Article

Subjective Costs of Tax Compliance

Jonathan H. Choi[†] and Ariel Jurow Kleiman^{††}

This Article introduces and estimates the "subjective costs" of tax compliance, which are costs of tax compliance that people experience directly and individually. To measure these costs, we conducted a survey experiment assessing how much taxpayers would pay to reduce the unpleasantness associated with filing a tax return. The experiment revealed that taxpayers are more concerned about inadvertent mistakes in their tax filings than the time spent on compliance. Respondents also only ascribed meaningful value to eliminating all tax compliance work; they ascribed essentially no value to marginal time savings. Additionally, taxpayers were indifferent between simplification services offered by a private company versus the government.

[†] Jonathan H. Choi is a Professor of Law at the University of Southern California Gould School of Law.

^{††} Ariel Jurow Kleiman is a Professor of Law at Loyola Law School, Los Angeles.

For comments on the draft at various stages, thank you to Trudy Cameron, June Carbone, Chris Evans, Ed Fox, Brian Galle, Ari Glogower, Jacob Goldin, Kristin Hickman, Claire Hill, James Hines, Zachary Liscow, Michael Love, Ruth Mason, Susie Morse, Joe Mitchell-Nelson, Jeesoo Nam, Philip Postlewaite, Alex Raskolnikov, David Schizer, Manoj Viswanathan, and Heydon Wardell-Burrus. This Article was presented at the Davis Polk & Wardell Tax Policy Colloquium at Columbia Law School, the Junior Tax Scholars Workshop, the Max Planck Institute for Tax Law and Public Finance, the Northwestern University Tax Colloquium, the Oxford-MIT-Michigan-Munich-Georgetown (OMG) Transatlantic Tax Talks, the University of Florida College of Law Faculty Workshop, the University of Michigan Legal Theory Workshop, the University of San Francisco Law School Tax Policy Colloquium, the National Tax Association annual conference, and the University of Virginia Tax Conference. We are grateful for the comments we received at those events. For their careful work, thank you to Toph Beach, Philip de Sa e Silva, Mary Fleming, Carly Heying, Adam Kolb, E. Isabel Park, and the rest of the editorial team at the University of Minnesota Law Review. Copyright © 2024 by Jonathan H. Choi and Ariel Jurow Kleiman.

These findings have important implications for theory and policy. From a theoretical perspective, these survey results call into question the nearly universal practice of using market wages to monetize the time that people spend on tax compliance work. Indeed, our results suggest that people value their tax compliance time at a rate much lower than their hourly wage. Regarding policy, these findings counsel policymakers to think big when it comes to reducing tax compliance costs and to focus on simplifications that reduce mistakes rather than merely saving time. They also suggest that policymakers need not be overly concerned about mistrust of government in the context of tax simplification and automation services.

TABLE OF CONTENTS

Introduct	tion	1259			
I. What We Do. Don't. and Should Know1264					
A.	What We Know: Objective Compliance Costs	1265			
B. What We Don't Know: Subjective Compliance					
	Costs	1267			
	1. Empirical Research on Subjective Costs	1270			
	2. Potential Subjective Benefits of Taxation	1272			
С.	What We Should Know: Monetizing Subjective				
	Costs	.1273			
II. Surve	ey and Results	1276			
A.	Methodology	1277			
	1. Discrete Choice Surveys Generally	1277			
	2. Survey Details and Implicit Prices	1280			
	3. Survey Sample	1282			
В.	Results	1283			
	1. Discrete Choice Experiment Results	1283			
	2. Qualitative Context	1295			
С.	Potential Objections and Robustness Checks	1301			
	1. External Validity	1301			
	2. Cheap Talk	1304			
	3. Statistical Validity, Survey Fatigue, and				
	Attenuation Bias	1307			
III. Impl	lications for Theory and Policy	1308			
Â.	Implications for Theory	1308			
В.	Implications for Policy	1310			
	1. Think Big	1311			
	2. Minimize Mistakes	1314			
	3. Public or Private, Taxpayers Don't Care	1317			
Conclusio	on	1319			
Appendix	Χ	1321			
А.	Survey Contents	1321			
	1. Respondent-Specific Data	1321			
	2. Disclosure of Possible Choice Sets	1322			
	3. Choice Card Sets	1322			
	4. Fractional Factorial Design	1323			
	5. Power Analysis and Sample Size	1324			
	6. Qualitative Questions				
B.	Attrition and Rejected Respondents	1326			
<i>D</i> .	1. Comprehension Check	1326			
	2 Attention Check	1327			

