratively Cumbersome to Police*

Sixth, even if a legal response is the only way to stop a tax planning strategy, a response isn't always worth the effort: in some cases, the administrative costs of identifying these transactions are too high.[78] As emphasized above, policymakers should balance these costs against the distortions and distributional effects caused by the transaction.[79] If the costs of stopping a planning strategy aren't justified, policymakers shouldn't target it, and can even use a preliminary filter to exempt it. Again, in making these judgments, policymakers should be guided by the normative presumption.

---

[76] *See* Schizer, *Frictions, supra* note 4, at 1319-33.

[77] "Dynamic" hedging, the strategy described above in which dealers constantly adjust the size of their position, is feasible only for publicly-traded assets. *Id.* at 1372-90.

[78] *See supra* Part III.B.1.

[79] *See supra* Part IV.C.3.

7. Transaction is Already Policed By Other Rules

Finally, even if an abuse is clearly worth preventing, a new rule isn't needed if an existing rule is already doing the job. For example, derivatives are not the only way to attain deferral and long-term capital gain for an actively traded portfolio: taxpayers also can trade through an offshore corporation. Yet Section 1260 does not have to address this strategy because other rules, including the PFIC regime, already target it. [80]

To sum up, preliminary filters can be used to winnow out transactions that do not warrant the expense of a rigorous test. In some cases, the tax treatment is appropriate, well settled, widely available, or already blocked by frictions. In other cases, the abuse is too costly to stop or already is policed with other rules.

C. Defining the Scope of the Relevant Transaction

While we believe that preliminary filters can reduce administrative costs, they cannot fix every problem. An important one that they don't solve is the need to define the relevant transaction. This challenge arises whenever a transaction must be tested, whether in the preliminary filters discussed in this Part or in the sophisticated tests discussed below in Part VI.[81] What counts as part of the transaction? What lies outside its scope?

This is a key issue in transaction-specific reforms, since they test–and then change the treatment of–particular transactions. Often, the issue is whether one transaction is enough like another. In Section 1260, for example, the treatment of a hedge fund derivative is recharacterized if–and only if–it is similar enough to the underlying fund. Like Section 1260, many rules require a comparison of two transactions: the one taxpayers actually do (e.g., the derivative) and the hypothetical one they could have done (e.g., the direct investment in the fund).[82] For clarity of exposition, let's call the transaction that taxpayers actually do the "actual transaction," and the one they could have done the "alternative transaction."

Yet the scope of these transactions isn't always self-evident. What specifically should be compared to what? Which cash flows are included? Which are not? This can be a very difficult question because cash flows can be packaged in different ways.

This reality can help tax planners avoid a transaction-specific reform. In some cases, they graft the planning strategy onto another transaction. In making the necessary comparison, can taxpayers use the whole thing? Or just a piece of it? In other cases, tax planners divide a transaction into components, which are formally separate. When should these components be analyzed separately? When should they be grouped together?

---

[80] *See* I.R.C. § 1297.

[81] *See infra* Part VI.C.1.

[82] There are many other examples as well. For instance, a derivative triggers tax withholding if it is enough like common stock. *See* Treas. Reg. § 1.871-15. A hedge causes a taxpayer to lose the dividend-received deduction if it reduces too much risk in the underlying stock. *See* Treas. Reg. § 1.246–5(c). The treatment of a hedge is recharacterized if it is enough like a sale. *See* I.R.C. § 1259. The treatment of offsetting positions is recharacterized if they are sufficiently offsetting, *see* I.R.C. § 1092, or if in combination they are too much like a debt instrument, *see* I.R.C. § 1258.

To keep taxpayers from manipulating the rule, policymakers need to give guidance on these questions. We raise this issue not to offer a definitive resolution, but to highlight that it is present here, just as it is present in countless other contexts.

1. *Scope of the Actual Transaction*

For example, under Section 1260 a forward contract to buy a hedge fund interest is covered, since it provides both the opportunity for gain and the risk of loss in the fund. In contrast, if the taxpayer acquires only the opportunity for gain, without the risk of loss, Section 1260 doesn't apply. As a result, an option to buy a hedge fund interest generally is not covered, as long as there is a meaningful possibility that the option will not be exercised. So if the fund interest is worth $100, and a taxpayer acquires an option to buy it for $100, this option won't be used if the fund's value declines below $100. This means the taxpayer is not exposed to the full risk of loss in the hedge fund.

But what if the taxpayer uses a second transaction to take on this risk of loss? For example, what if a day later she sells a put option on the hedge fund, obligating her to buy it for $100 if the hedge fund's value declines below $100? If the two options are evaluated separately, neither is similar enough to the underlying fund to trigger the statute. But if they are evaluated as a unit, Section 1260 would apply.

So, should they be evaluated separately or together?[83] Does it matter whether the taxpayer waits a few days (or a few months) before entering into the second transaction? What if the options are with different counterparties? What if it was not the taxpayer herself–but a fund in which she is a passive investor–that entered into one of these option transactions? In that case, would the taxpayer even know what the fund is doing? Rules are needed to define the scope of the actual transaction.

The issue arises in deciding not just whether to combine two formally separate transactions, but also whether to split a transaction into components. For example, what if a taxpayer enters into a derivative to buy an interest in two different hedge funds? As a whole, the derivative does not track either one. But if we treat it as two derivatives, each would trigger the statute. Should we bifurcate the derivative in this way? What if the derivative is based on the value of thirty-five different funds?

In answering these questions, policymakers face a tradeoff between blocking avoidance strategies, on the one hand, and minimizing administrative costs, on the other. To ensure that a rule reaches the targeted transaction and other similar transactions, at least some aggregation and bifurcation is needed. But requiring taxpayers to take these additional steps–and, for that matter, requiring the government to audit them–adds to the administrative burdens of the rule.

These burdens are especially unappealing in a preliminary filter, which is supposed to be easy to apply. After all, the whole idea of a filter is to keep things simple for most transactions, and to require more rigorous (and costly) scrutiny only of a subset of transactions. But if taxpayers can avoid the filter by either adding

---

[83] Section 1260 specifically requires options to be aggregated when they have "substantially contemporaneous maturity dates." I.R.C. § 1260(d)(1)(C).

to or splitting up the relevant transaction, the filter will allow too many false negatives. As a result, some degree of aggregation and bifurcation is needed at this stage, and policymakers should err on the side of including more transactions, so the more rigorous test can be applied to them.

One way to reduce the cost of requiring bifurcation and aggregation at the preliminary filter stage is to use more than one filter, and to ensure that at least one of the filters does not require this more burdensome analysis. For example, assume that Congress decides to use both a taxpayer-based filter (which exempts taxpayers with income below a minimum threshold) and a transaction-based filter (which uses aggregation and bifurcation in some circumstances when defining the scope of the transaction). Although the transaction-based filter is more costly to apply because it uses aggregation and bifurcation, the income-based threshold spares most taxpayers from having to apply it.

In any event, when policymakers decide how much aggregation and bifurcation to require, the normative presumption is relevant. Again, the more problematic the relevant planning is, the more motivated policymakers should be to block it.

2. Scope of the Alternative Transaction

This question of scope is relevant not just in the actual transaction, but also in the alternative one the taxpayer could have done. For example, assume a taxpayer enters into a derivative based on the S&P 500.

In this case, there are at least three alternative transactions: first, an investment in 500 individual stocks; second, an investment in a mutual fund or ETF that tracks the S&P 500; and, third, an investment in an "index future," which is a derivative based on the index.

So which is it? Although the answer is not self-evident–since these alternatives are economically comparable–it actually determines whether Section 1260 applies. On the one hand, if policymakers choose the first answer (individual stocks), Section 1260 doesn't apply (because it doesn't cover derivatives based on common stock).[84] On the other hand, if policymakers choose the second answer (a mutual fund or ETF), the statute does apply (because derivatives based on pass-throughs are covered).[85] If policymakers choose the third answer (an index future), the answer is unclear, arguably turning on whether the future is more like the stocks themselves or a mutual fund.

Notably, a lot is at stake in deciding whether index futures are covered. If they aren't, this can become a significant loophole. For example, instead of a derivative based on the value of a hedge fund, taxpayers could enter into derivatives based on an index that tracks the hedge fund's performance.

Likewise, many hedge funds use specific trading strategies that can be reduced to an algorithm. What if the derivative is based on the algorithm, rather than on the hedge fund itself?

Is it possible to cover these derivatives, but not more "plain vanilla" derivatives that are based on, say, the S&P 500? Arguably, this seems like the right

---

[84] Specifically, the statute covers common stock only under regulations, which Treasury has not promulgated as of this writing. *See* I.R.C. § 1260(c)(1)(B)(ii).
[85] *See* I.R.C. § 1260(c)(2)(A).

result, since the latter are conventional commercial transactions that usually are not tax motivated. But to accomplish that, the rule cannot require all indices (or, for that matter, none of them) to be covered. Presumably, the line should be drawn based on how common and widely followed an index is. Something that is customized–in effect, the trading strategy of a particular taxpayer–should be covered, even if more conventional indices are not.

But regardless of the answer to this specific issue, the more general point should be clear: rules are needed to define and characterize the alternative transaction. This is a daunting problem because there is no such thing as "the true underlying." The relevant cash flows can be packaged in different ways and, at least at a conceptual level, it's hard to view one as more authentic than the others.

3. *What is the Solution?*

But although this inquiry is difficult, it can't be avoided. So what should policymakers do? How should they define the scope of the relevant transactions?

As a start, policymakers can draw on general tax principles that deal with this issue. Admittedly, these principles are not always effective, but we are not seeking to add to or critique these rules here. Rather, we mean to piggyback on existing rules.

Yet these rules can be applied in either a tough or a lenient way. In deciding which approach to take, policymakers should be guided, once again, by the normative presumption in Part IV. When false negatives are a particular concern, aggregation, bifurcation, and other strict rules should be deployed more aggressively. But when false positives are a particular concern–so that an inclusive definition of scope would require vetting for a substantial number of "good" transactions–a more permissive approach is warranted.

D. Interactions Among Filters and Adjustments Over Time

The good news is that no single filter has to be perfect, since policymakers can use more than one. In addition, policymakers can refine them over time.

1. *Interaction of Filters*

As this Part has shown, filters based on taxpayers and transactions each have advantages and disadvantages. So which should be used? The answer is "both." When filters are cumulative, each one takes pressure off the others.

For example, if a taxpayer-based filter already screens out many transactions, there is less pressure on transaction-based filters to exempt transactions. Maybe these filters aren't needed at all? Or maybe they can be tougher, exempting only clear cases? Once one filter has already narrowed the field, others can fine tune at the edges, without having to exclude vast categories.

2. *Should Preliminary Filters Become More Lenient or Tougher Over Time?*

These judgments also can change over the years. A key question, though, is how to refine filters over time. Should they become more lenient or more tough? On the one hand, filters can start as narrow safe harbors–offering only limited exemptions from more rigorous vetting–and then broaden over time. On the other

hand, filters can move in the opposite direction, listing specific situations that require careful vetting, and then adding to this "naughty list" over time. Each approach has advantages and disadvantages.

a. Safe Harbors: More Lenient Over Time

A virtue in starting tough–that is, beginning with narrow preliminary filters and expanding them over time–is its information-forcing effect. Taxpayers are more willing to tell the government about false positives than false negatives, as noted above. Harnessing this impulse, this approach invites them to come forward about "good" transactions that aren't (yet) exempted.

But this approach has familiar downsides as well. If preliminary filters don't screen out many transactions, the more rigorous test (the third step in our three-step analysis) applies more broadly, increasing compliance and enforcement costs. Taxpayers have to invest more in this rigorous screening, and the government has to sift through more "good" transactions to find "bad" ones, which can be like searching for needles in a haystack.[86] Over time, the government can reduce these costs by gradually broadening the safe harbors, but this process can take a long time.

b. Listed Transactions: Tougher Over Time

Alternatively, starting off easy and getting tougher offers the mirror image of these costs and benefits. The government has to work harder to find the false negatives–that is, to identify "bad" transactions that are wrongly filtered out at the preliminary stage. Policymakers can try to add these transactions over time, but they are likely to be a step behind.[87]

This means policymakers have to pick their poison. At the preliminary filter stage, do they prefer too many false positives or too many false negatives? Like other aspects of this analysis, the answer turns in part on the normative presumption.

VI. STEP 3: THE ANALYTICAL STAGE

The last two Parts have proposed two steps to craft transaction-specific reforms like Section 1260. First, when policymakers learn of a new tax planning strategy, they should consider how problematic it is, as well as the collateral effects from targeting it. Second, to distinguish this planning strategy from "good" transactions that should not be targeted, policymakers should start with preliminary filters. These "quick and dirty" tests should exempt a large volume of transactions, narrowing the pool that requires more rigorous (and costly) vetting.

This brings us to the third step in our methodology. This "analytical stage" should use sophisticated tools to determine whether a transaction is covered. As a

---

[86] This problem has arisen, for instance, with the rules for so-called "listed" transactions–an effort to identify potential tax shelters.

[87] For example, the government has taken a long time to focus on potential abuses of Section 1260. Although the statute was enacted in 1999, some versions of these transactions were not identified as "of concern" for sixteen years. *See, e.g.*, I.R.S. Notice 2015-47, 2015-30 I.R.B. 76; I.R.S. Notice 2015-48, 2015-30 I.R.B. 77; I.R.S. Notice 2015-73, 2015-46 I.R.B. 660; I.R.S. Notice 2015-74, 2015-46 I.R.B. 663. It probably is no accident that these notices followed a congressional study of the issue.

case study, we propose a new test for Section 1260. We show why our test is better than alternatives others have suggested, highlight challenges in applying it, and show how it would apply in pending litigation.

A. More Analytically Rigorous Comparisons: Our Difference-Value Test

Once our first two steps have identified the actual and alternative transactions,[88] the fundamental question in Section 1260 (and, indeed, in many other transaction-specific reforms) is whether they are similar enough. If they are, taxpayers essentially are treated as if they engaged in the alternative transaction instead (i.e., investing in the underlying fund). This Section proposes a novel way to compare these two transactions, which has key advantages over other methodologies.

We frequently refer to the actual transaction as the "derivative" and the alternative transaction as the "underlying." This reflects the paradigmatic situation (like in Section 1260) where the actual transaction is a derivative contract because taxpayers are trying to use derivatives to get better tax treatment.

1. *Valuing Differences*

In essence, our approach is to identify differences between the derivative and the underlying, quantify the economic value of these differences, and then compare this value with the overall value of the underlying. If the discrepancy is modest, representing only a small percentage of the total value of the underlying, we would tax holders of the derivative as if they owned the underlying.

In ways, our difference methodology resembles a test, "the option-pricing methodology," which the NYSBA Tax Section proposed for a different rule: constructive sales under Section 1259.[89] This provision targets planning strategies that simulate a sale of an appreciated asset without triggering tax. The key question is whether taxpayers have gotten rid of too much economic exposure (e.g., to appreciated stock). In contrast, Section 1260–which uses language similar to Section 1259–asks whether taxpayers have taken on too much exposure (e.g., to a hedge fund).

