Shelter Check: Proactively Finding Tax Minimization Strategies via AI

Andrew Blair-Stanek, Nils Holzenberger, Benjamin van Durme

Shelter Check: Proactively Finding Tax Minimization Strategies via AI

by Andrew Blair-Stanek, Nils Holzenberger, and Benjamin Van Durme

Andrew Blair-Stanek

Nils Holzenberger

Benjamin Van Durme

Andrew Blair-Stanek is a professor of law at the University of Maryland, Nils Holzenberger is an associate professor in computer science at Institut Polytechnique de Paris, and Benjamin Van Durme is an associate professor in computer science at Johns Hopkins University. This article is based on work supported by the National Science Foundation under grant No. 2204926.

In this article, the authors explore how artificial intelligence could be used to automatically find tax minimization strategies in the tax law. Congress or Treasury could then proactively shut down such strategies. But, if large accounting or law firms develop the technology first, the result could be a huge, silent hit to the treasury.

Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

In *Peracchi*,[1] the Ninth Circuit addressed what basis a taxpayer had in a note in which he was the debtor. At the end of the opinion, the court did something unusual but wise, saying, "We take a final look at the result to make sure we have not placed our stamp of approval on some sort of exotic tax shelter."[2] Yet the court had neither the tax expertise nor the resources to fully consider whether its holding would enable future tax shelters. The court's tax shelter analysis took just two paragraphs, covering only a few possible scenarios.

Artificial intelligence will hopefully one day allow judges deciding tax cases to do a thorough check to ensure that their opinions do not enable new tax minimization strategies. We dub this future software "Shelter Check," and we are now doing the research to make it a reality. A judge's law clerk would upload a draft opinion to Shelter Check, which would see whether the opinion

[1] *Peracchi v. Commissioner*, 143 F.3d 487 (9th Cir. 1998).

[2] *Id.* at 496.

would allow new tax minimization strategies by going through its interaction with the millions of words in existing tax law authorities. If the opinion did allow tax minimization, the judge could revise the opinion to prevent that, such as by narrowing the holding.

Other branches of government could also use Shelter Check. Before congressional votes on a tax bill, staffers could run the bill's text through Shelter Check to see if it created new tax minimization opportunities. If so, the language could be amended. Similarly, before the IRS issues new Treasury regulations, revenue rulings, or other tax guidance, they could be run through Shelter Check and modified if needed.

Shelter Check could also be used by Treasury or academic researchers to search the existing body of tax law for tax minimization strategies. The IRS or Congress could then proactively shut these shelters down before well-advised taxpayers used them.

Although the term "shelter" is sometimes used only for egregious tax avoidance schemes, we use it to refer to all tax minimization strategies, including those that a court might find to be permissible tax planning. Shelter Check should be able to capture not only exotic tax shelters but also legitimate tax planning opportunities created by new tax law authorities. (Shelter Check is a catchier name than "Tax Minimization Check.")

This article explains the benefits of the Shelter Check approach, which runs contrary to the consensus about how to use AI in tax law. We include an underlying theory of tax minimization, and also make one policy proposal: The IRS should act now to address the possibility that AI researchers working for big accounting or law firms get Shelter Check working first.

Bottom-Up vs. Top-Down

How can AI identify tax minimization strategies? The consensus is that the best approach is bottom-up,[3] feeding the large quantities of data available to the IRS (for example, tax returns, Form 1099s, and public records) into machine-learning models to tease out patterns that human auditors would miss. This consensus is well founded since machine learning can extract useful insights from massive data sets. This bottom-up approach has been taken by the IRS[4] and most researchers applying AI to tax law.[5] But bottom-up attempts rarely make any use of the actual text of tax law authorities like the IRC, Treasury regulations, IRS revenue rulings, and case law. When tax law authorities are used, they are often simplified and hand-coded by humans into the models,[6] which would be prohibitively expensive to do for all tax law authorities.

Shelter Check would take the opposite approach — top-down[7] — in which the raw text of all available tax law authorities, plus any proposed new authorities, is fed into models that extract its meaning. Then other computer models would try different combinations of facts and tax law authorities to identify tax minimization strategies.