1258 MINNESOTA LAW REVIEW [108:1255

	3. Protest Responses	1328
	4. Attrition Statistics	1329
С.	Discrete Choice Model	1329
D.	Regression Table	1331
Е.	Confirming Independence of Irrelevant	
	Alternatives	

The income tax return inspires aggravation, confusion, boredom, and even anger.¹ Tax season is a time of handwringing and foot-dragging that climaxes in a flurry of online forms submitted at 11:59 p.m. on April 15th.² Scholarship on tax compliance costs largely echoes this narrative. Commentators and policymakers criticize the large amounts of time and resources that taxpayers must devote to complying with their tax obligations.³ Despite the fact that voluntary tax compliance

2. See Amélie Poulain, The Simpsons – Homer on Tax Day, YOUTUBE (Apr. 15, 2014), https://www.youtube.com/watch?v=ZnJcZ-5P8hE (showing a clip from The Simpsons: The Trouble with Trillions, Season 9, Episode 20, depicting Homer's last-minute effort to file his taxes).

3. E.g., Joseph Bankman, Who Should Bear Tax Compliance Costs? 2 (Stanford L. Sch., John M. Olin Program in L. & Econ., Working Paper No. 279, 2004), http://ssrn.com/abstract=519783 [https://perma.cc/NB2S-UC6E] ("Compliance costs substantially reduce the social gains from taxation"); Scott Hodge, The Tax Compliance Costs of IRS Regulation, TAX FOUND. (Aug. 23, 2022), https://taxfoundation.org/tax-compliance-costs-irs-regulations [https:// perma.cc/J4FA-6PUM] ("A less direct cost [of taxation] is the precious time taken out of our lives to comply with a Byzantine tax code that requires billions of hours completing mountains of IRS paperwork and tax returns."); Jason J. Fichtner & Jacob M. Feldman, The Hidden Costs of Tax Compliance, MERCATUS CTR. GEORGE MASON UNIV. (May 20, 2013), https://www.mercatus.org/research/ research-papers/hidden-costs-tax-compliance [https://perma.cc/9KFT-Q6R3] (tallying and critiquing the "hidden costs" of taxes, which include compliance costs); Press Release, Elizabeth Warren, Sen., U.S. Senate, Senator Warren Leads 22 Colleagues in Introducing the Tax Filing Simplification Act of 2022 (July 13, 2022), https://www.warren.senate.gov/newsroom/press-releases/

^{1.} See, e.g., Expert: Confusion Reigns Supreme as Americans Wait Until the Last Minute to File Their Tax Returns, CBS N.Y. (Apr. 17, 2022), https:// www.cbsnews.com/newyork/news/expert-confusion-reigns-supreme-as -americans-wait-until-the-last-minute-to-file-their-tax-returns [https://perma .cc/G5B2-BH2M] (reporting survey results that 21% of respondents were "too confused" about their personal tax situation); Beverly Moran, Why Can't the IRS Just Send Americans a Refund – Or a Bill?, CONVERSATION (Mar. 22, 2021), https://theconversation.com/why-cant-the-irs-just-send-americans-a-refund-or -a-bill-156733 [https://perma.cc/5SDP-KNEA] (describing tax filing as "unpleasant," "onerous," and "tedious"); Ryan Lasker, Why Are Taxes So Confusing?, MORNING BREW: MONEY SCOOP (Jan. 3, 2022), https://www.morningbrew.com/ money-scoop/stories/2022/01/03/why-are-taxes-so-confusing [https://perma.cc/ 5FP5-FMF5] (considering reasons why the U.S. tax system is confusing, in part by comparing the U.S. system to the relatively simpler system of Estonia); On the Media, Why So Tedious, Taxes?, WNYC STUDIOS (Apr. 6, 2017), https://www .wnycstudios.org/podcasts/otm/segments/taxes-tedious [https://perma.cc/N4Z7 -AURF] (ascribing U.S. tax system complexity, in part, to lobbying by tax preparation companies).

and tax morale in the United States are high relative to peer nations,⁴ there is a general sense that U.S. taxpayers are dissatisfied with the process of filing their income tax returns.