The NYSBA's option-pricing test focused on fairly simple instruments, which are based on publicly-traded assets. In contrast, our methodology is more ambitious. It accommodates more complex instruments, including ones with price-based contingencies, and applies more broadly (e.g., to derivatives based on hedge funds).[90]

2. *Illustrative Example*

To illustrate our difference methodology, assume that Taxpayer wishes to invest in an underlying asset such as a hedge fund interest or common stock ("Underlying"). But instead, Taxpayer buys an option to purchase Underlying at

---

[88] This terminology is introduced in Part V.C., *supra*.

[89] *See* N.Y. STATE BAR ASS'N TAX SECTION COMM. ON FIN. INSTRUMENTS, COMMENTS ON H.R. 846, NYSBA TAX SECTION REP. #901 28-30 (May 21, 1997) [hereinafter "NYSBA, REP. #901"]. One of us (David Schizer) was a principal author of this report.

[90] For a discussion of the differences between our difference value approach and the NYSBA's option-pricing approach, see *infra* note 98.

any time in the next five years. This "call option" entitles Taxpayer to pay a discounted price: Taxpayer can buy Underlying for $70 (the so-called "strike" or "exercise" price), which is less than Underlying's current value of $100. This means that Taxpayers could make $30 if they were able to exercise the option immediately (paying only $70 for something worth $100).[91] An option that offers this sort of favorable pricing is known as "in the money."

What if Underlying's price falls below $70? Taxpayer does not have to buy it for $70. Unlike a forward contract, an option offers a choice or option–hence the name–either to use it or to let it expire.

Notably, there is another difference between Underlying and the call option: the timing of payment. To buy Underlying, investors have to make an up-front payment of $100 (the initial value of Underlying). In contrast, the option spares them from paying in full right away. Although they have to buy the option itself, they don't (yet) have to buy Underlying. The option lets them wait, so they pay $70 (the price specified in the option) only if and when they actually use the option. To make the cash flows equivalent, then, we assume that the taxpayer not only buys the option, but also buys a bond to fund the $70 purchase price.[92] Through this bond, the taxpayer will still receive $70, even if the call option expires worthless.[93]

3. Step #1: Identify the "Difference Contract"

The first step in our difference-value methodology is to compare the economic return on Underlying with the return on the derivative (which, in this case, is a call option). How are they different? How much of the opportunity for gain does the derivative provide? What about risk of loss?

Assuming these investments are different, how can investors fill the gap? What additional investment would they need? We call this investment the "difference contract." By definition, the return on Underlying can be perfectly replicated with a combination of the derivative and the difference contract (or contracts).[94]

In our illustrative example, the call option provides the same opportunity for gain as Underlying: any increase above the current value of $100 benefits a holder of this option, just as it benefits an investor in Underlying. In contrast, the call option only imposes a portion of the risk of loss–from $100 down to $70–but not the risk of loss below $70. (If the option expires worthless, they will still receive $70 from the bond.)

---

[91] Taxpayer would not be getting something for nothing in this situation. The price Taxpayer pays for the option is necessarily at least $30 when immediate exercise of the option would yield $30. For the sake of simplicity, we assume that the option is European, meaning that it can only be exercised upon expiration in five years. This assumption precludes the possibility of immediate exercise.

[92] Another way to align cash flows is to assume that a taxpayer who buys Underlying would borrow $70 of the purchase price.

[93] For the sake of simplicity, we assume that the asset does not pay dividends or make other distributions. We also assume that the risk-free interest rate is 5%.

[94] As discussed below, the difference contract may consist of more than a single simple option. This is the case, for instance, if the derivative omits both opportunity for gain and risk of loss. In that circumstance, the values of these contracts are not netted. Rather, to account for both risk of loss and opportunity for gain, the absolute value should be used. *See infra* Part VI.A.6.

The key difference, then, is that an investor in the option is protected from risk of loss below $70. What separate contract can offer this protection? The answer is an option to sell (a so-called "put option") entitling the holder to sell the underlying asset for $70 (so the "exercise price" is $70). What is this avoided exposure worth? The answer, of course, is the value of this put option–something that is feasible to compute.

4.   Step #2: Compute the "Difference Value"

This brings us to the second step in our difference methodology. Once the "difference contract" has been identified–a put option with an exercise price of $70 in this example–the next step is to value it. How much would someone pay for it?

Notably, all put options are not created equal. They are worth more when they are more likely to be used. All else being equal, the higher the exercise price, the more likely it is to be used (since it offers a better sale price). When Underlying is worth $100, a put option with an exercise price of $90 is worth more than one with an exercise price of $70, since Underlying's price does not have to decline as much for the option to become "in the money" (i.e., to offer a positive payoff when exercised).

For the same reason, a longer term generally also makes the put option more valuable. Compared with an option that lasts one day, an option that lasts ten years gives the price of Underlying more time to decline from $100 to below $70.[95]

Likewise, the volatility of Underlying is also important. The more volatile it is, the more likely it is to trade below $70 (and, again, to generate a positive payoff).

So what is the value of a put option with a five-year term and an exercise price of $70 (when Underlying is worth $100)? The answer varies with the volatility of the underlying asset.[96] Table 1 compares the value when volatility is low (20%) and high (60%), showing that the answer is quite different:

Volatility	20%	60%
Value of Difference Contract (the Put Option)	$1.31	$18.90

Table 1: Difference Contract's Value With Varying Volatility

---

[95] Notably, a longer term is even more valuable when the option is "American style," which means it can be exercised at any time. In contrast, a European style option–the kind we assume here–can only be exercised at maturity. As a caveat, though, the effects of an option's term can be complicated and, at times, a bit counter intuitive. For example, when the strike price of a European put is well above the price of the Underlying (a so-called "deep in-the-money" situation), the value of the option can become less than its intrinsic value (the value that would be obtained if immediate exercise were possible). In this case, a longer time to expiration is a disadvantage to the holder and corresponds to a lower option value.

[96] Under the Black Scholes model, interest rates also affect an option's value. In this example, the risk-free rate is assumed to be 5%.

5. Step #3: Compute the "Difference Percentage"

Once the difference value has been computed, it should be compared to the value of Underlying.[97] What percentage of the total does the difference contract represent? This ratio provides a precise way to measure the significance of the omitted exposure. A smaller percentage makes the derivative a closer substitute for the underlying, so Section 1260 is more likely to apply.[98]

In the example above, the Underlying is $100 when Taxpayer enters into the derivative contract. If the volatility is 60%, the difference value is $18.90. As a result, the "difference percentage" is 18.90/100 or 18.9%. Alternatively, if the volatility is 20%, the difference percentage is only 1.31/100 or 1.31%.

Is a difference of 1.31% small enough to trigger constructive ownership? What about 18.9%? How small must this percentage be? In other words, how similar must the contract be to the Underlying? In setting this threshold, Congress should be guided by the normative presumption, discussed above.[99] Depending on

---

[97] Notably, in our example (and in many other cases), we may take the size of the transaction to be the initial price of Underlying ($100). Yet in more complex situations in which payment for underlying would not necessarily be paid all up-front and may be contingent, it would generally be appropriate to take the full value of payments made into account. For example, the current price of all future payments needed to purchase Underlying could be determined, and this aggregate price could be used.

[98] Instead of using the value of the underlying as the denominator, there is another possibility, which the NYSBA proposed for Section 1259. *See* NYSBA, REP. #901, *supra* note 89, at 28-30. Section 1259 tests whether a hedge should be treated as a sale, evaluating how much risk of loss and opportunity for gain a taxpayer has *retained*. Section 1260 is the mirror image, asking how much the taxpayer has *transferred*. Either way, the effort is to assess how significant the relevant exposure is. But the NYSBA makes a different comparison than we do–in part because they are developing a narrow test for the (simpler) context of publicly-traded securities, while our test is more general and, therefore, more ambitious.

Like our test, the NYSBA options pricing approach computes the value of the difference between the contract and Underlying and puts it in the numerator. But their denominator is different. Instead of using the value of Underlying, as we recommend, the NYSBA does an additional calculation. In essence, they value the total risk-based exposure in Underlying: their denominator is the sum of the (absolute) value of the opportunity for gain in Underlying (measured with an at-the-money call option) and the risk of loss in Underlying (measured with an at-the-money put option).

This approach has one potential advantage over our "difference value" proposal. It strips out the capital invested, so that the test focuses on risk alone. As a result, the calculation no longer can be manipulated by either stuffing in or stripping out capital with a fixed return, a danger that must be managed in our approach.

However, this advantage comes at a significant cost: unlike our approach, the NYSBA's method requires an additional set of calculations (i.e., valuing total risk of loss and opportunity for gain in Underlying). This presumably seemed easy in the narrow setting in which they proposed this approach: hedging publicly traded stock. But this effort would be much more daunting in valuing hedge funds and their trading strategies. What is the value of the upside and downside in using a particular long-short strategy? This strikes us as a potentially intractable problem.

In contrast, our denominator is much easier to compute: it is simply the investment the taxpayer could have made. As a practical matter, this is likely to be the investment that the taxpayer's counterparty actually is making. Again, our approach requires the government to police efforts to manipulate the denominator with capital stuffing or stripping, but we think that on balance this is a more administrable approach than the NYSBA's methodology, at least in this context.

[99] *See supra* Part III.

policymakers' assessment of the relative harm from false negatives and false positives, they can set the cut-off at, for example, 5% or 10%.

With either of these cutoffs, volatility is a decisive factor in our example. Section 1260 is triggered if the Underlying's volatility is 20%, but not 60% (even though the contract has the same exercise price, duration, and other economic terms). This result makes sense. Again, for an in-the-money call option to diverge from Underlying, the latter's value has to fall below the $70 exercise price. A more volatile Underlying is more likely to trade below $70.[100]

6.   *Other Difference Contracts: Using Absolute Value*

As the above example shows, our methodology quantifies the omitted exposure on a derivative. Yet this omitted exposure can take different forms. In the example above, the derivative omits only risk of loss. This Subsection uses two other simple examples to illustrate how our methodology applies if instead the derivative omits only opportunity for gain or, alternatively, a combination of opportunity for gain and risk of loss.

a.   Omitting Only Opportunity for Gain

For example, assume that Underlying is still worth $100, and a taxpayer enters into a derivative that omits opportunity for gain above $150. Specifically, the derivative offers opportunity for gain up to $150, as well as all the risk of loss below $100. As in our last example, the derivative is paired with a bond, which in this instance pays $100.

To fill in the omitted exposure, the difference contract has to provide opportunity for gain above $150. This would be a call option to buy Underlying at $150. Our test would value this call option and compare it to Underlying's $100 value, yielding a result of 11.90 or 45.85 if the volatility of Underlying is 20% or 60%, respectively.[101]

b.   Omitting Both Opportunity for Gain and Risk of Loss

Instead of omitting only risk of loss or only opportunity for gain, a derivative can omit a portion of both. For example, assume it provides risk of loss below $95 and opportunity for gain above $115.[102] This derivative omits exposure to price changes between $95 and $115, which a direct investment in Underlying obviously would provide.

---

[100] If Underlying's price falls below $70 upon expiration, the taxpayer has $70. Although the call option is worthless, recall our assumption that the taxpayer also buys a bond, which pays $70 at maturity. So, if Underlying is worth $50, the investor still has $70 (from the bond). Thus, there is a divergence between the price of Underlying and the call (with bond) position, which is the excess of $70 over the price of Underlying (e.g., $20 when the price of Underlying is $50). In contrast, when the price of Underlying is at or above $70 upon expiration, the payout on the call option plus $70 has the same value as Underlying. For example, if Underlying's price is $110, the call generates $40 and the bond generates $70 for a total of $110.

[101] As in our prior example, we assume a risk-free rate of 5% and a term of 5 years.

[102] A taxpayer can get this exposure, for instance, by purchasing an option to buy the underlying for $115 (i.e., a long call option with an exercise price of $115), and by selling an option to sell the underlying for $95 (i.e., a short put option with an exercise price of $95).

To fill in these gaps, taxpayers actually need two difference contracts, instead of just one. The first would give them the opportunity for gain they do not already have (from $100 to $115). Since this is a way to make more money–an asset, not a liability–they would have to buy it. Specifically, they would buy a derivative known as a "call spread."[103]

Meanwhile, the second contract would force them to take on the risk of loss they are not already bearing (from $100 to $95).[104] Unlike opportunity for gain, this extra risk of loss is a way to lose more money. Since it is a liability–not an asset– they would be paid to take it on. They would sell a derivative known as a "put spread."[105] (Again, the taxpayer also would need to buy a bond, as noted above.)

Again, the two difference contracts would have to be valued (step two of the difference value methodology). Then, their values would be compared with the value of Underlying to compute the difference percentage (step three).

This brings us to an important point. From the taxpayer's perspective, the value of one of these contracts is positive (the call spread), while the value of the other is negative (the put spread). They pay for the first and get paid for the second.

Yet these values should not be netted against each other. What matters here is how different the derivative is from Underlying. As this example illustrates, some differences have positive value, while others have negative value. When this is the case, these differences should not cancel each other out. Otherwise, our methodology would understate the differences. In an extreme case–if the omitted opportunity for gain has exactly the same value as the omitted risk of loss–the (net) omitted exposure is zero.[106] This implies that the derivative is identical to Underlying, which clearly is not the case.

The solution to this problem is simple: disregard whether retained exposure is an asset or a liability–since this doesn't matter–and just add their values together. In other words, use the absolute value of each component when computing the difference value.

---

[103] A call spread actually is the combination of two options. Here, taxpayers would *buy* a call option entitling them to buy Underlying at $100 (to acquire all appreciation above $100), and then *sell* a call option, entitling their counterparty to buy Underlying at $115 (so the taxpayers give up all appreciation above $115). In combination, these two call options provide taxpayers with appreciation between $100 and $115.

[104] In general, we use the term difference contract to refer to the single contract that represents the overall difference between a derivative and the underlying. In this example, however, when we refer to two difference contracts, we are using the term to speak of two distinct components of the usual overall difference contract: omitted opportunity for gain and omitted risk of loss. This facilitates our discussion and enables us to talk separately about the aspects of each component.

[105] Like a call spread, a put spread also is a combination of two options. Here, taxpayers would sell a put option entitling their counterparty to sell Underlying to them for $100 (so the taxpayers are exposed to all risk of loss below $100), and then *buy* a put option entitling them to sell Underlying at $95 (so the taxpayers avoid risk of loss below $95). In combination, these two put options expose taxpayers to risk of loss between $100 and $95.

[106] For example, imagine a derivative that provides opportunity for gain above only $115, while imposing risk of loss only below $56, so the derivative offers a final payoff of $100 in between $56 and $115. This is arguably very different from the economics of the underlying. But if we use netting, the value of the difference contract may be small or even zero. For example, in a Black-Scholes model with a volatility of 20%, a risk-free rate of 5%, and a term of 5 years, the call spread and the put spread both have a price of approximately $6.65.