One advantage of Shelter Check's top-down approach is that it can be proactive, preventing the use of tax minimization strategies. By contrast, the bottom-up approach is reactive, discovering taxpayers' strategies only after they have been used by enough taxpayers who have filed returns reflecting the tax minimization, which may have occurred years before filing.

Top-Down Proactive Approaches

There are 1,738,185 words in the IRC and 11,350,156 words in the Treasury regulations. There are 2,340 IRS revenue rulings, 69,418 IRS letter rulings, 7,779 IRS technical advice

[3] Sarah Lawsky, "Form as Formalization," 16 *Ohio State Tech. L.J.* 114, 115 (2020).

[4] William Hoffman, "Artificial Intelligence Yields Promises and Risks for IRS," *Tax Notes*, Oct. 15, 2018, p. 406. Cara Griffith et al., "Artificial Intelligence Isn't Here Yet, but It's Already Changing Tax: Transcript," *Tax Notes Federal*, Dec. 21, 2020, p. 1880 ("governments . . . parse incredible amounts of data to find the fraudsters" (Jeff Saviano)).

[5] *E.g.*, Erik Hemberg et al., "Tax Non-Compliance Detection Using Co-Evolution of Tax Evasion Risk and Audit Likelihood," ICAIL '15: Proceedings of the 15th International Conference on Artificial Intelligence and Law (2015).

[6] Benjamin Alarie and Bettina Xue Griffin, "Using Machine Learning to Crack the Tax Code," *Tax Notes Federal*, Jan. 31, 2022, p. 661 ("Once the legal topic is selected, cases discussing the relevant legal question are then reviewed by our in-house legal research team to *translate them into structured data*.") (emphasis added).

[7] Lawsky, *supra* note 3.

memoranda, and 51,230 decided federal tax cases.[8] That vast body of law allows for many tax minimization strategies that involve combining two or more existing authorities in a way that arguably produces substantial tax savings.

No human can be familiar with all the existing tax law authorities, let alone consider all the possible ways to combine two or more of them to minimize taxes. Computers may never understand the subtleties of tax law as well as tax lawyers. But computers can handle vast amounts of data, draw potential connections between distant parts of tax law, and experiment with millions of combinations of tax law authorities to see whether they produce tax minimization strategies. There are two basic ways to approach this data in Shelter Check: New-Authority Shelter Check and Existing-Authority Shelter Check.

New-Authority Shelter Check would automate the inspection of new tax law authorities, doing what the *Peracchi* court attempted in its opinion. It would be available to judges to check draft tax case opinions, to Congress to check draft tax legislation, and to the IRS to check both draft Treasury regulations and rulings. It would examine the new authority — whether statute, regulation, or ruling — against all the existing tax law authorities to flag whether it might create a new tax minimization strategy. If a strategy were flagged and the drafters reviewed it and decided it was a serious concern, the fix would be to change the draft authority to prevent the new strategy.

Although the *Peracchi* court held for the taxpayer, it would make sense to run Shelter Check on all draft tax law authorities, even those that appear to be unfavorable to taxpayers. Professor Martin Ginsburg observed that "every stick crafted to beat on the head of a taxpayer will metamorphose sooner or later into a large green snake and bite the Commissioner on the hind part."[9] Indeed, many provisions and rulings intended to be unfavorable to taxpayers end up being used by creative taxpayers in conjunction with other tax law authorities to minimize taxes.

Existing-Authority Shelter Check could be run by anyone, including the IRS and academic researchers, to find whether existing tax law authorities can be combined to create tax minimization strategies. Such a broad search within existing law would be a white-hat search. In computer security, hackers looking to steal money or cause harm are called black-hat hackers, while those looking to find vulnerabilities with the intent of reporting them to get them fixed are called white-hat hackers. In this scenario, lawyers and accountants developing tax strategies for their clients wear the black hats.

The possibility of white-hat researchers running Shelter Check on existing authorities is one benefit of our top-down approach. The consensus bottom-up approach relies on the large amount of data the IRS collects every year, which is protected by strict confidentiality laws,[10] meaning that only a fraction of the AI researchers who want to work with it can have access. By contrast, top-down approaches like Shelter Check would work with the raw text of tax law authorities that are available to all researchers.