Given the level of attention and resources that policymakers and scholars devote to tax compliance costs, one might reasonably assume that we have a decent sense of how significant these costs are for individual taxpayers. It turns out, we don't. Although researchers know a good deal about the time and money that people spend on tax compliance activities,⁵ we know almost nothing about exactly how unpleasant or burdensome people find these activities to be.⁶

This Article fills that gap, asking just how burdensome the tax filing process is to ordinary taxpayers. We refer to this burden as the "subjective costs" of taxation. Subjective costs are the costs of tax compliance that people experience directly and individually. To estimate these costs, we use something called a discrete choice experiment to measure taxpayers' willingness to pay to reduce or eliminate the burdens (or benefits) of tax compliance activities.⁷ By asking taxpayers whether they would be

5. *E.g.*, Joshua D. McCaherty, *The Cost of Tax Compliance*, TAX FOUND. (Sept. 11, 2014), https://taxfoundation.org/data/all/federal/cost-tax-compliance [https://perma.cc/NE3H-23JG] (providing tax compliance time and expenditure estimates).

senator-warren-leads-22-colleagues-in-introducing-the-tax-filing-simplification -act-of-2022 [https://perma.cc/3BMA-MUWF] (introducing proposed legislation to make the tax filing process less time-consuming and expensive).

^{4.} E.g., Yair Listokin & David M. Schizer, I Like to Pay Taxes: Taxpayer Support for Government Spending and the Efficiency of the Tax System, 66 TAX L. REV. 179, 185–86 (2013) (discussing high tax morale among U.S. taxpayers); James Alm & Benno Torgler, Culture Differences and Tax Morale in the United States and in Europe, 27 J. ECON. PSYCH. 224, 229–44 (2006) (finding U.S. tax morale to be high relative to peer European nations); Tax Administration 2021: Comparative Information on OECD and Other Advanced and Emerging Economies, ORG. FOR ECON. COOP. & DEV. 198–99 tbl.D.12 (2021), https://doi.org/10 .1787/cef472b9-en (providing on-time filing rates for the personal income tax among OECD members and listing the United States with one of the highest rates (99.7% in 2019)).

^{6.} See Bankman, supra note 3, at 2 (noting that compliance costs estimates fail to account for the "anxiety many taxpayers feel when filing their return").

^{7.} Discrete choice experiments ask survey respondents to choose between hypothetical goods/services with varying attributes, one of which is typically price. By presenting respondents with multiple options, including the option to purchase nothing, discrete choice experiments mimic the type of decision-

willing to pay to reduce various tax compliance burdens, we can impute a monetary value to those burdens.

Based on existing literature, we hypothesized that people might find tax compliance activities to be particularly aggravating, tedious, and generally unpleasant. Separately, people might worry about making mistakes on their return or being audited.⁸ Therefore, we asked respondents whether they would be willing to purchase tax services that would reduce or eliminate their time spent on taxes, and/or eliminate their risk of making a mistake or being audited, at various hypothetical prices.

We have three primary findings, all of which add nuance to the conventional wisdom about compliance costs of taxation. First, the sources of aversion to tax compliance are different than previously understood. Commentators tend to focus on the time spent on tax filing and the aggravation and tedium associated with the task.⁹ Accordingly, prominent simplification reforms often seek to reduce the time people spend on their return—for example, by reducing paperwork or the calculations necessary to claim itemized deductions by increasing the standard deduction.¹⁰ But we found that taxpayers are more bothered by possible mistakes than by any general aggravation or unpleasantness associated with the compliance tasks themselves. That is, respondents in our survey were willing to pay more to eliminate their risk of making a mistake or being audited (about \$72, on average) than they were willing to pay to

making that people face in real-life market decisions. By altering the attributes of the various goods or services offered, respondents' choices can reveal how they value each attribute. See *infra* Part II.A.1 for further discussion of discrete choice experiments.

^{8.} For a brief discussion of other possible subjective costs aside from aggravation and anxiety, see *infra* note 45 and accompanying text.

^{9.} See Hodge, supra note 3 ("Americans will spend more than 6.5 billion hours complying with IRS tax filing and reporting requirements in 2022."); On the Media, supra note 1, at 01:13 (providing an example of a taxpayer's meticulous and time-consuming tax preparation process and the associated feelings).