Table 2 shows the results for this example. The key point is that the value of these two contracts should not be netted. The net value is misleadingly low (5.30 in the third line below): Instead, the absolute value should be used (8.00 in the fourth line below).[107]

	Value
Call Spread	6.65
Put Spread[108]	−1.35
Sum	5.30
Difference Contract **(Sum of Absolute Values)**	8.00

Table 2: Values of Components and Overall Difference Contract for Derivative Omitting Both Risk of Loss and Opportunity for Gain

Notably, although our methodology needs to be able to cover derivatives that omit both risk of loss and opportunity for gain, they are less common for hedge funds than for other types of underlying assets. Indeed, they are a staple of another form of tax-planning: hedging appreciated assets, without triggering a sale for tax purposes. But they are largely unavailable for hedge funds. Why the difference? As one of us has emphasized elsewhere, a bank can easily offer this sort of derivative when the underlying asset is publicly traded (since the bank can hedge dynamically), but not for an underlying (like a hedge fund) that isn't publicly traded.[109]

7. *Valuing the Difference Contract: Alternative Methods*

So far, this Section has developed a methodology that depends on valuing what we call "difference contracts." But how should these hypothetical contracts be valued? When comparable contracts are available in the market, market value should be used.

When there is no readily ascertainable market price, the value can be determined using familiar techniques for valuing illiquid assets. Since the difference contract is likely to include options (and option spreads), valuations based on theoretical models like Black-Scholes are likely to feature in this effort, as in the example above.

---

[107] These are the approximate values for the put and call spread using the Black-Scholes pricing framework with volatility 20% for Underlying, a risk-free return of 5%, and a contract term of 5 years.

[108] We write the value of the put spread as negative because it represents downside exposure that must be added to the difference contract to replicate the underlying. By contrast, the call spread is positive because it represents upside exposure.

[109] *See* Schizer, *Frictions*, *supra* note 4.

In general, the difference contract's value should be determined in a commercially reasonable manner in light of the facts, perhaps informed by expert opinions. These opinions should be explicit about their assumptions, so the analysis can be tested and challenged.

A court assessing the value would make credibility determinations and judgments as it would for any determination of fair market value. Indeed, one of the advantages of our approach is that it puts courts in a familiar position. Valuation disputes arise quite frequently in litigation, so courts know how to adjudicate them.

Admittedly, valuations can be costly and disputes about them can be daunting to litigate. But again, this is why we recommend preliminary filters to reduce the number of cases that require these valuations.

8. Choosing the Underlying and Risks of Manipulation

The essence of the test we recommend is, of course, to compare the derivative with the underlying. Yet as emphasized above, to compare two transactions, we need to define the scope of each of them. What should be compared to what?

Our methodology does not avoid this question–and, of course, alternative methodologies don't either. Any comparison, however sophisticated, must first define what is being compared.

In our view, the right moment to address this issue is the preliminary filter stage, as noted above. Guidance is needed, for example, about when to bifurcate a position, when to lump together positions that are formally separate, and the like. If these rules don't completely resolve the question, judgment needs to be exercised.[110] Since these choices can affect the outcome, the government has to police them.

B. Comparison with Alternative Tests

Our difference-value methodology is not the only test for comparing investments. At least four other methodologies, which were proposed for other statutory provisions, can be adapted to Section 1260:

- First, a simple test is based on the amount of exposure or "spread" that one position offers but the other does not (the "spread" test).[111]

- Second, a more complicated test compares how much the value of one position changes when the other position's value changes by a

---

[110] For example, as noted above, we have matched the cash flows on the derivative and Underlying by pairing the derivative with a bond. Yet a different approach would be to assume that Underlying is purchased with borrowed funds.

[111] This test was proposed by the New York State Bar Association Tax Section for constructive sales under Section 1259, a topic closely related to constructive ownership. *See* NYSBA, REP. #901, *supra* note 89, at 28. The NYSBA's approach with both provisions was for taxpayers to avoid the rule by either retaining (in the case of Section 1259) or avoiding (in the case of Section 1260) an adequate amount of exposure. In both cases, they relied on the sort of "band" of exposure described in text, while caveating that the test assumed that the band includes the current price, the underlying is not unusually volatile, and the instrument will not last more than a specified term of years. The government picked up on aspects of this approach, at least for Section 1259, in Rev. Rul. 2003-7. 2003-1 C.B. 363.

dollar. Since this ratio is known as "delta," this is called the "delta test."[112]

- Third, there is a more sophisticated variation of the delta test, which the government implemented in regulations[113] after critiques were offered of prior proposed regulations.[114]

- Fourth, another test, applied by the Second Circuit, seeks to determine the probability that two positions will track each other.[115]

1. Spread Test

In comparing a derivative with an underlying, a "spread test" measures the differences in a crude way, identifying the range of prices in which they offer different returns. For an underlying worth $100 ("Underlying"), for instance, how much risk of loss does the derivative offer below $100? How much opportunity for gain above $100? How much, if any, of this exposure is omitted?

a. Size of the Spread

Before returning to our main illustrative example (the in-the-money call option), let's begin with one that is better suited to the spread test: the derivative, mentioned above, that provides risk of loss below $95 and opportunity for gain above $115.[116] This derivative omits exposure to price changes between $95 and $115, which a direct investment in Underlying obviously would provide.

This means that modest moves away from the current price of $100 affect an investor in Underlying, but not in this derivative. Indeed, if the price doesn't change, this derivative won't either make or require a payment. (In the parlance of derivatives traders, it is "out-of-the-money.")

Does this 95-115 price range (or "spread") render the two sufficiently different? To answer this question, the spread test asks what percentage this spread represents of Underlying's value.[117] When the NYSBA proposed this test, they recommended a minimum percentage of 20%, which is what this derivative provides ([115-95]/100).[118]

---

[112] This test also was proposed by the NYSBA for constructive sales. *See* NYSBA, REP. #868, *supra* note 12, at 18-22, and has also been refined and enhanced by academics. *See, e.g.*, Thomas J. Brennan, *Law and Finance: The Case of Constructive Sales*, 5 ANN. REV. OF FIN. ECON. 259 (2013).

[113] Treas. Reg. § 1.871-15.

[114] *See, e.g.*, Thomas J. Brennan & Robert L. McDonald, *The Problematic Delta Test for Dividend Equivalents*, 146 TAX NOTES 525 (2015).

[115] This test was proposed by the Second Circuit for constructive sales. *See* Estate of McKelvey v. Comm'r, 906 F.3d 26 (2d Cir. 2018).

[116] *See supra* Part VI.A.6.b. A taxpayer can get this exposure, for instance, by purchasing an option to buy the underlying for $115 (i.e., a long call option with an exercise price of $115), and by selling an option to sell the underlying for $95 (i.e., a short put option with an exercise price of $95).

[117] For this purpose, the NYSBA uses Underlying's value at the moment when the taxpayer enters into the derivative.

[118] The NYSBA offered this test for constructive sales. A taxpayer who *retained* exposure to gain and loss in a specified "spread" around the current price would not trigger a constructive sale. *See* NYSBA, REP. #901, *supra* note 89, at 28. In constructive ownership, by contrast, a taxpayer who *eliminated* exposure in this range would not trigger constructive ownership.

b. No Exposure Near Current Price: Derivative Must Be "Out-of-the-Money"

Notably, the NYSBA also required the omitted exposure to include Underlying's current price.[119] For example, when Underlying's price is $100, a derivative that omits exposure between $200 and $220 would not satisfy the test, even though the $20 of omitted exposure does, in fact, represent 20% of Underlying's current price ($100).

Why do the returns on the derivative and Underlying have to diverge near the current price? In other words, why must the derivative be out-of-the-money? The intuition is that these differences are more immediate–arising as soon as the price moves even a penny–so they are more significant. Instead of just potential differences, the spread causes actual differences.

In this example, a derivative that omits exposure from $90 to $110 is almost certain to offer a different return than Underlying; after all, Underlying's price inevitably will move away from $100, even if it doesn't move very much.

In contrast, if the derivative omits exposure between $200 and $220–or, for that matter, between $10 and $30–Underlying's price would have to change a lot (either doubling or losing 70% of its value) before differences between it and the derivative begin to matter. (In other words, these derivatives are "in-the-money."). As a result, fluctuations near the current price don't cause the returns on the derivative and Underlying to diverge.

c. NYSBA Spread Test Can't Analyze In-the-Money Derivatives

As the NYSBA emphasized, the main virtue of the spread test is that it is easy to administer. Yet as the NYSBA acknowledged, this simplicity comes at a cost.[120]

For one thing, this test puts too much weight on whether the spread includes the current price. In other words, it can't approve in-the-money derivatives.

Unfortunately, this requirement can generate arbitrary results. For example, consider two options to buy Underlying, which currently is worth $100, at any time during the next three years. One option entitles holders to pay $101, while the other entitles them to pay $99. While these options are quite similar–and both differ markedly from Underlying–only the first passes the spread test. Even though both have spreads that are larger than 20% of Underlying's current market ($100)–101% in the first ($101 to $0) and 99% in the second ($99 to $0)–only the first has a spread that includes this $100 price. In other words, only the first is out-of-the-money.

Because of this requirement, the spread test also flunks our illustrative example above: a call option entitling holders to pay $70 for an Underlying worth $100. Although the option and Underlying offer different returns when the price falls below $70–a 70% spread–this spread, once again, doesn't include the current $100 price.

But although the option to pay $99 and the option to pay $70 both fail this test, they are quite different. The option to pay $70 is a closer substitute for

---

[119] *See id.* at 28.
[120] *See id.* at 28

Underlying, since it exposes the investor to greater risk of loss as Underlying's price declines (i.e., from $99 down to $70).[121] Put another way, the option to pay $99 is only slightly in-the-money, while the option to pay $70 is much more in-the-money.

d. A Variation of the Spread Test: How In-the-Money Is the Derivative?

In principle, the spread test could be modified to account for this difference. The key issue is how much the price has to change before further changes stop affecting the derivative's payout. Put another way, the spread test could be adjusted to ask how in-the-money the derivative is.

Returning to our example, when an option entitles the holder to pay $99, Underlying's price has to decline by only $1 before further declines become irrelevant. This $1 decline represents only 1% of the Underlying's current $100 value (so this option is 1% in-the-money). In contrast, when an option entitles the holder to pay $70, Underlying's price has to fall by $30 or 30% (so this option is 30% in-the-money).

As this percentage increases, the option becomes a closer substitute for Underlying; like Underlying, it exposes the holder to more price changes near the current price. To implement this new variation of the spread test, policymakers can pick a percentage that triggers Section 1260 (e.g., 25%).

e. Other Limitations of the Spread Test: Volatility and Term

But even with this adjustment, the spread test is still a very blunt instrument. The problem is that all spreads with a specified size are not created equal. Giving up exposure to a range of prices matters more when the investment is likely to trade only within this range.

With a very stable investment, for instance, giving up the first 20% of appreciation may mean, in effect, giving up all the appreciation there is likely to be. In contrast, giving up 20% is much less meaningful with a very risky investment, whose price is likely either to double or to fall to zero. As these examples show, the more volatile an investment is, the less significant a particular spread becomes.

The same also can be true of the maturity of the derivative. Regardless of the underlying's volatility, it is more likely to trade outside the spread if the contract lasts for 10 years, instead of 10 days.[122]

So unfortunately, the NYSBA's spread test has several problems. In contrast, our difference-value test avoids these deficiencies. Our test works better because it focuses–not on the size of a spread–but on its value. As a result, our test is able to analyze in-the-money derivatives and, more generally, to account for

---

[121] The holders of both options lose money as the price falls to $99, but only the holder of the latter option loses money as the price falls from $99 to $70. Once the price falls below $99 (e.g., to $90), there's no longer a benefit to using an option to pay $99. Why pay $99 for something now worth $90? In contrast, there is still a benefit to using the option to pay $70. As a result, a price decline from $99 to $70 affects the return on an option to pay $70, but not on an option to pay $99.

[122] The risk-free rate of return is relevant as well.

volatility and duration. Admittedly, our difference-value test is harder to administer, but it avoids the spread test's limitations.

2. Delta Test: Basic Case

Our test also is better than another alternative, the "delta test," which the NYSBA and other commentators[123] have proposed and the government has tried to implement.[124] Like our approach, the delta test is harder to administer than the spread test, and is supposed to avoid its limitations. But unlike our approach, the delta test is easy to manipulate. This Subsection shows the delta test's advantages, while the next shows how to confound it.

"Delta" is a concept in finance, which measures how much the value of one position changes when the price of another changes by a dollar.[125] A delta of one means these changes are dollar-for-dollar.

In principle, Section 1260 can use delta to assess the similarity of a hedge fund derivative and the underlying fund interest. When a taxpayer enters into this derivative, what is its delta with Underlying? Is it close enough to one to count as constructive ownership?

To illustrate this test, let's return to our recurring example of a call option to buy Underlying for $70. When Underlying is worth $100, this in-the-money option is worth $46.79 (assuming Underlying's volatility is 20%).[126] What happens when the stock price increases by $1 to $101? The value of the call option increases by 95 cents to $47.74. In other words, the delta is 0.95 (i.e., 0.95/1.00).[127] Notably, the value of delta keeps changing as the price changes.[128]

How close to 1.0 must delta be to count as constructive ownership? To answer this question, policymakers need to designate a threshold, such as .90. The lower the threshold, the more transactions will be covered. In setting this level, policymakers should consider the equity and efficiency implications of hedge fund derivatives, the risks of an overbroad rule, and other factors that influence the normative presumption, as discussed in Part IV.

---

[123] Brennan, *supra* note 112; NYSBA, REP. #868, *supra* note 12, at 18-22   .

[124] *See* Treas. Reg. § 1.871–15 (testing dividend equivalent payments).

[125] In other words, delta is the ratio of the incremental change in price of the contract to a corresponding incremental change in the underlying asset price. The exact value of delta is the limit of this calculation for arbitrarily small changes in the price of the underlying. It is thus the derivative, in the sense of calculus, of the price of the contract with respect to the price of the underlying.

[126] We compute this value using Black Scholes with a 5-year term and 5% risk-free rate. Note that this is only the value of the call option component of the contract and does not include the price of the bond that pays $100 at expiration. For purposes of computing delta, the bond is irrelevant because its value does not change with the price of the underlying. Hence, it adds zero to the value of delta.

[127] This is a rounded approximation, with each call value rounded, and then the difference of the rounded numbers taken to compute delta. Computation without intermediate rounding would yield a result of 0.9442. In addition, this is a discrete approximation to delta. The true delta value is the limit of the result of such computations as the incremental adjustment to the underlying price tends to zero. That figure is 0.9430.

[128] The value of delta may also be interpreted as the amount of asset that would be necessary to hedge exactly the exposure to the asset represented by the contract. This hedge is as of the time of entry into the contract. In general, the amount of asset needed for a hedge, and hence the value of delta, changes over time.

The delta test works well for "plain vanilla" options with no contingencies. Unlike the spread test, it accounts for Underlying's volatility. As Table 3 shows, delta in our recurring example varies with Underlying's volatility:

Volatility	20%	60%
Delta	0.943	0.869

Table 3: Delta for Varying Volatility

The delta test correctly treats the low-volatility case as more similar to the underlying than the high-volatility case, and thus more likely to trigger constructive ownership. After all, when Underlying is not volatile, bearing risk of loss below $70 is less significant. So as this intuition suggests, if the delta threshold is .90, the low-volatility case triggers Section 1260, while the high-volatility case does not. Like volatility, the time left to maturity also affects delta, so a delta test is sensitive to this variable as well.[129]

While our difference-value test also accounts for these differences in volatility and term, the spread test does not, as emphasized above. In other words, the delta test and our difference-value methodology are both more nuanced and accurate than the spread test.