What if Existing-Authority Shelter Check finds a tax strategy? First, it would have to be reviewed by human tax lawyers to determine whether it is plausible. Many considerations might make a strategy implausible, such as nontax legal restrictions or economics. But if human review showed that the strategy was a plausible threat to tax revenue, there are several possible fixes. The IRS might make the strategy a listed transaction[11] or a transaction of interest,[12] either of which would require taxpayers to report its use to the IRS or face penalties.[13] Another possible fix would be to amend the IRC or Treasury regulations to explicitly prevent the strategy from working. If the strategy relied on a revenue ruling, the IRS could simply revoke the ruling.

AI models, like humans, are not perfect. Some errors are false positives: For Shelter Check, that would be strategies that it identified but that are

[8] Document counts from Tax Analysts' website.

[9] Martin D. Ginsburg, "Rethinking Limited Partnership Taxation," *Tax Notes*, Mar. 3, 1986, p. 877.

[10] Section 6103. Agents, including researchers, can have access to tax return data, but with strict limits. Section 6103(b)(5)(B)(iii).

[11] Reg. section 1.6011-4(b)(2).

[12] Reg. section 1.6011-4(b)(6).

[13] Reg. section 1.6011-4(d).

not plausible. Other errors are false negatives: For Shelter Check, that would be plausible tax strategies that it never identified. Human tax lawyers are good at determining whether a tax strategy is plausible, but not good at foreseeing how authorities might be misused to create new tax strategies. So Shelter Check's false negatives (that is, missed plausible strategies) are more dangerous, and decreasing them should be the priority.[14]

All AI models require data for training and for fine-tuning to minimize errors. There are several sources of attempted tax strategies for training and verifying Shelter Check. All existing listed transactions and transactions of interest cite the tax authorities relied on to avoid taxes. Similarly, all court cases discussing tax avoidance or judicial doctrines like substance over form cite the authorities the taxpayers rely on and describe the facts of the attempted strategy.

New-Authority Shelter Check will be substantially faster than Existing-Authority Shelter Check. Suppose that there are 100,000 existing tax law authorities and, for simplicity, that tax minimization schemes are built using just two authorities. Running one new authority through Shelter Check to see if it interacts with any of these existing authorities to create a tax minimization scheme requires checking 100,000 possible interactions. By contrast, each of the 100,000 existing authorities can interact with each of the 99,999 other authorities in 4,999,950,000 different ways. Although clever heuristics likely can reduce the computational burden without substantially increasing the error rate, Existing-Authority Shelter Check will still likely be thousands of times slower than New-Authority Shelter Check. This is good, because judges' law clerks, congressional staffers, and IRS attorneys will want Shelter Check to run quickly. Running Existing-Authority Shelter Check will likely require extensive computational resources. Unfortunately, black-hat tax advisers looking for tax minimization strategies for their clients can afford those computer resources more easily than academic researchers or the IRS. This mismatch leads us to our one policy suggestion.

[14] In statistical terms, Shelter Check should be high recall at the expense of lower precision.

Make AI-Found Strategies Reportable

Our only immediate policy suggestion is for the IRS or Congress to create a new category of reportable transactions for tax minimization strategies that taxpayers or their tax advisers find using AI. Reportable transactions are those that the IRS, through regulations, "determines as having a potential for tax avoidance or evasion."[15] A taxpayer engaging in anything the IRS has designated a reportable transaction must explicitly report the transaction to the IRS.[16] Some reportable transactions are those in which the IRS has specified the substance of the transaction,[17] such as tax shelters the IRS has come across. But several categories of reportable transactions focus on the process behind the transaction. For example, if a tax adviser tells a taxpayer about a transaction and requests that the taxpayer keep its details confidential, that's a reportable transaction.[18] As another example, if a tax adviser promises to return some of its fees if the IRS successfully challenges a strategy devised by the adviser, that's a reportable transaction.[19]

The definition of reportable transactions should be expanded to include strategies that taxpayers or their tax advisers discover using AI tools not available to the general public. We outline Shelter Check with the intent that it be a white-hat tool to prevent tax minimization. But we are still years away from having a working Shelter Check, and we plan to publish our intermediate results. (We hope to attract other white-hat AI researchers to the area, and they too will likely publish their intermediate results.) Yet savvy tax advisers, like the big accounting firms, may have a working Existing-Authority Shelter Check before we and other white-hat researchers do. If the black hats win this race, making the strategies they find into reportable transactions will prevent a massive, silent hit to the treasury.