^{10.} See, e.g., Erica York & Alex Muresianu, The Tax Cuts and Jobs Act Simplified the Tax Filing Process for Millions of Households, TAX FOUND. 4 (July 2018), https://files.taxfoundation.org/20180806164308/Tax-Cuts-and-Jobs-Act -Simplified-the-Tax-Filing-Process-for-Millions-of-Households.pdf [https://perma.cc/L56Z-8FWX] (explaining how increasing the standard deduction simplifies tax filing).

eliminate all the time they spend on tax compliance activities (about \$53, on average).¹¹

Second, taxpayers don't value marginal time savings. Respondents in our survey ascribed essentially no value to a service that would shave an hour off their tax compliance time.¹² They ascribed only a slightly higher value—but still not much to a service that would cut their tax compliance time in half.¹³ Respondents only seemed to meaningfully value a service that would eliminate *all* tax compliance work, with a willingness to pay \$53 total (i.e., about \$6.40 per hour).¹⁴ When it comes to saving time on their taxes, it's an all-or-nothing calculus for taxpayers.

Third, respondents were indifferent as to whether the government provides tax simplification services directly or outsources such services to a private third party, so long as the service itself is of equivalent quality.¹⁵ This finding challenges the assumption that people inherently mistrust the government to provide tax simplification and automation services.¹⁶

In addition to our three primary findings, our results differ in magnitude from prevailing compliance-cost estimates. Our survey suggests that people value their tax compliance time at an amount less than their marginal wage rate. As mentioned above, researchers have a good sense of the amount of time that people spend on tax compliance.¹⁷ To monetize this time, analysts typically multiply it by some ascribed hourly wage.¹⁸ This method is used to value people's time for nearly all cost-benefit

^{11.} See infra Part II.B.1.a.

^{12.} See infra Part II.B.1.b.

^{13.} See infra Part II.B.1.b.

^{14.} See infra Part II.B.1.

^{15.} See infra Part II.B.1.c.

^{16.} Editorial, *Would You Trust the IRS to Do Your Taxes?*, BOS. HERALD (Mar. 23, 2023), https://www.bostonherald.com/2023/03/23/editorial-would-you -trust-the-irs-to-do-your-taxes [https://perma.cc/HF3S-Q9FT] (discussing trust in government in the context of a free public tax filing system).

^{17.} See McCaherty, *supra* note 5 (estimating that, in 2012, Americans collectively spent 1.35 billion hours filing individual taxes, and businesses spent another 1.89 billion hours filing taxes).

^{18.} See Fichtner & Feldman, supra note 3, at 10 tbl.2 (showing variations in wage assumptions across multiple studies).

analyses across government agencies.¹⁹ Our findings suggest that these market-wage-based estimates do *not* reflect the welfare burden that people experience.²⁰ This result comports with other behavioral-economics research on how people value time.²¹ It also suggests that the dollar-value benefit of large-scale government reforms may be lower than other commonly cited estimates.²²

Our findings have important implications for both theory and policy. Regarding theory, our findings challenge the nearly universal practice of using market wages to value the time that people spend on tax compliance work.²³ Regarding policy, perhaps the most important takeaway is that policymakers should think big when it comes to reducing tax compliance costs.²⁴ Since taxpayers ascribe little value to marginal time savings, policymakers should not bother with reforms that only save taxpayers a small amount of time. Reforms that would eliminate all (or nearly all) tax compliance activities—whether for all taxpayers or for certain subgroups of taxpayers—will be disproportionately more valuable than reforms that merely shave an hour or two off their tax preparation work.

20. See infra Part III.A.

21. See, e.g., Patrick Lloyd-Smith et al., Decoupling the Value of Leisure Time from Labor Market Returns in Travel Cost Models, 6 J. ASS'N ENV'T & RES. ECONOMISTS 215, 215 (2019) ("[R]espondents value their leisure time heterogeneously and substantially differently from their implied wage rate."); Jay R. Cherlow, Measuring Values of Travel Time Savings, 7 J. CONSUMER RSCH. 360, 362–63 (1981) (summarizing studies valuing travel-time savings, the majority of which find that travel time is worth some fraction of wage rate).