3. Delta Test: Barrier Options

Even so, a delta test has a problem of its own: unlike our difference-value approach, it is easy to manipulate. The problem is that delta is "hyper local," drawing comparisons at a specific price and moment in time. This quality makes delta especially sensitive to small changes in contract terms, even ones that are unimportant in the long run, and thus don't influence our difference-value approach.

a. Confounding the Delta Test With Contingencies

To exploit this hypersensitivity, taxpayers can add contingencies to their derivatives. They can engineer results under the delta test that are quite counter-intuitive–even a bit wacky.

To illustrate the point, let's add a contingency to our recurring example, a five-year call option to pay $70 for Underlying, which currently is worth $100. Unlike before, this call is now subject to a contingency: it can be used only after Underlying's price has declined below $98. In other words, this contingent call is activated if, and only if, Underlying's price declines from $100 to below $98 sometime during the option's five-year term. If this "knock in" contingency does

---

[129] For example, if the term of the call option is reduced from 5 years to 1 year, and volatility is 20%, the delta value is 0.984. This increase from 0.943 reflects the higher likelihood that the underlying price will end above $70 if only a 1-year term is allowed.

not occur, the option expires worthless.[130] An option with this sort of contingency is known as a "barrier option."[131]

What is the delta of this contingent call?[132] In principle, one might expect this contingency not to alter delta very much: after all, if Underlying is a risky asset (like a hedge fund interest), its price is almost certain to decline more two percent at some point over five years. Therefore, this contingency is unlikely to affect the option.

Nevertheless, this contingency has a thoroughly unexpected effect: it turns delta negative! Ordinarily, a call option appreciates when Underlying's value increases (e.g., from 100 to 101), since buying for a fixed price (e.g., of $70) becomes more appealing. But bizarrely, this contingency causes the call to lose value. Likewise, call options usually decline in value as Underlying's price falls, since buying for a fixed price (e.g., of $70) becomes less appealing. But weirdly, this contingency causes the call to increase in value.

So what is going on? Why has this contingency turned things upside down? The key point is that this barrier option can't be used unless Underlying's price falls below $98 at some point. If this never happens, the option is worthless. So as Underlying's price rises–moving further away from $98–the risk increases that the option will never be activated.

This means that increases in Underlying's price have competing effects. On the one hand, these increases tend to make any call option more valuable, as noted above. On the other hand, these increases tend to make this call option less valuable, since it is less likely to "knock in." When the latter effect dominates, increases in Underlying's price actually reduce the barrier option's value, turning delta negative.[133]

This counterintuitive effect is more pronounced when Underlying is less volatile, presumably because as Underlying's price moves further above the $98 threshold, it is less likely to come back down. As Table 4 shows, delta in the low volatility case actually is negative 1.52: when Underlying's price increases from 100 to 101, the call option's value actually declines not just by a few cents, but by a full $1.52!

Volatility	20%	60%
Delta	-1.520	-0.395

Table 4: Delta for Knock-In Call

---

[130] In this example, even if the option cannot be used, the taxpayer will still get $70. In effect, the bond portion of this transaction is not subject to this "knock in" contingency.

[131] James Chen, *What Is a Barrier Option? Definition, and Knock-In Vs. Knock-Out*, INVESTOPEDIA (Apr. 5, 2022), https://www.investopedia.com/terms/b/barrieroption.asp [https://perma.cc/3PJX-Z8 PT] ("A barrier option is a type of derivative where the payoff depends on whether or not the underlying asset has reached or exceeded a predetermined price.").

[132] A contingency that adds additional features to a derivative is called a "knock in" feature; in contrast, a condition that causes a derivative to expire prematurely is called a "knock out" feature.

[133] Likewise, declines in Underlying's price also causes competing effects. A call option generally loses value when the price of the underlying declines, but here the knock-in call comes closer to $98, and thus is more likely to become usable.

As this bizarre result shows, the delta test has utterly short-circuited. It is treating the call as very different from the Underlying, when this just isn't the case. Without the contingency, the call actually is a close substitute, as shown above.[134] Nor should the contingency make much of a difference. After all, it has no effect as long as Underlying's price declines to $98 at some point in five years–something that is extremely likely to happen.[135]

As this example shows, delta is not an effective test for instruments with payouts based not just on the final asset value, but also on the price path to get there. Delta is very sensitive to small effects at specific prices, which can obscure broader correlations. Put more colloquially, it often misses the forest for the trees.

b.   Potential Fixes For the Delta Test

Admittedly, there are ways to manage these problems with a delta test, at least to an extent. Perhaps a more sophisticated model could deemphasize contingencies, although we doubt it.

Alternatively, the test could ignore some contingencies, as other tax rules do.[136] So instead of analyzing what taxpayers actually did, the rule could test a simplified hypothetical version. The government took an analogous approach in regulations under Section 871, which use delta to police a strategy for avoiding withholding tax: investing in derivatives instead of common stock.[137] But testing a hypothetical deal has obvious problems. Among the many variations, which should be tested? What (potentially distorting) assumptions are taxpayers allowed to make?

In our view, these "fixes" seem more plausible when everyone is trying to make delta work, such as when traders hedge. But here delta is used not to finetune a portfolio, but to block tax planning. So instead of looking for a reasonable solution, some taxpayers will make aggressive assumptions to try to game the system.

---

[134] *See supra* Part VI.A.2 & VI.B.2 (showing similarity using difference value and delta tests).

[135] The risk-neutral probability of reaching $98 is 94.8% in our low-volatility case (20%) and 99.4% in our high-volatility case (60%). Note that risk-neutral probabilities are a construct of the pricing model used and are not the same as real-world (so-called "physical") probabilities. The risk-neutral values may be thought of as real probabilities adjusted to account for the risk aversion of a marginal investor, as reflected in prices. In general, real-world probabilities would correspond to a higher probability of an increase in asset price, and hence a lower probability of hitting the barrier. However, because we are interested in understanding the behavior of delta, which is also a construct of the pricing model, the risk-neutral probabilities are useful to consider.

[136] By analogy, regulations on various types of debt instruments specify how various types of contingencies should be treated. *See* Treas. Reg. § 1.1272-1(c) (specifying presumptions for contingencies in computing yields).

[137] Specifically, taxpayers who enter into complex contracts are required to construct a simplified alternative with a delta of .8. *See* Treas. Reg. § 1.871-15(h)(2). Unfortunately, the regulations fail to give clear guidance about *which* alternative to use (since a number of different ones could have a delta of .8). The regulations then direct taxpayers to compare the alternative–whichever it is–with the transaction they actually used. Specifically, taxpayers have to compare percentage changes in the delta of what they actually did with percentage changes in the simplified version. This comparison turns out to be not just exceedingly complicated, but also manipulable. Unfortunately, in the same way that knock-in and knock-out features allow for manipulation of delta, as discussed above, they also allow for manipulation of *percentage changes* in delta.

c. Key Advantage of the Difference-Value Test: Treatment of Contingencies

Fortunately, our difference-value test avoids this problem: "knock-in" and "knock-out" contingencies don't skew the results. The reason is that our test focuses on overall value, not on fleeting correlations. It considers the full range of outcomes, using probability weighting to measure the likely impact of contingencies. This keeps improbable ones from distorting the result. Again, to put the point more colloquially, our test compares forests, instead of individual trees.

To illustrate the difference between our test and the delta test, let's return to the contingent call option in the example above. Unlike its noncontingent counterpart, it cannot be used if Underlying's price never falls below $98. Yet this outcome is quite unlikely, as long as Underlying is volatile enough and the option's term is long enough. So should this contingency affect the option's value? The answer is "yes, but only modestly." This is exactly what we find with our difference-value approach. As Table 5 shows, the contingent call has a slightly higher difference percentage–that is, it is a bit less like Underlying–because of its "knock-in" feature.

Volatility	20%	60%
Difference percentage for non-contingent call[138]	1.31%	18.90%
Difference percentage for knock-in call[139]	9.22%	21.77%

Table 5: Difference Percentage for Simple Call and Knock-In Call

The point here is not to bash delta, but to clarify how it should be used. The key is that it focuses on an instant in time. This quality makes delta especially valuable for trading strategies that require constant adjustments, such as hedging.[140]

For the same reason, the tax law can use it to test conditions at a specific moment. Indeed, the delta test was first proposed for Section 1259, which taxes a

---

[138] The difference contract in this case is, as discussed above, simply a put option with strike $70. The difference percentages are the ratios of the put prices in the two volatility scenarios to the $100 price of the Underlying.

[139] The difference contract in this case is a knock-in put option with strike $70, together with a knock-out position in the asset. The asset position is the difference if the barrier level of $98 is never hit. The price of the knock-in put in this case is the same as an unconditional put, because a simple put with strike $70 does not provide a payoff if the barrier of $98 is never crossed.

[140] Delta helps clarify the trader's short-term exposure and figure out what position will be offsetting at that moment. Over time, delta keeps changing, so the trader updates the hedge in a process known as "dynamic hedging." Such an update is even self-financing, under idealized modeling assumptions. Notably, though, although dynamic hedging with positions in the Underlying is possible for barrier options, it is more challenging than with simple options. For barrier options, it can be useful to hedge with positions in simple options in the Underlying instead of just Underlying positions directly. For more information regarding hedging barrier options, see, e.g., Peter Carr & Andrew Chou, *Breaking Barriers*, 10 RISK 139 (1997).

hedge that is too much like a sale.[141] Since sales happen at an instant in time, delta is a good fit. Arguably, the same is true when hedging (even just for a moment) has other tax consequences, such as tolling holding periods, deferring losses,[142] and denying the dividends-received deduction.[143] In these contexts, delta answers the relevant question: how much does one position offset another at a particular moment in time?

But the relevant time horizon is different for constructive ownership under Section 1260: the question is how closely one position tracks another–not just for an instant–but for at least a year, so gains can be taxed at long-term gains rates.[144] Delta is less effective in testing this sort of enduring correlation.

More generally, when the tax law needs to assess similarity over time, we are better off focusing on factors that are likely and enduring, instead of ones that are improbable and fleeting. Fortunately, our difference value methodology is designed to focus–more than delta can–on these lasting correlations over time. As a result, our proposal can't be manipulated as easily by contingencies.[145]

4.   *The McKelvey Test*

Finally, instead of focusing on correlation (like delta) or value (like our proposal), another alternative is to focus on probabilities. How likely is a derivative to deliver the same economic return as the underlying?

The Second Circuit used this sort of probability-based test in another context: assessing whether a hedge was perfect enough to be taxed as a sale under Section 1259.[146] In theory, the hedge in *McKelvey* was not perfect. The taxpayer remained exposed to some changes in the stock price. But in practice, the price would have to increase significantly before the taxpayer actually would be affected.[147] Although this was theoretically possible, it was quite unlikely.

---

[141] *See* I.R.C. § 1259; NYSBA, Rep. #868, *supra* note 12, at 18-22   .

[142] *See* I.R.C. § 1092 (straddle rules).

[143] *See* Treas. Reg. § 1246–5.

[144] It is possible to create tests that are delta-based and attempt to incorporate the fact that delta may change over time. This is the approach taken in the regulations for complex contracts under § 871(m), for example. *See* Treas. Reg. § 1.871-15. Still, there is a conceptual mismatch for the reasons we have explained.

[145] To be sure, the value of delta is not totally irrelevant to the long-run perspective. Delta may be viewed as representing an average of future expected differences in outcomes, essentially a measure of correlation between the outcomes of the contract and the underlying. But this measure can mask important issues. For example, it nets future deviations in one direction with offsetting ones in the other direction. A delta value close to 1 may show tracking that is close *on average*, but not uniform across potential outcomes. In contrast, our difference value methodology does a better job of identifying this sort of uniform similarity.

[146] I.R.C. § 1259(d)(1). The case dealt with a contract that originally did not give rise to a constructive sale but was later modified. The court treated this modification as if it gave rise to a new contract. We do not focus on this issue here.

[147] To be precise, the taxpayer entered into a "variable prepaid forward contract." Although the taxpayer committed to sell *some* shares, the number would vary with the stock price. The problem for the taxpayer was that this quantity would vary only if the stock price increased significantly. Estate of McKelvey v. Comm'r, 906 F.3d 26 (2d Cir. 2018).

To resolve the case, the court tried to estimate the probability that this would occur. Concluding that the probability was only 15%, the Court treated the hedge as a constructive sale.[148]

In principle, policymakers could use a similar approach for constructive ownership, but we do not recommend it. The problem is that this approach focuses just on the likelihood that something will occur, but without also considering how important this scenario is. Put another way, the *McKelvey* test considers the probability of a scenario, but not its magnitude.

For example, suppose that 85% of the time there is exact equality between the underlying and a derivative. However, in the remaining 15% of the time, the derivative actually behaves like negative 10 times the underlying. Arguably, this contract should not be considered equivalent to the underlying. Nevertheless, the probability-only approach would ignore this large magnitude, but low probability, possibility.[149]

The good news is that, unlike the *McKelvey* test, both the difference-value test and the delta test incorporate magnitudes as well as probabilities. These tests do not share this defect of the *McKelvey* test.[150]

C. Issues In Making Comparisons

Section B showed that our difference-value approach is more precise and nuanced than the spread, delta, and *McKelvey* tests. Even so, our proposal still faces three of the same challenges: first, defining the relevant transaction; second, accounting for contingencies; and third, considering the taxpayer's control over the underlying assets. This Section briefly considers these issues in turn.

1. *Scope*

In our view, the main "Achilles heel" of our proposal is that it does not answer the question, emphasized above, of what should be compared to what?[151] Like the spread, delta, and *McKelvey* tests, our approach has to define the relevant transactions in order to compare them. To manipulate any of these tests, taxpayers can tweak the transactions, as discussed above, by adding additional cash flows (so bifurcation is needed) or splitting them into components (so aggregation is needed).

Even so, there is one way in which our test is easier to manipulate than the delta test: the addition (or subtraction) of a fixed amount from either the derivative or Underlying. For example, instead of paying the value of Underlying at maturity, a derivative could pay this value minus $70.[152]

---

[148] The court also dealt with a companion contract for which the probability was 13%. For our purposes, it is sufficient to consider the treatment of only one of these contracts. *Id*. at 32-33.

[149] Even if policymakers want to focus on probabilities, the court did not clarify the type of probability that should be used. Should it be risk-neutral probabilities? Or real-world probabilities? There are arguments for either type, but the court failed to resolve this issue clearly. *Id*.

[150] In the court's defense, this concern arguably did not arise in *McKelvey*. The case dealt with a specialized contract, whose payoff was largely fixed. As a result, the outcome in the low probability situation was not radically different. *Id*.

[151] *See supra* Part V.C.

[152] This is one way to describe the in-the-money option in our recurring example, which requires a payment of $70 to exercise the option. Thus, for example, if the final price of the Underlying is $110, the option pays $40 at expiration, rather than $110. This results in a substantial gap between

Arguably, adding or subtracting fixed amounts shouldn't matter to investors. They can get essentially the same return (the performance of Underlying) by putting up either more or less cash. Put another way, this is a difference in leverage, not return.[153] As a result, this difference arguably should not be relevant under Section 1260, which assesses economic similarity without accounting for leverage.[154]

So ideally, the addition or subtraction of a constant amount should be just as irrelevant under the tests we have analyzed. Is this the case? The answer is "yes" for the delta test. The correlation between Underlying and a derivative should not change if a lump sum is added (or subtracted) from one or the other.