AI is increasingly woven into many tools used by lawyers, including the search functions of Westlaw or Lexis and spreadsheet programs like

[15] Section 6707A.

[16] *Id.*; reg. section 1.6011-4(b).

[17] Reg. section 1.6011-4(b)(2) and (6).

[18] Reg. section 1.6011-4(b)(3).

[19] Reg. section 1.6011-4(b)(4).

Microsoft Excel. If a tax adviser does legal research on Westlaw, combined with modeling in Excel, and creates a tax planning strategy for a client, that should not be a reportable transaction. Westlaw and Excel are available to the general public, and only the use of AI tools not available to the general public should make a strategy a reportable transaction. The biggest danger for tax administration is a black-hat Shelter Check developed by one of the large accounting or law firms to minimize their best-paying clients' taxes substantially but quietly.

Theory of Tax Minimization

There are three categories of tax minimization strategies: (1) tax planning, which would be upheld by a court if the IRS challenged it; (2) tax avoidance, which a court would not uphold if challenged, resulting in the taxpayer's paying more in taxes and potentially penalties; and (3) tax evasion, which typically involves a badge of fraud, such as lying to the IRS, and can be criminally prosecuted. We see Shelter Check as not really having a role in fighting evasion, which relies on fraud rather than the creative combination of tax law authorities. Rather, we focus only on planning and avoidance.

The difference between tax planning and tax avoidance is simply whether the IRS would win in court in attacking the transaction using a judicial doctrine like substance over form or on a question of statutory interpretation. Shelter Check could (and should) be used to identify strategies that would fall under either.

It is not clear whether tax planning or tax avoidance is worse for the tax system. Both reduce tax revenues. Tax planning generally involves less aggressive strategies, but the IRS cannot shut down tax planning by challenging it in court. So we believe that Shelter Check should identify strategies that would be either tax avoidance or tax planning.

There are four basic approaches to minimizing tax. The first three were laid out by the economist Joseph Stiglitz,[20] while the fourth involves nuances of tax law and escaped Stiglitz's notice.

Postponement of taxes. Also known as deferral, this reduces the present value of taxes paid. Taxpayers might arrange a deduction now, with the corresponding gross income coming only in a future year. Or taxpayers might move income from this year into a future year. Another example is buying an asset like corporate stock that is expected to grow in economic value, with no tax on the gain until the stock is sold.

Tax arbitrage between two or more taxpayers. This can happen between two individuals with different tax brackets, such as when a high-bracket taxpayer transfers income to a low-bracket taxpayer. The low-bracket taxpayer might even have a U.S. tax rate of 0 percent, as with foreign corporations, domestic tax-exempt entities, or state or local governments.[21] One approach is to have losses allocated to a U.S. taxpayer, while the corresponding gains are allocated to an entity with a 0 percent U.S. tax rate.

Tax arbitrage with one taxpayer, between two or more rate schedules. Many tax systems have different rate schedules for different types of income, and this sort of strategy involves shifting income from a higher-rate schedule to a lower-rate schedule. In the U.S. system, the most common example of this is turning ordinary income or short-term capital gains into long-term capital gains or qualified dividends.

Legal cleverness. Stiglitz failed to consider this fourth possibility, probably because he was an economist and not a tax lawyer. Some tax strategies take advantage of rules that allow taxpayers to avoid tax permanently. For example, a basis calculation formula in the IRC might give an inappropriately high basis, allowing some economic gain to avoid taxation permanently. Similarly, a formula for calculating income inclusion might give an inappropriately small result. Or a transaction might inappropriately qualify for an exclusion, meaning the income will never be included. Those transactions are not postponement because they do not increase taxes in a future year. They do not involve tax arbitrage between taxpayers or rate schedules, as they often

[20] Joseph E. Stiglitz, "The General Theory of Tax Avoidance," 38 *National Tax J.* 325-337 (Sept. 1985).