22. See infra Part II.B.1.d.

23. See generally Erica York, Reviewing Different Methods of Calculating Tax Compliance Costs, TAX FOUND. 3–4 (Aug. 2018), https://files.taxfoundation .org/20180821100528/Reviewing-Different-Methods-of-Calculating-Tax -Compliance-Costs.pdf [https://perma.cc/XRK2-DN32] (describing different hourly rates used to monetize time costs).

24. See infra Part III.B.1.

^{19.} This valuation method has been challenged elsewhere. See, e.g., Adam M. Samaha, Death and Paperwork Reduction, 65 DUKE L.J. 279, 328–36 (2015) (explaining why hourly wages are not an accurate basis for monetizing time burdens); Peter Feather & W. Douglass Shaw, Estimating the Cost of Leisure Time for Recreation Demand Models, 38 J. ENV'T ECON. & MGMT. 49 (1999) (proposing an alternative, non-wage-based method for measuring an individual's opportunity cost of leisure time); Kenneth E. McConnell & Ivar Strand, Measuring the Cost of Time in Recreation Demand Analysis: An Application to Sportfishing, 63 AM. J. AGRIC. ECON. 153, 153–54 (1981) (arguing that time cost should be valued at some fraction of the wage rate less than one).

Our survey results also suggest that policymakers should focus on reforms that reduce concerns about inadvertent mistakes in tax filings.²⁵ Congress could, for instance, simplify complex rules and eliminate "traps for the unwary" in the tax code.²⁶ The IRS could reduce taxpayer anxiety by better publicizing remediation programs for taxpayers who make (good faith) mistakes on their tax returns, as well as programs for those who can't afford to pay their taxes right away.²⁷

Finally, we also observe that respondents in our survey did not place a premium on the nature (public or private) of the tax service provider. We draw two related policy implications from this finding. First, policymakers should decide on the public/private nature of tax services based on factors other than taxpayers' preferences. Second, mistrust of government is unlikely to dissuade taxpayers from using a government-provided service.

The Article proceeds as follows. In Part I, we provide a broad overview of the scholarship on tax compliance costs. This scholarship focuses almost exclusively on what we term the "objective costs" of tax compliance, that is, costs that are measured by an external value such as market wages. We also highlight a gap in the literature related to "subjective costs"—the personal costs that people experience in preparing and filing their taxes—and survey the limited research on these costs. Part II describes our survey methodology and results, including both quantitative results and qualitative survey responses. We also discuss potential objections to and robustness checks of our methods. Part III offers implications for theory and policy.

I. WHAT WE DO, DON'T, AND SHOULD KNOW

This Part describes the current state of the research on the compliance costs of taxation. A large literature estimates these costs by aggregating the amount of money spent on tax assistance, as well as ascribing a monetary value to the time that taxpayers spend complying with tax laws. However, these estimates

^{25.} See infra Part III.B.2.

^{26.} See, e.g., Robert Greenstein et al., Reducing Overpayments in the Earned Income Tax Credit, CTR. ON BUDGET & POL'Y PRIORITIES 1 (Jan. 31, 2019), https://www.cbpp.org/sites/default/files/atoms/files/4-30-13tax.pdf [https://perma.cc/U2UH-A69H] ("EITC errors occur primarily because of the complexity of the rules surrounding the credit. Most of them reflect unintentional errors, not fraud.").

^{27.} See infra notes 190-98 and accompanying text.

can't capture the personal, idiosyncratic costs that people experience when filing a tax return, which we term the "subjective costs" of tax compliance. This Part defines "subjective costs" and explores the limitations of the current compliance-costs research in addressing them.

A. WHAT WE KNOW: OBJECTIVE COMPLIANCE COSTS

Compliance costs are costs that taxpayers incur in complying with their tax obligations. These costs traditionally include time spent planning for, preparing, and filing taxes, as well as money spent on tax software or professional preparers.²⁸ Compliance costs are one of the three main costs that the tax system imposes on taxpayers, alongside the taxes themselves and the efficiency costs incurred when taxpayers change their behavior to avoid taxes.²⁹ Across tax types, methodologies, and countries, researchers have found that compliance costs are "high" and "significant."³⁰

As the word "cost" implies, measuring compliance costs requires monetizing the time and effort expended in complying with tax laws.³¹ Thus, the vast majority of the research in this area focuses on ascribing a monetary value (sometimes