However, adding or subtracting a lump sum could potentially change the difference-value calculation–at least if taxpayers are deliberately trying to manipulate it. For example, recall that the difference percentage is the value of the difference contract divided by the value of Underlying. But should these transactions be defined to include the fixed payment? For example, if the derivative pays $70 less than the value of Underlying, should the difference contract also include a payment of $70? Or should $70 be removed from Underlying? The answers to these questions can affect the outcome.

As a result, policymakers need to give guidance on this issue. As noted above, they can rely on general tax principles, and can tweak those based on the normative presumption.

2. Manipulative Contingencies

Another possible way to game these tests–to make a derivative seem less like the underlying asset–is to introduce contingencies. The good news, noted above, is that some contingencies are less of a problem for our proposal than they are for the delta test.[155] These contingencies have less effect on a derivative's value (and thus on our approach) because they are either unimportant or unlikely to happen.[156] The difference-value is able to shrug off these minor contingencies–in a

---

the Underlying and the derivative. We eliminated this gap for payoffs above the strike price by requiring that the contract pay $70 in addition to the simple payoff of the call option.

[153] By leverage, we mean financing that must be repaid regardless of the performance of the Underlying. For example, if an investor borrows $70 to buy a risky asset, and must pay the $70 without regard to what happens to the risky asset, the $70 is a quintessential example of what we mean by leverage. As it turns out, some features of the GWA contract that at first seem to be pure leverage–and have no connection to Underlying's price–actually do have a connection. For a discussion, see *infra* Appendix A. This means that this apparent advantage of delta over the difference-value method may be less than meets the eye.

[154] In defining a constructive ownership transaction, Section 1260(d)(1) does not account for leverage. For example, it specifically lists "a forward contract to acquire" the underlying asset without regard to whether the forward contract requires an up-front payment. I.R.C. § 1260(d)(1). Similarly, it also includes "a holder of a call" and a "grantor of a put," even though options usually involve deferred payments. I.R.C. § 1260(d)(1)(C).

[155] *See infra* Part VI.B.3.c.

[156] We speak in general terms like "unimportant" and "unlikely" to convey the gist of the idea without technical details. Note that in order for our difference method not to change much as a result of a contingency, both the probability of the contingency must be low, and the alternate outcome must not be too extreme. Contrast this with the *McKelvey* test that focused only on probability and not on the size of the low-probability event.

way the delta cannot–because we focus on value over time, not on correlations at a point in time, as noted above.[157]

In contrast, contingencies pose more of a challenge to our difference-value test when they are more likely to affect the derivative's value–that is, when either their probability or magnitude is more significant. Let's turn to some examples to illustrate this challenge and to show how the difference-value test should deal with it.

a. Coin Flips

For example, assume that Taxpayer enters into a quintessential constructive ownership transaction–a forward contract to acquire an interest in Underlying (a hedge fund)[158]–but with a twist: on the day after Taxpayer enters into this contract, there is a coin flip to determine whether Taxpayer is the buyer or the seller on this contract. Because of this contingency, this derivative is 50% likely to be a perfect substitute for Underlying (when Taxpayer is the buyer), and 50% likely to be the exact opposite (when Taxpayer is the seller).[159]

This sort of contingency can confuse both the delta and difference-value tests. Before the contingency is resolved, the delta of this derivative is essentially zero, because both the "buyer" and "seller" outcomes are equally likely (and thus cancel each other out). For the same reason, its initial value also is essentially zero, so the difference-value test would generally treat it as very different from Underlying.[160]

But this is the wrong result. This derivative can actually be a perfect substitute for Underlying, as long as taxpayers are patient: if the coin toss makes them the seller, they can cancel this derivative and try again.[161] Eventually, the coin toss will go the right way, giving them a derivative that is a perfect substitute.

b. More Realistic Contingencies

Admittedly, the coin flip seems tax-motivated, since there is no commercial rationale for it. Yet other contingencies are easier to justify, including the level of unemployment or inflation, the closing of a specific transaction, the price of

---

[157] As our knock-in call example in Part VI.B.3 shows, a low probability contingency with a not-very-extreme impact on the outcome can move delta from close to *positive* one to well below *negative* one.

[158] Assume the term is three years, the amount of hedge fund interests to be delivered is fixed, and the purchase price is equal to the current value of the interest in the hedge fund, grown at the rate of return on Treasury bills over the course of the three-year period.

[159] This coin-flip example is based on an example in Brennan & McDonald, *supra* note 114.

[160] Specifically, this contract provides ownership of the Underlying half the time. But the other half of the time, it provides the exact opposite, i.e., a short position in the Underlying. This means the difference contract provides full risk of loss and opportunity for gain half the time, so the difference contract is zero. The other half of the time, the difference contract is actually double the value of the Underlying. This is the absolute value of the difference between the Underlying and its negative. The price of the difference contract is thus the same as the Underlying (*i.e.*, one-half of double the price of the Underlying). Thus, our difference percentage would be 100%.

[161] Note that this possibility implicates the question of scope discussed above as well. If we take the derivative to be not just a single coin-flip contract, but rather the entire series of contracts to be used, then we can get to the correct answer under either our test or the delta test. Considerations of scope and manipulability often intersect and overlap. We discuss each in turn to highlight different facets of the overall challenge in assessing similarity.

Underlying, and the like. To defend these contingencies, taxpayers can claim that they are relevant in deciding whether to invest in Underlying.

Alternatively, taxpayers can add another contingency with a tax advantage of its own: they can invest in Underlying through a variable life insurance policy, with a death benefit equal to the value of Underlying on the date of death. To buy this insurance, taxpayers can pay the current value of Underlying (which the insurance company uses to invest in Underlying), plus a generous fee. If the form of this transaction is respected, the proceeds are tax-free, so any profit on Underlying is never taxed.[162]

In each of the examples in this Subsection, the derivative and Underlying are likely to be quite similar, but a contingency makes them seem different. How should our "difference-value" methodology treat these contingencies? This is really a two-part question. First, as a matter of policy, what should the answer be? Second, what are the options for implementing this policy judgment?

c. Normative Assessment

In answering the first question, policymakers should once again be guided by the normative presumption. On the one hand, how troubling are false negatives? How problematic is it for taxpayers to change their tax treatment by including these conditions? On the other hand, how harmful are false positives? To what extent will a broad rule chill socially useful transactions that are not tax-motivated?

This analysis will not be the same for every contingency. If there is no commercial rationale for it (e.g., the coin toss), policymakers should discount it. But if a contingency has social value (e.g., life insurance), it warrants more deference.

d. Probability-Based Assumptions

If policymakers do not want contingencies to prevent constructive ownership, they can adjust the "difference-value" methodology in two ways. First, they can require taxpayers to make assumptions about the contingency. If it is virtually certain to be satisfied, the relevant valuations should assume that it already has been satisfied. Likewise, if a contingency is unlikely, the valuations should ignore it. The tax law already uses this approach, for instance, with contingent payments on debt instruments.[163]

e. Updated Analysis After Contingencies Are Resolved

Second, instead of predicting whether a contingency will be satisfied, another alternative is for taxpayers to wait and see. If the relevant condition ultimately is satisfied, a (partially) updated valuation can be required. To be clear, the only update would be the inclusion of the contingency. Otherwise, the analysis would still use the conditions in effect when Taxpayer entered into the derivative.

To illustrate this "updated ex ante" approach, assume that on January 1 of Year 1, when an interest in Underlying is worth $100, Taxpayer enters into a three-year call option to buy Underlying for $200, which has a contingency: after two

---

[162] *See* I.R.C. § 101.
[163] *See* Treas. Reg. § 1.1275-2(h); Treas. Reg. § 1.1275-4(a)(5).

years, a coin toss will determine whether this purchase price on the option (the "exercise price") is reduced from $200 to $50. Imagine that this coin toss does, indeed, reduce the exercise price to $50, and over this two-year period Underlying's price has declined to $45.

Should Section 1260 apply to this option? If the contingency is ignored in Year 1, the answer probably is "no." A three-year option to buy Underlying for twice its current value (i.e., $200 when it is trading at $100) is not a close substitute for Underlying.

What about after the coin toss two years later? Taxpayers can be required to rerun the analysis, since the exercise price on the option has now fallen from $200 to $50. But the answer depends on what other values are used. Specifically, should Underlying be valued at $100 (the value when Taxpayer originally got the option) or at $45 (the value two years later when the exercise price was reduced). We recommend using the initial value ($100), which (in this case) makes this option more likely to trigger Section 1260. In other words, the only fact we would update is the contingency itself (i.e., the new $50 purchase price on the option), but not anything else (e.g., the value of Underlying).

The advantage of this updating approach is that it doesn't require assumptions that might prove incorrect. But the disadvantage is that it doesn't provide certainty about the relevant treatment. Taxpayers and the government have to wait to find out whether this derivative triggers Section 1260. (Presumably, interest would be imposed to compensate the government if the derivative should have been taxed less favorably, but Taxpayer's earlier returns did not reflect this less favorable treatment.)

3. Investor Control

So far, this Subsection has shown that our difference-value approach is not immune to challenges that arise in any effort to compare two transactions: the need to define their scope, and to decide how to treat contingencies. In the same spirit, there is another issue that our difference-value approach does address: whether a taxpayer exerts enough control over an asset to qualify as its owner for tax purposes.

Actually, control is not relevant under Section 1260, which focuses instead on economic return. This statute applies when a derivative is a sufficiently close substitute for the underlying asset, regardless of whether the taxpayer is managing or controlling this asset. We agree with Section 1260's focus on return and adopt it in this Article, since return usually is what matters most to investors.

Yet we recognize that the government sometimes focuses on control–not under Section 1260–but under the law of tax ownership. For example, when a taxpayer enters into a derivative contract with a bank, and the bank hedges by purchasing particular assets, the government sometimes invokes control to argue that "the real owner" of these assets for tax purposes is the taxpayer, not the bank.

Knowing this, well advised taxpayers are careful not to link the derivative too closely with the bank's hedge. For example, the payout on the derivative should not be defined as the price the bank gets in liquidating its hedge. Likewise, if the derivative is based on a portfolio managed by the bank, the taxpayer should be

careful not to get too involved in managing this portfolio. Under the caselaw, this sort of control can affect the result.[164]

We flag these issues mainly to emphasize that our "difference-value" approach does not address them. Questions of ownership are pervasive in the tax law, and they are fairly well understood.[165] We are not seeking to contribute to that understanding here.

We should note, though, that it is not obvious to us why, as a matter of policy, a taxpayer's active participation makes the transaction more problematic. Admittedly, the optics may be less favorable when taxpayers claim not to be the owner, but still play an active management role. But so what? If the issue is really whether particular types of economic returns should be taxed one way or the other, why does it matter who is making the decisions?

Moreover, if control actually does matter, the government faces a difficult challenge in policing it: many trading strategies can run on "autopilot" once they have been developed. Indeed, a trading strategy often can be distilled into an algorithm that is shared with the bank–a step that the government could not easily monitor.

In any event, we do not seek to analyze these issues here. The significance of investor control–and its application to derivatives based on pass-through entities or managed accounts–are beyond this Article's scope. Instead, our focus is on economic similarity and how to measure it.

VII.   BASKET OPTIONS AND THE GWA LITIGATION

This Part applies our difference-value approach to a type of hedge fund derivative, known as a basket option, which is the subject of a multi-billion dollar litigation, *GWA, LLC v. Commissioner*. After describing the planning strategy in this case, this Part explains why Section 1260 does not reach it (at least until Treasury promulgates regulations), and then turns to a key issue, which looms large both in the case and in any regulations the Treasury ultimately adopts: how does the economic return on the derivative compare with the return on the underlying? We flag problems in analyzing this issue with the spread, delta, and *McKelvey* tests, and show that our difference-value test avoids these problems, offering a more nuanced and robust analysis of the issue.

A. Basket Options: A Potential End Run Around Section 1260

Like the first generation of hedge fund derivatives described above–forward contracts to buy a fund interest[166]–basket options are supposed to avoid the tax cost

---

[164] *See* Webber v. Comm'r, 144 T.C. 324 (2015); Steven Horowitz et al., *The Tangled Web: Substance vs. Form, Webber, and the Revenge of the Investor Control Doctrine*, 34 J. TAX'N INV. 19 (2017)

[165] For an insightful discussion, *see generally* Alex Raskolnikov, *Contextual Analysis of Tax Ownership*, 85 B.U. L. REV. 431 (2005).

[166] *See supra* Part II.A.

of frequent trading.[167] After describing basket options, this Section surveys issues they raise under Section 1260, as well as under tax rules governing ownership.

1. Differences From Earlier Hedge Fund Derivatives

Compared with earlier hedge fund derivatives, basket options are different in two ways. First, this trading is outsourced to derivatives dealers. Using the hedge fund's trading strategy, dealers manage a portfolio (or "basket") of securities and grant an option to buy this basket to the hedge fund (or a potential investor in the fund).[168] Second, this option (supposedly) offers some protection from risk of loss in the basket, and thus (arguably) does not track it as perfectly as a forward contract.[169]

In 2014, a Senate committee blew the whistle on basket options, claiming that they were avoiding billions of dollars in tax.[170] The Treasury followed up with a pair of notices, deeming basket options a "tax avoidance transaction" and requiring special disclosure about them.[171] The government also challenged basket options on audit. One hedge fund, Renaissance Technologies, LLC, reportedly settled by paying about $7 billion, which is one of the largest tax settlements in history.[172]

2. GWA Contract

Meanwhile, another hedge fund, George Weiss Associates ("GWA"), opted to litigate the issue. The trial was held in September 2022 and, as of this writing, the judge has not yet issued a decision.

---

[167] U.S. S. PERMANENT SUBCOMM. ON INVESTIGATION, *Abuse of Structured Financial Products: Misusing Basket Options to Avoid Taxes and Leverage Limits* 1 (2014) https://www.hsgac.senate.gov/wp-content/uploads/imo/media/doc/REPORT-Abuse%20of%20Structured%20Financial%20Products%20(Basket%20Options)%20(7-22-14,%20updated%209-30-14).pdf [https://perma.cc/3WK4-P6BU] ("the resulting short-term profits were frequently cast as long-term capital gains").

[168] *See* I.R.S. Notice 2015-73, 2015-46 I.R.B. 660 (noting that taxpayers "either determine the assets that comprise the reference basket or design or select a trading algorithm that determines the assets" and have the right to make (nonbinding) suggestions that the dealer generally accepts).

[169] *See id.* ("In some cases, taxpayers are also mischaracterizing a transaction as an option to avoid application of § 1260."). As discussed further below, the option terminates when the portfolio declines below a specified level, which is above the price the holder is entitled to pay for the Underlying (the "strike price"). This feature generally causes the option to function more like a forward contract. Yet this instrument does differ from a forward contract when the price declines very rapidly. In this scenario, some risk of loss can be shifted to the counterparty on the option, as noted below. *See infra* Part VII.C.4.