[21] Section 115.

involve just one taxpayer and no rate-schedule shifting.

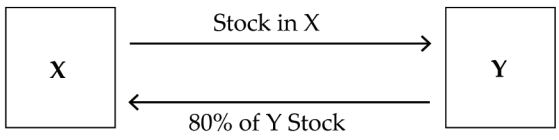
Examples

We now consider three former tax minimization strategies. For each, we discuss how Shelter Check, if it had been available, could have been used to prevent the strategy in the first place.

Example 1: PwC's Rev. Rul. 74-503, 1974-2 C.B. 117, repatriation strategy. This transaction is an example of what we called tax minimization via legal cleverness. This transaction was marketed by PwC to some of its clients that were U.S. companies with substantial cash accumulated in foreign subsidiaries, in which the earnings largely avoided U.S. taxation.[22] Had the clients ordered the foreign subsidiaries to pay this cash to them as dividends, the clients would have owed substantial U.S. corporate income tax under then-applicable law. If the foreign subsidiaries had instead used the cash to buy stock in the U.S. parent, the client would have had to pay tax because section 956 treats purchases of U.S. property (such as stock in the U.S. parent) as subpart F income. To avoid any tax, PwC devised a strategy taking advantage of part of section 956 and Rev. Rul. 74-503.

In that ruling, corporation X transferred some of its stock to corporation Y in exchange for 80 percent of the stock of Y in a section 351 transaction. (See Figure 1.) The IRS had to figure out the basis X had in the Y stock it received. Normally in a section 351 transaction, 358(a) would govern the shareholder's basis. But section 358(e) provides that section 358 "shall not apply to property acquired by a corporation by the exchange of its stock . . . as consideration in whole *or in part* for the transfer of the property to it." (Emphasis added.) So the IRS drew upon various principles to conclude that X's basis in the Y stock was $0. Normally a low basis is unfavorable to the taxpayer, resulting in greater gain when the taxpayer disposes of the property. But this seemingly taxpayer-unfavorable ruling came back to "bite the Commissioner on the hind part."[23]

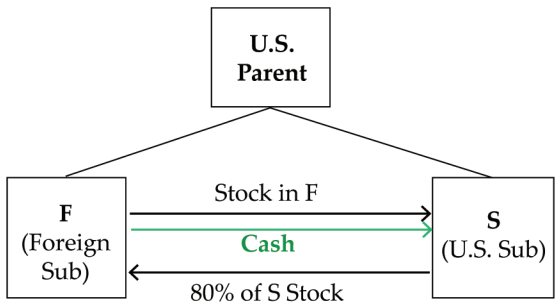
Figure 1. Facts of Rev. Rul. 74-503



Held: X Has Basis $0 in This Y Stock

The transaction that used Rev. Rul. 74-503 to avoid tax on repatriating cash from a foreign subsidiary involved setting up a new U.S. subsidiary (S) of the U.S. parent. The foreign subsidiary (F) that had the cash would then contribute F stock plus the cash to S in exchange for S's stock in a section 351 transaction, as shown in Figure 2.

Figure 2. PwC's Repatriation Strategy



Argued: F Has Basis $0 in This S Stock

PwC said that Rev. Rul. 74-503 applied, with F playing the part of X and S playing the part of Y. As a result, F's basis in the S stock received was supposedly $0. Because section 956 measures the subpart F income by the U.S. property's basis,[24] that meant that the U.S. parent had zero tax on the cash's repatriation.