30. Id. at 457.

^{28.} The Paperwork Reduction Acts (PRA) of 1980, 1995 and subsequent amendments require the IRS to track the time that taxpayers spent on "record-keeping," "tax planning," "form completion and submission," and "all other," as well as how much money they spend to have third parties perform these tasks. See 44 U.S.C. §§ 3502, 3506–3507 (requiring agencies to evaluate "burdens," which include time and expenditures, before collecting information from the public); INTERNAL REVENUE SERV., 1040 (AND 1040-SR): INSTRUCTIONS 106–07 (2022), https://www.irs.gov/pub/irs-pdf/i1040gi.pdf [https://perma.cc/HGA6-BE27] (identifying time burden estimates for each of the statutory tracking requirements).

^{29.} Chris Evans, *Taxation Compliance and Administrative Costs: An Overview, in* TAX COMPLIANCE COSTS FOR COMPANIES IN AN ENLARGED EUROPEAN COMMUNITY 447, 449 (Michael Lang et al. eds., 2008).

^{31.} *Id.* at 451 (defining tax compliance costs to include time and labor); Off. of Mgmt. & Budget, *Estimating Paperwork Burden*, WHITE HOUSE: PRESIDENT BARACK OBAMA ARCHIVES (Oct. 4, 1999), https://obamawhitehouse.archives .gov/omb/fedreg_5cfr1320 [https://perma.cc/TJ6T-XKQU] (discussing "monetizing burden hours" to comply with requirements of the Paperwork Reduction Act); 44 U.S.C. § 3502(2) (defining the "burden" of paperwork in the Paperwork Reduction Act to include "time, effort, or financial resources").

expressed as a percentage of GDP) to tax compliance activities.³² This value includes time spent on tax preparation, as well as money spent on software, tax preparation assistance, and tax expertise.³³ Tabulating expenditures is straightforward; monetizing time is somewhat more complicated.³⁴ To convert time spent on tax compliance into a dollar amount, policy analysts typically multiply the hours spent by some market-determined hourly rate—usually the average hourly wage for all workers or for professional workers.³⁵ The monetized cost of time is then added to the total money spent on tax preparation and filing to arrive at a monetary estimate for the total objective costs of tax

^{32.} See Evans, supra note 29, at 457 (describing possible ways to measure compliance costs including "in absolute money terms or relative to tax yield, GDP or administrative costs"—all of which refer to monetary values).

^{33.} Id. at 451 (referring to these as the "hard core" of compliance costs).

^{34.} See, e.g., Off. of Mgmt. & Budget, *supra* note 31 ("Monetizing burden hours would present a daunting methodological challenge and raises issues concerning certainty and ease of administration by agencies."); Samaha, *supra* note 19, at 319–36 (challenging the government's prevailing hourly-wage method of monetizing time).

^{35.} E.g., York, supra note 23, at 3-6 (describing different hourly rates to monetize time costs and using Bureau of Labor Statistics averages for full-time private sector workers (\$37.28) and professional workers (\$52.05) to calculate time-cost savings from individual and alternative minimum tax reforms, respectively); Samaha, supra note 19, at 298 (noting that the Institute of Museum and Library Services "used the national average per capita income of about \$20 per hour to convert respondent time into dollar cost"); DEP'T OF HEALTH & HUM. SERVS., FOOD & DRUG ADMIN., TOBACCO PRODUCT STANDARD FOR N-NITRO-SONORNICOTINE LEVEL IN FINISHED SMOKELESS TOBACCO PRODUCTS 78 (2017), https://www.fda.gov/media/102728/download [https://perma.cc/ZU3R-MNRR] ("Labor hours are valued at the current market wage as reported by the May 2015 Occupational Employment Statistics published by the Bureau of Labor Statistics"). But see Daniel J. Phaneuf, Can Consumption of Convenience Products Reveal the Opportunity Cost of Time?, 113 ECON. LETTERS 92, 94 (2011) (arguing for an alternative measure of time that uses consumers' choices for time-saving products rather than the market wage rate to infer individuals' "shadow value of time").

compliance.³⁶ Current estimates suggest that these objective costs are huge, annually totaling hundreds of billions of dollars.³⁷

We consider the actual money spent and the monetized cost of time spent, taken together, the "objective costs" of tax compliance. These costs are "objective" because they are measured according to some external value, like market wages or product prices, rather than personal values, like utility or willingness to pay.³⁸