[170] *See* U.S. S. PERMANENT SUBCOMM. ON INVESTIGATION, *supra* note 167 (noting that strategy had been "used by at least 13 hedge funds to conduct over $100 billion in securities trades").

[171] *See* I.R.S. Notice 2015-73, 2015-46 I.R.B. 660 (deeming basket options in effect on or after January 1, 2011 to be "listed" transactions beginning on October 21, 2015); I.R.S. Notice 2015-74, 2015-46 I.R.B. 663 (deeming basket options entered into on or after November 2, 2006 and still in effect on Jan. 1, 2011 to be "transactions of interest").
These two notices revoked Notice 2015-47 and Notice 2015-48. The later notices were similar to the earlier ones, with the difference that the later ones provided more detailed descriptions of the transactions in question to avoid overbreadth.

[172] Manojna Maddipatla et al., *Renaissance Executives Agree to Pay Around $7 bln to Settle Tax Dispute With IRS*, REUTERS (Sept. 2, 2021), https://www.reuters.com/business/finance/renaissance-executives-pay-about-7-bln-settle-tax-probe-wsj-2021-09-02/ [https://perma.cc/8SEL-RPKW].

In this case, Deutsche Bank ("the Bank") managed a portfolio, which it fine-tuned with constant trading ("the Underlying Portfolio" or "the Underlying").[173] GWA entered into a derivative contract with the Bank ("the GWA contract" or "the contract"), which was based on the value of this portfolio. This subsection gives a somewhat high-level overview of the contract.[174]

GWA paid a cash premium of $10 to the Bank for this contract, which was documented as a call option with a stated term of 12 years. It entitled GWA to pay $90 for a portfolio that was then worth $100. In other words, this option was somewhat in-the-money. If this were the entire story, the contract would have offered all the opportunity for gain in the managed portfolio and provided protection against losses below $90.

Yet unlike a conventional call option, the contract had two additional features that arguably rendered it a "forward contract in disguise." First, the contract would terminate (or "knock out") if the portfolio's value declined below a preset expiration price ("EP") of $97–that is, $3 (or 3%) below the portfolio's initial value. Second, GWA had the right to keep the contract from terminating by putting up more money and thereby lowering the knock-out threshold.[175] For example, if GWA invested another $3, the EP declined from $97 to $94. If the portfolio kept declining and reached $94, GWA could put in another $3 to reduce the EP to $91, and so on.

These cash infusions were tracked in a premium account, which started at $10 (reflecting GWA's up-front premium) and increased by any amounts GWA invested to "buy down" the EP.[176] If GWA allowed the contract to terminate, the Bank would keep a portion of the premium account to cover declines in the value of the portfolio, and GWA would get the rest. For example, if the portfolio declined

---

[173] The account is managed based on a specified investment strategy. As a result, the case presents investor control issues, such as those discussed in Part VI.C.3. But the analysis here focuses instead on economic return.

[174] Two caveats are in order here. First, in order to simplify our calculations and to be consistent with our prior examples, we assume that the investment strategy for the Underlying of the GWA contract follows a lognormal process, even though this is not necessarily the best model for hedge fund returns (which are often understood to be relatively stable, except for occasional large downward deviations).

Second, to keep the analysis tractable and the exposition clear, we oversimplify the transaction at some points. In addition, we do not have access to the full range of relevant evidence, including the trial testimony. Rather, our source is the contract itself, which is described in Joint Trial Exhibit 23-J. No. 6981-19 (U.S.T.C. filed Nov. 20, 2020). As a result, our analysis is intended to be suggestive, not definitive.

[175] In the actual contract, 97 is the initial "expiration notice price level" rather than the "expiration price." When this level is reached, there is generally an opportunity for the seller to provide notice to the buyer and for the buyer to buy down the knock-out level, as described in the text. Such a buy-down must occur quickly upon notice, generally within a few hours. The initial "expiration price" in the contract is 94. If this level is reached before a buy-down has occurred, or before notice can be given, the contract ends. For simplicity of exposition, we treat 97 as the knock-out level, subject to the opportunity for the buyer to pay an additional buy-down amount, and we refer to this as the "expiration price."

[176] The premium account also was reduced by an amortized premium amount of 0.1 for every year during the term of the contract, reflecting a permanent payment from the investor to the bank, and a total payment of 1.2 over the course of the 12-year term. For the sake of simplicity, we do not include this feature.

to $97 and GWA chose to terminate the option, Bank would take $3 from the premium account (to cover the portfolio's decline from $100 to $97), leaving GWA with the remaining $7. In effect, GWA would absorb the $3 loss.

Because of these contingencies, GWA's contract arguably was less like a call option than a margin account or a forward contract. The contract did not insulate GWA from risk of loss as the portfolio's value declined, as long as GWA exercised its buy-down rights. If GWA actually desired full exposure to the underlying, it could get it by simply continuing to cover these losses. This risk of loss usually arises–not in a call option–but in a forward contract or an investment in the underlying.

A key question, then, is whether GWA regularly used the buy-down feature to keep the contract from terminating. On the one hand, if the answer was "yes," the contract was a close substitute for the Underlying, as noted above. On the other hand, if the answer was "no," the contract was less like the Underlying: GWA had opportunity for gain–as long as the contract didn't "knock out"–but GWA did not have risk of loss below $90 (since the contract would terminate at $97).[177]

3. Planning Around Section 1260

Even though GWA's contract had a lot in common with the hedge fund derivatives covered under Section 1260, there was an important difference: The contract was based on the value–not of a hedge fund–but of a managed account. As a result, it fell outside the language of Section 1260, which applies only to derivatives based on the value of "any equity interest in any pass-thru entity."[178] In GWA's transaction, there was no "pass-thru entity."

This is not to say that derivatives on managed accounts are "in the clear." Section 1260 gave Treasury regulatory authority to cover them, although no regulations have been issued yet.[179] Even so, Treasury has used notices to deter these transactions, as noted above.[180]

When Treasury ultimately does write regulations, a basket option presumably will be covered only if it conveys substantially all of the risk of loss and opportunity for gain in the basket. In other words, the answer will turn on a comparison, and this Part shows how it should be done.

4. Economic Similarity and General Principles

The same comparison also is relevant in the GWA litigation. Even if the government cannot invoke Section 1260, it can rely on general principles to claim that GWA–not the dealer–was the real owner of the portfolio. If the court agrees, GWA would be taxed on (short-term) gains from the Underlying Portfolio, instead of on (long-term) gain from the basket option.

---

[177] As a middle ground, if buy-downs will only be used to a point, then a more complex analysis is needed to address how the contingencies will be resolved in various situations.

[178] I.R.C. § 1260(c)(1)(A).

[179] Specifically, this authority covers derivatives based on common stock and debt securities. *See* I.R.C. § 1260(c)(1)(B) & (d)(1)(D). Presumably, this authority extends to managed accounts comprised of these investments. An interesting question is whether it extends to accounts with other assets, such as commodities and cryptocurrency. If not, could taxpayers avoid the rule by including these other assets in the portfolio? How much must they include?

[180] I.R.S. Notice 2015-73, 2015-46 I.R.B. 660; I.R.S. Notice 2015-74, 2015-46 I.R.B. 663.

In this analysis, the key question is, once again, how similar the GWA contract was to the Underlying Portfolio. If these two investments offered essentially the same economic return,[181] a court is likely to disregard the contract and treat the Underlying Portfolio as owned by GWA, not the Bank. But if the contract is meaningfully different–and thus has economic substance and a business purpose–a court is likely to respect the transaction's form, taxing GWA on the contract instead of on the Underlying Portfolio.[182]

In short, this case turns to a significant extent on how similar these two positions were. This brings us back to the main issue in this Part: how should the tax law assess the similarity of these two positions? Should it use the spread, delta, *McKelvey*, or difference-value test? Let's consider each alternative in turn.

B. Inadequacy of Other Tests

This Section applies the spread, delta, and *McKelvey* tests to the GWA contract. Unfortunately, none of them are well suited to analyzing this planning strategy.

1. Spread Test

Although the spread test has the virtue of simplicity, it is a poor fit for the GWA contract. In form, this contract is an option to buy an underlying for less than its current price (i.e., an "in-the-money" call). In this way, the contract resembles our recurring example of an option to pay $70 for an asset worth $100. Like that option, the GWA contract fails the spread test by tracking the underlying too closely at the Underlying's current price: in other words, the spread of omitted exposure ($0 to $90) doesn't include $100.[183]

Perhaps failing the spread test, and hence treating the GWA contract like actual ownership, is the correct result. Or perhaps it is not. It is hard to know because the spread test ignores important factors such as the volatility of Underlying and the term of the option, as discussed above.[184]

In principle, a modified spread test could measure the discount on the option (i.e., how in-the-money it is), as noted above.[185] The GWA contract purports to offer only a 10% discount (i.e., charging $90 for something worth $100), which is less than the 30% discount in our recurring example (i.e., charging only $70). This implies that the GWA option is less like the underlying.

But once again, the spread test ignores important information about the contract: the "knock out" and "buy-down" features. These contingencies arguably

---

[181] In arguing that GWA is the true owner, the government can try to offer evidence that GWA, not Bank, is deciding how to invest the Underlying Portfolio. As noted above, this "investor control" analysis is not our focus here. *See supra* Part VI.C.3.

[182] Along with showing differences in risk of loss an opportunity for gain, GWA can also claim that the contract was more leveraged than the Underlying Portfolio. Arguably, a business purpose for using it was to circumvent regulatory limits on leverage. *Cf.* U.S. S. PERMANENT SUBCOMM. ON INVESTIGATION, *supra* note 167 (criticizing use of basket options to evade limits on leverage).

[183] *See supra* Part VI.B.1.

[184] *See supra* Part VI.B.1. As proposed by the NYSBA, the spread test would have rejected the GWA contract for another reason as well: it's unusually long term of twelve years. NYSBA, REP. #901, *supra* note 89, at 28.

[185] *See id.*

make the GWA contract a closer substitute for the Underlying, as explained above.[186] Admittedly, the test can be modified to consider more variables, but this may well defeat its purpose, which is to keep things simple.

2. Delta Test

While a more sophisticated test is needed, the delta test may not be up to the job either. To be fair, the test works well with a simple in-the-money call, which doesn't have contingencies like knock-outs and buy-downs. Unlike the spread test, the delta test can account for a "plain vanilla" option's volatility and term, as explained above.[187]

To see this, imagine that the GWA contract was just a simple call option, with no fancy contingencies, entitling GWA to pay $90 for an Underlying that initially was worth $100: in other words, a call that was 10% in-the-money. This option would give essentially the same economic return as the Underlying as long as Underlying's price did not decline by more than 10%.

The delta test is effective in gauging the significance of this difference. As Table 6 shows, delta is a bit higher with a 60% volatility than a 20% volatility.[188]

Volatility	20%	60%
Delta	0.914	0.916

Table 6: Delta of GWA Contract Without Contingencies

In addition, an Underlying currently worth $100 is less likely to end up below $90 as the years go by. This scenario becomes less likely with the passage of time (if only because the risk-free rate is positive, so assets generally should appreciate by at least that rate on average). As a result, options that are 10% in the money become a closer substitute for the Underlying as their terms get longer. Perhaps this is the reason why the GWA contract has the unusually long term of twelve years. Once again, the delta test is sensitive to this difference. As Table 7 shows, delta is higher for a 12-year option than for a one-year option.

Option Term	Vol=20%	Vol=60%
12 Years	0.914	0.916
1 Year	0.810	0.712

Table 7: Delta of GWA Contract without Contingencies:
12-Year Versus 1-Year Term

---

[186] *See supra* Part VII.A.2.

[187] *See supra* Part VI.B.2.

[188] As with our earlier examples, we use the Black-Scholes pricing model with a risk-free rate of 5% and a term of 12 years.

But although the delta test is reliable for these simple options, it misfires badly once contingencies are introduced, as explained above.[189] This potential for manipulation is on full display in the GWA contract.

To highlight this problem, let's add the knock-in feature to our in-the-money option, so it terminates if Underlying's price falls below $97. Let's also assume that there's no buy-down. In other words, GWA won't make an additional investment to save the option. (This assumption is critical, as we discuss below.) On these assumptions, the delta test produces a crazy result: a delta greater than one. As the price of Underlying declines by a dollar from $100 to $99, the contract's value declines by more than a dollar:

Volatility	20%	60%
Delta of Knock-out Call	2.42	1.24

Table 8: Delta of GWA Contract With Knock-Out (assuming no Buy-Down)[190]

This result is truly bizarre. After all, the contract will never–indeed, it can never–pay more than the value of the underlying. Thus, the high delta value is hardly a prediction of a potential payoff.

Instead, it is just the product of the weird interaction between small price changes, on the one hand, and the knock-out condition, on the other. Specifically, a decline from $100 down to $97 has only modest implications for the Underlying Portfolio, but major–potentially fatal–implications for the knock-out call. So if the Underlying's value drops from $100 to $99, the derivative's value will decline by more than a dollar–because, again, it's coming dangerously close to being terminated. As a result, since a decline of a dollar in the underlying leads to a decline of more than a dollar in derivative, delta is greater than 1.

There are other wacky effects as well, which derive from other aspects of the GWA contract. For example, the contract includes an implicit loan from the Bank to GWA (in sparing GWA from paying the full price of Underlying up front). While a simple loan ordinarily would not affect delta–since the interest and principal payments do not change with Underlying's price–this loan is different: the timing of GWA's payments do, in fact, depend in part on the Underlying's price.[191] As a result, these payments have a delta of their own. As the Appendix

---

[189] *See supra* Part VI.B.3.

[190] This table reflects the delta of the knock-out call, but the contract actually has a second component: the $7 payment the holder receives when the option is terminated (i.e., from their premium account). As Appendix A shows, accounting for this component reduces delta, but not by very much. Appendix A also analyzes additional interest payments due under the contract. As Appendix A shows, the timing of these payments depends on the Underlying price. As a result, these payments have a delta of their own, which helps offset the overly-large delta caused by the contingency.

[191] For instance, declines in the price cause GWA to get a portion of their premium account back (if the contract terminates) or to pay additional amounts to buy-down the expiration price.

shows, the delta of these payments can offset the unexpectedly large delta caused by the contingency.[192]

But these "adventures in delta" arguably are not a persuasive analysis of the GWA contract, at least once the buy-down feature is taken into account. This right to save the option allows GWA to take on risk of loss below $90, as emphasized above.[193] So if the delta test worked as it should, it would find a closer correlation between the GWA contract and the Underlying Portfolio. Delta should be much closer to 1.0.[194]

To sum up, although the delta test is effective for simple instruments, it can make a real hash of more complicated ones, especially when contingencies are involved. For example, it is confounded by the knock-out call option and yields delta values much greater than 1.0.

There are ways for the delta test to get to the right answer–for instance, by simply assuming a buy-down and thus ignoring the knock-out[195]–but this requires a more nuanced analysis, which taxpayers may not be motivated to provide. After all, if their goal is to justify their tax planning, they may well want a misleading result, and the delta test gives them discretion to engineer it (or, at least, the illusion of it).