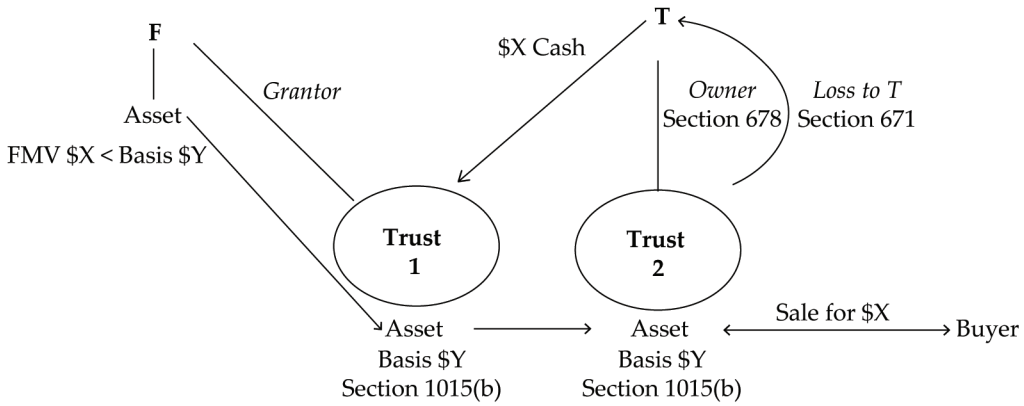
How would Shelter Check have prevented this strategy? Suppose that back in 1974 the IRS had had New-Authority Shelter Check to scan its draft of Rev. Rul. 74-503. Shelter Check would have flagged that the ruling's zero-basis holding could be combined with section 956 in the strategy later devised by PwC. The IRS could then have decided not to issue Rev. Rul. 74-503. (Indeed, the IRS revoked it in 2006 after PwC's

[22] *Barnes Group Inc. v. Commissioner*, T.C. Memo. 2013-109 (noting that this strategy came from PwC's internal "Ideasource database"). One of the authors briefly did some legal research for the taxpayer in that case.

[23] Ginsburg, *supra* note 9.

[24] Section 956(a) flush language.

Figure 3. Distressed Asset Trusts Strategy



strategy came to light.[25]) Alternatively, the IRS could have issued the ruling, but with an explicit limitation that its holding would not apply if any consideration other than X stock had been contributed to Y. That also would have foiled PwC's later strategy.

This example also illustrates how tax law is particularly amenable to using AI to find abusive strategies. Tax law is unique in that it boils down to a concrete number — money owed to the government.[26]

Example 2: *Summa Holdings* arrangement. This transaction is another example of what we call the tax minimization strategy of legal cleverness. It involved the combination of the Roth IRA provisions with the domestic international sales corporation provisions. Congress granted DISCs an explicit tax exemption for commissions on exports to cut taxes on exports and reduce the trade deficit. The taxpayers in *Summa Holdings*[27] and other taxpayers who used this strategy owned businesses that exported goods. They also owned Roth IRAs, tax-favored retirement accounts explicitly allowed by Congress that allow tax-free growth and withdrawals. They had their Roth IRAs own DISCs, which collected commissions on their businesses' exports. These commissions were deductible by the taxpayers' businesses and excluded from the DISCs' income.[28]

The IRS challenged these tax savings, attempting to recharacterize the commissions as nondeductible deemed dividends to the businesses' owners, followed by excess Roth IRA contributions. But the Sixth Circuit allowed the tax savings, noting that Congress had explicitly enacted both provisions the taxpayers were using. In other words, the taxpayers' strategy was valid tax planning, not tax avoidance.

Shelter Check would seek to catch all tax minimization strategies — not only tax avoidance like PwC's combination of Rev. Rul. 74-503 and section 956, but also tax planning like a Roth IRA holding a DISC. Both tax avoidance and tax planning siphon money from the treasury, and both could be caught before being used — with the proper software.

Suppose congressional staffers had had access to New-Authority Shelter Check when Congress was preparing to enact the Roth IRA provisions in 1997. Shelter Check would have identified the possible combination of those provisions with the DISC provisions that Congress had added over two decades earlier. Congress could have expressly barred Roth IRAs from directly or indirectly owning DISCs. That would have prevented the *Summa Holdings* strategy entirely.

Example 3: Distressed asset trusts (Notice 2008-34, 2008-1 C.B. 645). This transaction, as

[25] Rev. Rul. 2006-2, 2006-1 C.B. 261.

[26] *Cf.* Lawsky, *supra* note 3, at 122 ("Other forms — immigration forms, for example — primarily collect and organize information. Tax forms, however, do things with numbers.").