3. McKelvey Test

So far, we have shown that the spread and delta tests are not reliable ways to analyze the GWA contract. Unfortunately, the *McKelvey* test is no better. While this test focuses on probabilities, as explained above,[196] the right way to apply it to the GWA contract is not obvious. After all, the context is different. *McKelvey* was about a specific issue–the number of shares to be delivered–that does not really arise in GWA.[197]

By analogy, perhaps the right question is the probability that the GWA contract will diverge from the Underlying. Since this happens when the Underlying's value falls below $97, this test can ask, "how likely is Underlying's

---

[192] If the interest payments on the loan were unrelated to the performance of the Underlying, then the loan would be irrelevant and the problems with the delta would remain unchanged.

[193] *See supra* Part VII.A.2.

[194] Just how close it gets depends in part on how precisely the contract accounts for the time value of various payments, including the deposit returned to GWA and the payment of the exercise price to the Bank. For a discussion, see *infra* Appendix A. Indeed, once these (and other) factors are taken into account, if the bank receives ongoing compensation at the risk-free rate for the outstanding balance on its implicit loan to the taxpayer, then there is in fact a delta of 1.0. This type of compensation is arguably present in the GWA contract, although it is not part of a knock-out call option in isolation. For further discussion, see *infra* Appendix A.

[195] Another way to do this is to account in a more precise way for other time-value-related components of the contract. *See infra* Appendix A.

[196] *See supra* Part VI.B.4.

[197] If we get creative, we can try to characterize the issue in *GWA* as the number of units, but this analysis is forced. The argument would be that the knock-out produces a fixed number of units: GWA gets $7 of its deposit back, and (in a sense) this represents 7/97 (or .072) of a unit of the Underlying Portfolio. Since the probability of a knock-out is high, as noted below, the test can find a high probability of producing a fixed amount (i.e., $7 worth of the Underlying). But the real issue in *GWA* isn't how likely the GWA contract is to yield a *fixed amount*, but how likely it is to *diverge from the Underlying*.

value to decline below $97?" As Table 9 shows, the risk-neutral probability of that outcome is extremely high:

Volatility	20%	60%
Probability	0.9389	0.9961

Table 9: Risk-Neutral Probability of Knock-Out

Applied in this way, the *McKelvey* test finds that the GWA contract is quite likely to diverge from the Underlying. But this isn't necessarily right. Like the delta test, this analysis fails to account for buy-downs. If GWA responds by investing more money each time the threshold is triggered, there actually isn't much difference in economic return. In short, *McKelvey* also doesn't provide the precision we need.

C. Difference-Value Test

In contrast, our difference-value test is better at capturing the real economics of the GWA contract. Our approach is harder to manipulate with contingencies, as noted above,[198] because it reflects the range of potential outcomes in a more thorough and practical way. This Section applies our difference-value test to the GWA contract and responds to potential concerns about this approach.

1. Defining the Benchmark: Similar to What?

In comparing the GWA contract with the Underlying Portfolio, the first step is to define the scope of the two transactions. Again, what is being compared to what? To compare "apples to apples," we make two assumptions.

First, like most derivatives, the GWA contract is more leveraged than the Underlying. GWA can bet on Underlying without paying the $90 exercise up front. Yet a difference in leverage generally should not avoid constructive ownership because the statute focuses–not on leverage–but on opportunity for gain and risk of loss, and rightly so.[199] To conform the leverage of the GWA contract and Underlying Portfolio, we assume the Underlying was purchased with $90 of borrowed money.[200]

Second, just as the derivative and underlying should have the same leverage, they also should end at the same time. As a matter of form, this is not the case. On the one hand, there is no time limit on owning the Underlying Portfolio. On the other hand, the GWA option has a fixed term (12 years) and can terminate earlier if the price falls below $97 (unless the investor pays to avoid this knock out). So to compare "apples to apples," we assume that our benchmark–an investment in the Underlying–would be sold whenever the GWA contract terminates (whether at maturity or after a "knock out").

---

[198] *See supra* Part VI.C.2.

[199] *See supra* Part VI.C.1; *supra* note 154.

[200] We assume that any loans in the GWA contract and the Underlying bear interest at a market rate, such as the risk-free rate, and are unrelated to the performance of the Underlying. If these interest payments match, our analysis can ignore them. *See infra* Appendix A.

In our view, this assumption is justified for three reasons. First, the investment in Underlying is hypothetical–after all, GWA is investing in the contract, not the Underlying–so an assumption of some sort is needed. Second, if we assume this hypothetical investment would have been either shorter or longer, we would have to make other (potentially messy) assumptions, for instance, about how GWA invests its proceeds when the contract terminates (e.g., which investment, for how long, etc.). Third, the assumption that the term is the same is especially plausible because of the contract's "buy down" feature: GWA has discretion to continue the contract, just as it would have had discretion to continue a direct investment in Underlying. Presumably, GWA would have made the same choice either way about how long to invest.[201]

2. Comparing Values: GWA and the (Properly Delineated) Underlying

Now that we have defined the relevant transactions, what is the difference in their values? Like with the delta test, let's start by imagining that the GWA contract was just a simple call option with no contingencies. Again, by allowing GWA to buy an Underlying currently worth $100 for only $90, this contract gives essentially the same economic return as the Underlying, as long as Underlying's price does not decline by more than 10%.

How meaningful is this difference? Under our difference-value approach, the first step is to define the difference contract, as explained above.[202] To fill in the missing exposure, GWA would have to take on risk of loss below $90: in other words, GWA would have to sell a put option with an exercise price of $90. Like the GWA contract, this put would have a twelve-year term.

The second step in our methodology is to value this put option.[203] Like the delta test, our methodology is sensitive to the Underlying's volatility. As Table 10 shows, the put is much more valuable–so the contract is more different from the Underlying– when the Underlying is volatile:

Volatility	20%	60%
Value of Difference Contract	3.77	29.04

Table 10: Value of Difference Contract Without Contingencies

Like the delta test, the difference-value test also is sensitive to the term of the contract. As Table 11 shows, a longer duration makes the difference contract more valuable.

---

[201] Admittedly, the contract has a twelve-year term, while a direct investment would not. But the parties presumably could extend the contract and, even if they didn't, twelve years is a long time.
[202] *See supra* Part VI.A.3.
[203] *See supra* Part VI.A.4.

Option Term	Vol=20%	Vol=60%
1 Year	2.31	15.40
12 Years	3.77	29.04

**Table 11: Value of Difference Contract without Contingencies:
12-Year Versus 1-Year Term**

The third step is to calculate the difference percentage.[204] For a twelve-year option, it is large (29.04%) if the Underlying's volatility is 60%, but modest (3.77%) if the volatility is 20%. This makes intuitive sense. Retaining exposure to declines beyond 10% is more meaningful if sharp declines in the Underlying's price are more likely. In other words, it matters whether the Underlying Portfolio focuses on tech startups or regulated industries. So the contract is more similar–and thus should be more likely to trigger constructive ownership–with a low volatility Underlying.

But this calculation is still incomplete because it does not account for the knock-out contingency. After all, the GWA contract is not really a simple call option that is 10% in the money (i.e., with a $90 exercise price when Underlying is worth $100). Rather, it terminates when the price of Underlying declines below $97. This means GWA is protected from risk of loss below $90–not necessarily for the full twelve years–but only as long as Underlying's price does not fall below $97.

Let's pause and consider the economic significance of this protection–or, really, its insignificance. If you were GWA's risk-management officer, you might initially be pleased to know that you were protected from risk of loss below $90. That sounds good, right? But once you hear that this protection (and, indeed, the whole deal) goes away when the price hits $97, you would likely ask, "Wait, this protection can never actually help us, right? After all, in order for us to need protection below $90, the value of Underlying has to fall below $97, right? So doesn't this protection terminate before we ever actually need it?" The answer is "yes" to all three questions. In other words, this protection is virtually always meaningless.[205]

So, returning to our difference value method, the first step is to identify a contract that would nullify this protection–and, thus, put GWA in the same position as if it was investing directly in Underlying.[206] Like in the noncontingent case, GWA needs to sell a put with an exercise price of $90. But here, there is a contingency: this put "knocks-out" when the price falls below $97.

The second step is to value this knock-out put.[207] The answer, of course, is that this put has essentially no value. This is the answer whether the volatility is 60

---

[204] *See supra* Part VI.A.5.

[205] A remote scenario in which it might actually matter, which we call "slippage," is discussed below.

[206] *See supra* Part VI.A.3.

[207] *See supra* Part VI.A.4.

or 20 and, indeed, whether the term is twelve years or one year. This result should be clear not only from a Black Scholes analysis, but also from intuition: as noted above, protection below $90 is useless if this protection (and the investment being protected) terminates when the price falls below $97. As a result, the third step–the difference percentage–is also zero.[208]

In other words, compared with the spread, delta, and *McKelvey* tests, our difference value method is more effective in grappling with contingencies. Unlike these other tests, it gives the right answer for the contingent GWA call: the contract is likely to be a close substitute for the Underlying.

But is our approach really better? Or are we just using better assumptions? Or–even worse–are we assuming they are similar, and then letting this assumption dictate the result? The rest of this Subsection considers these questions, focusing on two assumptions: first, GWA does not "buy down" the contract; and second, the Underlying and Contract can be liquidated for the same value.

3. First Assumption: No Buy-Down

Our analysis assumes that GWA does not use its "buy-down" right (i.e., to keep the contract from terminating). Instead, we assume that the deal (and the downside protection) terminates when the Underlying's price falls to $97, as noted above.[209]

This assumption about buy downs is critical, even dispositive, in the delta test: the more buy downs we assume, the closer delta gets to one, as noted above.[210] After all, if GWA always has to buy down the contract, however low Underlying's price goes, GWA bears essentially all the risk of loss. This makes the GWA contract more like a forward contract than an option, as noted above.[211]

However, this assumption does not loom as large in the difference-value test. In our proposal, the key assumption is not whether this contract will be extended, but whether it has the same duration as Underlying. As long as these investments end at the same time–whether both are bought down or neither are bought down–there is an exact match under our methodology. Put another way, once we assume that the Underlying and GWA contract always end at the same time–an assumption we defend earlier[212]–we don't need to know why they end or, for that matter, whether a buy-down has delayed this end.

4. Second Assumption: No "Slippage"

Finally, along with assumptions about why these investments end, an assumption is needed about what happens when they end. Does GWA get a different result by investing in the contract, instead of in the Underlying Portfolio? Specifically, does the contract provide any protection from losses? The general answer is "no," as noted above, since protection below $90 is supposed to terminate when Underlying's price falls to $97.[213]

---

[208] *See supra* Part VI.A.5.
[209] *See supra* Part VII.A.2
[210] *See supra* Part VII.B.2. Delta also can be influenced in various ways by time value. For a discussion, see *infra* Appendix A.
[211] *See supra* Part VII.A.2.
[212] *See supra* Part VII.C.1.
[213] *See supra* Part VII.C.2.

But although this protection usually terminates, it does not always. There is a scenario–admittedly, an unlikely one–when this protection turns out to be helpful: a market crash or other disruption that makes it hard to liquidate the Underlying Portfolio, so GWA doesn't get the price it wants. As an example of this problem, which we call "slippage," imagine that GWA wants to terminate its investment when the Underlying Portfolio's value falls to $97, but the relevant sale is implemented too slowly, so it yields only $85, instead of $97.

Who bears this $12 of loss? Actually, GWA bears less of it with the GWA contract than with a direct investment. On the one hand, if GWA makes this investment through the contract, the loss from $90 to $85 is borne by the Bank, not GWA. On the other hand, if GWA buys the Underlying instead, then GWA–not the Bank–bears this $5 of loss. So even though the contract usually does not provide meaningful protection below $90, it actually does in this rare situation.

A key issue, then, is the value of this protection. The good news is that our difference-value test is better than the other tests at valuing it. Indeed, valuing this sort of difference is exactly what our test is designed to do.

To measure the cost of this slippage, we can value the right to sell for $90 when an asset is trading at $97. In other words, to value the risk of loss below $90, which the contract shifts to the bank, we can use the value of an out-of-the-money put option, whose exercise price is $90 when the Underlying price is $97. The term of this put option is the time it might take to liquidate the Underlying. Ordinarily, this would take minutes or even seconds. But during a market disruption, this process can take longer, so we consider 1-day and 5-day periods. As Table 12 shows, the value of this protection is quite low:

Volatility	20%	60%
No Slippage	0.00%	0.00%
1-Day Slippage	0.00%	0.31%
5-Day Slippage	0.03%	7.87%

Table 12: Cost of Slippage to Bank[214]

Not surprisingly, the cost of slippage increases with volatility, as well as with the time needed to liquidate the Underlying Portfolio. But it generally won't be an issue, except in extreme situations like the so-called quant meltdowns in August 2007 and March 2020.[215] As a result, the GWA contract usually is still a

---

[214] We express the prices in this table as percentages of the initial taxpayer capital investment of 10. This corresponds to a choice of Underlying equal to an investment of 10 by the taxpayer, with debt financing of 90, to fund overall investments with a cost of 100. If we instead chose the overall investment as Underlying, dividing by 100 rather than 10 would be appropriate, and the percentages would accordingly be one-tenth of the values reported in the table.

[215] To value execution risk in extreme conditions, the simple lognormal model used here is not the best fit. Instead, the better choice would be a more sophisticated model that accounts for the trading strategy used in the Underlying.

close substitute for the Underlying. More importantly, the difference-value test is the most effective way to measure how close a substitute it actually is.

VIII.   CONCLUSION

To sum up, the tax treatment of investments is riddled with inconsistencies, but Congress has little appetite for ambitious reforms, at least for now. Instead, Congress regularly targets specific planning strategies, such as hedge fund derivatives.

To make this sort of transaction-specific reform more effective, this Article recommends a three-step process. First, policymakers should ask whether a response is really necessary. For example, Section 1260 defends a questionable policy: rewarding taxpayers not just for investing, but for sticking with a specific investment. Arguably, instead of policing this holding period rule, Congress should liberalize it, as discussed above.[216] Yet if policymakers still want to defend this line (e.g., to avoid a tax cut for wealthy people), they need to steer clear of "good" transactions (e.g., hedges of business risks, call options on growth stocks, etc.).

This brings us to the second step: preliminary filters should screen out transactions that do not pose the relevant abuse. For example, filters should exempt taxpayers with income and assets below a specified level, common business transactions that are not tax motivated, and the like.

With the right preliminary filters, the third step–a sophisticated test–applies more narrowly, and thus is less burdensome to administer. For Section 1260, we recommend a "difference-value" test, which compares the value of exposure omitted from a hedge fund derivative to the total value of the underlying fund interest. We show why this approach is better than a simple "spread test," as well as tests based on delta and probability. We also offer novel insights about delta tests, showing how they can be manipulated fairly easily with contingencies. More generally, we recommend tests that are precise and rigorous because they can provide greater certainty, while also limiting a taxpayer's ability to use new variations of the targeted planning strategy.

---

[216] *See supra* Part IV.B.2.

IX. APPENDIX A: DELTA AND THE TIMING OF PAYMENTS IN THE GWA CONTRACT

As discussed in Part VI, the GWA contract includes a knock-out call, as well as a buy-down option. Unfortunately, the delta test does a poor job of dealing with contingencies in general, as noted above.[217] As a result, this test is not effective in analyzing GWA's knock-out call.[218] Again, the problem is that the delta test is overly focused on correlations at a moment in time.

This Appendix broadens the analysis to include two other features of the GWA contract: first, the Bank's implicit loan of $90 to GWA; and, second, GWA's right to receive $7 of its deposit when the contract terminates. As it turns out, these features have a somewhat unexpected effect on delta. This Appendix shows that if these components are precisely calibrated, they actually can bring delta close to 1.0.

We show that the sum of the deltas of GWA and the Bank in the various components of the contract is 1.0. This is not surprising. At the end of the day, the parties are jointly investing in one unit of Underlying. Through the contract, they divide the economic return between them, with some going to GWA and some staying with the Bank. On net, the sum of this exposure should simply be the Underlying–that is, delta 1.

Perhaps more surprisingly, we show that if the Bank is compensated for the time-value of money with respect to its loan, the Bank has a delta of 0. Accordingly, GWA has a delta of 1.0 in this case. We show how multiple features of the contract combine to yield this result.

This Appendix begins by analyzing the contract from the perspective of the Bank, and then turns to the perspective of GWA. As before, we base our analysis on limited information available from Joint Trial Exhibit 23-J. No. 6981-19 (U.S.T.C. filed Nov. 20, 2020). For the sake of simplicity, we also assume that the buy-down option is not used.

A. The Bank's Perspective

Let's begin with the delta of the Bank's position. At first blush, it might seem as if the Bank has no exposure to the Underlying, so its delta on the contract should be zero. This makes sense if the Bank is simply serving as an intermediary here, so that GWA–not the Bank–is supposed to get the economics of investing in Underlying.

But this is not necessarily the case. The Bank actually does have exposure to changes in Underlying's value, which comes through an implicit loan to GWA. Specifically, the Bank invests $90 up front (to buy Underlying) and receives a final payment of $90 when the contract ends.

Since the Bank receives this payment regardless of whether Underlying appreciates or depreciates, the payment seems to have no connection to Underlying. If this were true, it would have a delta of zero.

But although the amount of this payment does not turn on Underlying's performance, the timing actually does. For instance, a decline below $97 could terminate the contract, and thus enable the Bank to receive its $90 sooner. Because of this linkage to the price of Underlying, the Bank's claim to $90 has a negative

---

[217] *See supra* Part VI.B.3.c.
[218] *See supra* Part VII.B.2.

delta. Notably, this effect arises if–and only if–the Bank is not receiving a market-rate of interest on the implicit loan, an issue we address below. Table A1 shows the components of delta from the Bank's point of view.

Volatility	20%	60%
Delta of $90 Cash Paid if and when Knock-out Occurs	-2.22	-0.29
Delta of $90 Cash Paid at Termination if no Knock-out Occurs	0.97	0.06
Total Delta of $90 Cash Paid	-1.25	-0.22

**Table A1: Delta for Bank with respect to $90 Payment
(Assuming no Buy-Downs)**

There are two different effects here. First, as noted above, the Bank gets its $90 back if the contract knocks out. This means that the Bank does better as Underlying's price falls–so the delta of this feature is negative–because a decrease in Underlying's value increases the likelihood of a knock-out. In other words, the Bank's position improves as the Underlying price falls.

Second, there is also a possibility that the contract never knocks out, so the Bank receives $90 when the contract expires after twelve years. Since increases in Underlying's price make this outcome more likely, this scenario has a positive delta, as the second line of the Table shows.

The Bank's overall delta is the net of these two effects. As Table A1 shows, it is negative. This is surprising if one thinks that the Bank does not want exposure to the Underlying and seeks to act solely as an intermediary. Despite this intention, it appears that the Bank actually is betting against the performance of GWA's investment strategy!

The picture is not yet complete, however. The bank receives not only the promised payment of $90, but also interest on this money.[219] These interest payments depend upon the outstanding loan balance, which in turn depends upon the performance of Underlying. As a result, the interest payments also have a delta of their own, which is positive: If the Underlying performs well–further deferring the $90 payment–the Bank earns more interest. As a result, the interest payments increase with the performance of the Underlying.

Calculating the precise delta of the interest payments is quite complicated. Instead, we can take a short-cut: If we assume that the interest perfectly compensates for the time-value of money, then the delta of the interest payments must exactly offset the delta of the promised payment of $90. Indeed, if the Bank

---

[219] See the definition of "Basket Base Performance" on page 9 of Joint Exhibit 23-J, reducing performance by "interest expense at the Specified Rate plus the applicable Spread … on the Basket Debit Balance." GWA v. Comm'r, No. 6981-19 (U.S.T.C. filed Nov. 20, 2020). The Basket Debit Balance keeps track of the net amount owed to the Bank. A similar interest adjustment is made in the other direction if there is a Basket Credit Balance.

is perfectly compensated in this way, it no longer cares when the $90 is repaid. In this situation, described below in Table A2, the delta of the interest perfectly offsets the delta of the $90 payment.

Volatility	20%	60%
Delta of $90 Cash Paid	-1.25	-0.22
Delta of Interest Payments	1.25	0.22
Total Delta of Bank	0.00	0.00

**Table A2: Delta for Bank with respect to GWA Contract
Including Compensation for Time-Value of Money
(Assuming no Buy-Downs)**

B. GWA's Perspective

Now that we have analyzed the Bank's exposure, let's turn to GWA. In Part VII, we focused on the knock-out call. To provide a fuller picture, we now include two other features: first, a knock-in "rebate"; and, second, the interest paid to the Bank as compensation for the $90 loan.

Let's begin with the knock-in rebate. When the call option knocks-out, GWA loses the call option but gets a portion of its "premium account" back. We call this payment of $7 a "knock-in rebate." GWA receives it when the Underlying price falls below $97.[220] Table A3 shows the delta values for the knock-out call and knock-in rebate.

Volatility	20%	60%
Delta of Knock-Out Call	2.42	1.24
Delta of $7 Knock-In Rebate	-0.17	-0.02
Delta of Knock-Out Call and $7 Knock-In Rebate	2.25	1.22

**Table A3: Delta for GWA of Knock-Out Call and Knock-in Rebate
(Assuming no Buy-Downs)**

The delta values for the knock-out call are the same as in Table 8. The delta of the knock-in rebate is calculated using our usual pricing model. This component of delta is negative because an increase in the value of the Underlying delays payment of $7 to GWA, and there is no compensation for the time-value of money

---

[220] The taxpayer also receives $7 as part of the value of the surviving call option if the barrier is never hit. We need not separately account for this – it is already incorporated into our analysis of the knock-out call from the main part of this Article.

with respect to this delay. As the table indicates, the delta values for the knock-in rebate are modest, and the delta for the knock-out call and knock-in rebate combined remains greater than 1.00, just as was the case for the knock-out call in isolation.

We can also add the interest that we discussed in Part IX.A to our analysis. The Bank receives its interest payments from GWA, and the two parties thus have opposite exposures with respect to the risks of these payments. We assume that the interest payment precisely compensates the Bank for the time-value of money, as explained in Part IX.A., and we use our prior calculations to determine the delta of interest for GWA.

Table A4 summarizes the deltas for the knock-out call, the knock-in rebate, and the interest, together with the overall effect of all three components. It is the same as Table A3, except that it adds the interest component. The delta values for the interest paid to the Bank are the negatives of the corresponding deltas for the Bank in Table A2.

Volatility	20%	60%
Delta of Knock-Out Call	2.42	1.24
Delta of $7 Knock-In Rebate	-0.17	-0.02
Delta of Interest Paid to Bank	-1.25	-0.22
Total Delta	1.00	1.00

**Table A4: Delta for GWA with respect to GWA Contract
Including Knock-Out Call, Rebate, and Interest
(Assuming no Buy-Downs)**

Notably, the net of all these effects is a delta of 1.0. The bottom line, then, is that the GWA contract can have a delta of 1.0 for two different reasons. First, as we discussed in Part VII, this is true when the buy-down feature is used (assuming appropriate compensation for the time-value of money). Second, this Appendix has shown another circumstance in which delta is 1.0: the key is to account for all the features of the GWA account–that is, not just the knock-out call, but also the $7 knock-in rebate, and the interest to the Bank.

Thus, although a knock-in call wreaks havoc with the delta test in the ways we have described, the GWA contract itself may not. The reason is that the GWA contract is not simply a knock-in call but has additional features that offset risk exposures from the knock-in feature, including both compensation to the Bank for the time-value of money and the buy-down right.

So in principle, the delta test can get the right answer, but only with a scrupulous accounting of all the components of the GWA contract. But again, a taxpayer who wants delta to be very different from 1.0–as a way to defend their tax planning strategy on audit–might instead focus only on the knock-out call, and thus generate a delta very different from 1.0.

X. APPENDIX B: DIFFERENCE VALUE AND THE TIMING OF PAYMENTS IN THE GWA CONTRACT

As we have shown above, the delta test is easily confused by contingencies. In Appendix A, we showed that there are other effects, aside from the skewed result for the knock-out call. When the contract's other components are evaluated, their deltas can offset the distorted delta of the knock-out call.

We have also shown that the difference-value test is more reliable than the delta test in evaluating contingencies. Its analysis of the knock-out call is not skewed in the same way. But what about the other components of the contract? How, if at all, do they affect the difference-value analysis?

The answer is "some, but not very much." As this Appendix shows, when the Bank implicitly loans money to GWA, the implicit interest on this loan–and, in particular, the way this interest charge is structured and the assumptions we make about it–can have modest effects on the difference value. But even without accounting for these effects, the difference-value test is still pretty reliable.

To show that this is the case, we need to identify and value the implicit interest. To do so, we consider the deal first from the Bank's perspective, and then from GWA's perspective. We then analyze the effect of this interest, showing how its structure can have a modest impact on the difference value.

A. The Bank's Perspective

As discussed in Part VII.A, the Bank gets $90 of capital back when the contract either terminates early or matures after twelve years. How much interest does the contract implicitly provide to the Bank for committing this capital?

To get the answer, we need to start by calculating the initial value of this $90 payment in each of the relevant scenarios–that is, when there is a knock-out and when there is not. Table B1 shows the price of each possibility, and then adds them together to compute the overall price.

Volatility	20%	60%
Price of Right to $90 Cash Payment if and when Knock-out Occurs	82.98	89.13
Price of $90 Cash Paid at Termination if no Knock-out Occurs	3.02	0.19
Price of Guaranteed $90 Payment	86.00	89.32

**Table B1: Price for Bank of Guaranteed $90 Payment
(Assuming no Buy-Downs)**

Notably, the price of the guaranteed payment is less than $90. This is not surprising. After all, the right to receive $90 in the future, with no compensation for the time-value of money, is necessarily less than $90 today. The difference between this price and $90 is, of course, the implicit interest in the contract. If we

assume that this interest is sufficient to make the Bank whole,[221] we can infer that the interest is simply the difference between $90 and the total price in Table B1. We calculate this amount and summarize the result in Table B2.

Volatility	20%	60%
Price to Bank of GWA Contract	90.00	90.00
Price of Guaranteed $90 Payment	86.00	89.32
Price of Interest Payments	4.00	0.68

Table B2: Price for Bank of GWA Contract
(Assuming no Buy-Downs)

B. GWA's Perspective

Are these estimates of the implicit interest correct? One way to check is to analyze the contract from GWA's perspective. As noted above, the contract offers GWA two potentially valuable rights: first, a knock-out call; and, second, a knock-in rebate (i.e., a payment of $7 when the contract terminates). What are these components worth? Using our usual pricing model, Table B3 values them:

Volatility	20%	60%
Price of Knock-out Call	7.55	3.75
Price of Knock-in $7 Rebate	6.45	6.93
Price of Knock-out Call and Knock-in $7 Rebate	14.00	10.68

Table B3: Price for GWA of Knock-out Call and Knock-in Rebate
(Assuming no Buy-Downs)

Notably, the total value of these two components is greater than $10 (whether the volatility is 20% or 60%). At first blush, this is odd. The problem is that GWA is paying only $10 to enter into the contract. So what accounts for the difference? The answer, of course, is the implicit interest GWA must pay the Bank. Table B4 adds this interest to the mix, using a negative number because–unlike the call and rebate–the interest is a liability to GWA, instead of an asset. In other words, GWA owes this interest to the Bank.

---

[221] We ignore any additional fees that may be paid to the bank above and beyond pure compensation for the time-value of money.

Volatility	20%	60%
Price of Knock-out Call	7.55	3.75
Price of Knock-in $7 Rebate	6.45	6.93
Price of Interest Payments	-4.00	-0.68
Price to GWA of GWA Contract[222]	10.00	10.00

**Table B4: Price for GWA of Components of GWA Contract
(Assuming no Buy-Downs)**

As Table B4 shows, once we account for the interest, the net value of the contract to GWA is, indeed, $10. In other words, the contract actually is worth what GWA pays for it.

C. The Difference Contract

Now that we have calculated this imputed interest–$4 when the volatility is 20% and $.68 when the volatility is 60%--the question is: what effect, if any, does this interest have on the difference value. The answer is, "not much, but the precise answer depends on how the interest is structured." Since the effect is modest, the difference-value test generally should offer a reasonably accurate assessment even if we don't account for this interest but, not surprisingly, the result is even more accurate if we do.

Specifically, what difference (if any) is there between the contract, on the one hand, and a leveraged investment in the underlying, on the other? If we ignore financing, the payoff from the knock-out call and the knock-in rebate give GWA the same result as it would get from owning the Underlying and selling if and when the price reaches $97. The only difference is thus in the financing payments.

When we incorporate the financing, the answer might not change at all–or it might change slightly–depending on how the interest on this leverage is structured. To show that this is the case, Table B5 considers two scenarios.

First, assume that when we compare the GWA contract with a leveraged investment in the Underlying, we assume that the interest on the Underlying's financing must be paid up-front as a lump sum.[223] In this scenario, the difference contract has a value equal to double this interest payment, as the first line of Table B5 shows: notably, the difference contract includes the price of the interest twice– that is, once for the up-front payment (which appears in the investment in the Underlying, but not in the contract), and again for the alternative that pays over time with the total amount depending upon the performance of the Underlying (which appears in the contract, but not in the investment in the Underlying).

---

[222] The sum of the components shown may be slightly different from the total shown because of rounding errors.

[223] In terms of our example, this would be an amount of $4.00 or $0.68, depending upon whether the Underlying volatility is 20% or 60%.

Second, assume instead that this difference between the financing arrangements is eliminated. In other words, the interest payment schedule is the same for both the Underlying and the GWA contract. In this scenario, the absolute value of the difference contract would have price $0, as the second line of Table B5 shows. This is because there is no difference between the two contracts.

Volatility	20%	60%
Price of Difference Contract if Underlying Comparison Uses Up-Front Lump Sum Payment of Interest	8.00	1.36
Price of Difference Contract if Underlying Comparison Uses Same Interest Schedule as Contract	0.00	0.00

Table B5: Difference Contract Prices
(Assuming no Buy-Downs)

The values in Table B5 may be divided by the $100 price of the Underlying to arrive at the corresponding difference percentages. The bottom line is that, as Table B5 shows, variations in the structure of the financing can yield somewhat different results, but our difference method still treats the GWA contract as fairly similar to the Underlying in either case.