[27] *Summa Holdings Inc. v. Commissioner*, 848 F.3d 779 (6th Cir. 2017).

[28] Section 995(b)(1)(F)(i) and (g).

shown in Figure 3, is an example of the tax minimization strategy of arbitrage between two taxpayers. Let's say you have a U.S. taxpayer T with income to shelter. Through advisers, T is put in touch with a foreign taxpayer F that is not subject to U.S. tax and owns an asset with substantial built-in loss. Let's say that F's basis in the asset is $Y, but it is now worth much less, $X.

F would contribute the asset to a grantor trust, called Trust 1. Section 1015(b) states that the asset still has basis $Y. T pays $X in cash to become the beneficiary of Trust 1. Then the asset is transferred to a second trust, Trust 2, designed so that T has the rights described in section 678(a)(1), which means T is treated as the owner of Trust 2. Again, section 1015(b) governs basis, saying that Trust 2 holds the asset with basis $Y. Then Trust 2 sells the asset to some unrelated third party for the $X it is worth, resulting in the recognition of the full built-in loss. Under sections 678(a)(1) and 671, this loss goes onto T's tax return, sheltering T's income. Other than taxes, this transaction is a wash for T, who paid $X cash to F but got back $X cash on the sale to a third party.

If Congress had New-Authority Shelter Check when it enacted section 678, it would have been alerted to this strategy. Congress could have amended section 1015(b) to set the asset's basis to $X in these circumstances.

This strategy also demonstrates how Shelter Check need not have a full understanding of all legal concepts. Section 678 applies if a beneficiary has "a power exercisable solely by himself to vest the corpus or the income therefrom in himself," which is likely a hard legal concept for AI to understand or represent. The Shelter Check modeling could simply assume that T might have that power over Trust 2 in finding the Notice 2008-34 strategy. But confirming that a trust can have that power in the real world is precisely why it would be necessary to have a human tax lawyer review potential tax minimization strategies identified by Shelter Check.

The Technological Challenges

The AI capabilities required for Shelter Check do not yet exist, or, if some accounting firm or law firm has developed them, their existence remains secret. We are working to develop the required capabilities and welcome the efforts of other white-hat researchers. The required AI capabilities fall into two basic categories — natural language understanding and strategy modeling.

We are focused on a form of natural language understanding called semantic parsing. Semantic parsing is concerned with extracting the logical, structured representation of the meaning of language.[29] For example, with a revenue ruling, a semantic parse would extract the relevant parties, their relations, their transactions, and the holdings. It might also extract the reasoning the IRS used to reach its holdings. There are also other approaches we are exploring that involve less explicit structure but rely on huge computational models with hundreds of billions of artificial neurons.[30]

There are two basic types of legal authority — statutory, which sets out rules in the abstract, and case-based, which gives facts and explains how the law applies to them. In tax law, the IRC and most of the Treasury regulations are the statutory type authorities. Tax law's case-based authorities include not only decisions by courts on tax law cases but also the Treasury regulations' examples, revenue rulings, private letter rulings, technical advice memoranda, field service advice, general counsel memoranda, and generic legal advice memoranda. The two types of legal authority pose different challenges for semantic parsing, and we are working on technologies for the semantic parsing of both.[31]

Any area of AI benefits from having lots of data to train predictive models, and semantic parsing is no exception. The raw text of many tax law authorities is now available, and that adds up to a substantial corpus of text. Even more promising is transfer learning, in which an AI

[29] Li Dong and Mirella Lapata, "Language to Logical Form With Neural Attention," Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers.

[30] Tom B. Brown et al., "Language Models Are Few-Shot Learners," 33 *Advances in Neural Information Processing Systems* 1877 (2020); Cade Metz, "Meet GPT-3. It Has Learned to Code (and Blog and Argue)," *The New York Times*, Nov. 24, 2020, at D6.

[31] *E.g.*, Nils Holzenberger et al., "A Dataset for Statutory Reasoning in Tax Law Entailment and Question Answering," Proceedings of the 2020 Natural Legal Language Processing Workshop (statutes); Andrew Blair-Stanek and Benjamin Van Durme, "Improved Induction of Narrative Chains via Cross-Document Relations," 11th Joint Conference on Lexical and Computational Semantics (2022) (case-based).

model is trained on a large set of related data before being trained on the domain of interest (here, tax law). There is now a huge trove of that related data for our purposes. The Harvard Law School Library has made the text of virtually all published U.S. case law available to researchers,[32] and the U.S. Code and Code of Federal Regulations are available from government sources in structured, computer-readable format.

This huge quantity of available text upends one of the arguments made in favor of the conventional bottom-up approach of using AI on the IRS's huge trove of return data, as opposed to the top-down approach we advocate. The bottom-up approach can use more data, since the IRS receives 261 million tax returns and 4.6 billion information returns each year.[33] But with several billion words of legal text now available, plus transfer learning technology to use nontax law text to train a tax-law-focused model, the top-down approach looks increasingly feasible.

Shelter Check's second technological challenge is strategy modeling. Once we have parsed several tax authorities, how can we combine them to create a tax minimization strategy? For example, in PwC's plan to combine Rev. Rul. 74-503 and section 956, simply understanding the meaning of the text of both authorities is not enough to know how to create the strategy. Rather, that also requires several modestly creative steps: creating a U.S. subsidiary to stand in for Y in the ruling, plus later transferring the cash from the U.S. subsidiary to the parent company. AI can exhibit this limited sort of creativity.

One possible approach to tax strategy modeling is reinforcement learning, which is a branch of AI concerned with agents that interact with an environment to fulfill goals. To fulfill their goals, agents take actions, which elicit reactions from the environment and may yield rewards. For tax strategy modeling, the agents are the taxpayers, whose goal is to maximize their net worth. The actions they can take are any legal action (for example, forming a subsidiary or transferring cash), and the environment consists of all tax law authorities. Reinforcement learning has managed amazing feats,[34] including the development of winning strategies for complex multi-player video games — making it a promising approach to shelter modeling.

A second possible AI approach to tax strategy modeling is language modeling. Language models can generate novel, coherent, logically plausible text, ranging from poetry to news stories. A language model trained on legal authorities and then tax law authorities could be used to systematically generate possible tax minimization strategies, which would then be evaluated using the semantic parses of all existing tax law authorities.

A third possible approach is modeling the associations between different areas of tax law. Computer scientists Jamshid Sourati and James Evans built a data set of links between scientific publications, authors, and scientific entities, such as types of physical materials.[35] They then used AI methods applied to that data set to mimic associations that researchers make. Astonishingly, this produced relevant scientific hypotheses. Shelter modeling faces a challenge similar to coming up with scientific hypotheses: associating tax law authorities from different areas of tax law that can interact in unexpected ways.

Conclusion

We have proposed a new approach to using AI in tax law. Rather than the consensus bottom-up approach of feeding the torrent of data the IRS receives into AI models, we propose the top-down approach of understanding the text of all tax law authorities and modeling how these authorities may be manipulated in new and unusual ways to minimize taxes. Our approach has several advantages, including being proactive, allowing outside researchers to help, and using the actual text of tax law authorities without requiring human lawyers to manually encode them. But our approach has a dark side —

[32] *See* Harvard Law School, Caselaw Access Project (2022).

[33] IRS Data Book 2021, at 4, Table 2 (2022); *id.* at 54, Table 22.

[34] Reinforcement learning has been applied to tax economics with promising results. Stephan Zheng et al., "The AI Economist: Taxation Policy Design Via Two-Level Deep Multiagent Reinforcement Learning," *Sci. Adv.* 8 (2022).

[35] Jamshid Sourati and James Evans, "Accelerating Science With Human Versus Alien Artificial Intelligences," arXiv (Apr. 12, 2021).

it also might be used by tax advisers to find new strategies. To counter that, we propose that the IRS immediately make tax strategies found using AI into reportable transactions.

We should note that the novel top-down approach we propose is not mutually exclusive with the bottom-up approach. Combining the two may turn out to be the most powerful approach to attacking tax minimization. For example, top-down semantic parses of tax law authorities and models of possible strategies may be entered into bottom-up models that review reams of tax return data. Conversely, the plentiful data available to the IRS might be used as input to help train top-down models. ■