B. WHAT WE DON'T KNOW: SUBJECTIVE COMPLIANCE COSTS

Although there is a large literature on the compliance costs of taxation, the research typically ignores the personal, idiosyncratic costs borne by individual taxpayers—what we call the "subjective costs" of tax compliance.³⁹ These costs are subjective because they vary taxpayer-to-taxpayer, depending on each person's attitude toward taxes, as well as general literacy, numeracy, recordkeeping habits, and so forth. Where the research does address subjective costs (usually framed as the psychological costs of taxation), it often does so via qualitative data that are

37. Hodge, supra note 3 (estimating total compliance costs at \$313 billion in 2022); Fichtner & Feldman, supra note 3, at 5 (summarizing empirical estimates of compliance costs at between \$67 billion and \$378 billion).

^{36.} See, e.g., Arthur B. Laffer et al., *The Economic Burden Caused by Tax Code Complexity*, LAFFER CTR. 20 (Apr. 2011), https://www.jec.senate.gov/public/ cache/files/12edbdc6-cac6-41be-bc0e-2b4338cf09ae/2011-laffer

⁻taxcodecomplexity.pdf [https://perma.cc/86UN-NSYS] (monetizing individual taxpayer time at \$68.42 per hour and business taxpayer time at \$55 per hour); J. Scott Moody et al., *The Rising Cost of Complying with the Federal Income Tax*, TAX FOUND. 9 (2005), https://files.taxfoundation.org/legacy/docs/sr138.pdf [https://perma.cc/QSF9-V4F9] (monetizing individual taxpayer time at \$39.18 per hour and business taxpayer time at \$47.96 per hour); *Simplification of the Tax System: Hearing Before the Subcomm. on Oversight of the H. Comm. on Ways & Means*, 108th Cong. (2004) (written statement of Joel Slemrod, Univ. of Mich.) (monetizing taxpayer time at \$20 per hour).

^{38.} Note that neither objective nor subjective measures of compliance costs take into account externalities from tax compliance. Tax compliance might incur negative externalities if, for example, time spent on taxes takes one away from time with family or prosocial work that would provide benefits to individuals other than the taxpayer. If we believe that tax compliance incurs mostly negative externalities, as opposed to positive externalities, all measures of compliance costs (including our own) may systematically understate the harm of tax compliance burdens.

^{39.} See Bankman, supra note 3, at 2 (noting this gap in the research).

difficult to compare with the quantitative compliance cost measures just described.⁴⁰

Many experts believe that conventional calculations of objective costs underestimate the true cost of tax compliance because they exclude important *psychological* costs of taxation.⁴¹ These scholars argue that tax compliance is uniquely unpleasant along two dimensions. First, tax compliance causes *aggravation* because it's tedious, difficult, and stressful.⁴² Second, tax compliance causes *anxiety* because taxpayers fear repercussions from making a mistake.⁴³ Researchers agree that these psychological costs exist and are significant.⁴⁴ Although we focus on aggravation and anxiety, subjective tax compliance costs surely include other experiences as well. For instance, some taxpayers might feel degradation or anger from being forced to complete an unpleasant task, apart from the unpleasantness of the task itself.⁴⁵

42. See Robin Woellner et al., *Taxation or Vexation — Measuring the Psychological Costs of Tax Compliance, in* TAXATION COMPLIANCE COSTS: A FEST-SCHRIFT FOR CEDRIC SANDFORD 35, 37 (Chris Evans et al. eds., 2001) (defining psychological costs to include "anxiety and frustration caused by complying with complicated revenue legislation"); Moynihan et al., *supra* note 41, at 46 tbl.1.

43. See *id.*; Evans, *supra* note 29, at 451 ("Taxpayers suffer stress, anxiety, and frustration as a result of attempting to comply with their tax obligations."). A related but distinct literature explores the psychological determinants of evading tax laws. See generally KEN DEVOS, FACTORS INFLUENCING INDIVIDUAL TAXPAYER COMPLIANCE BEHAVIOUR 13–66 (2014) (surveying the literature on taxpayer non-compliance behavior, focusing on theories from economics, psychology, and sociology).

44. See, e.g., Bankman, supra note 3, at 2 (acknowledging the existence of taxpayer anxiety that most cost estimates do not include); Binh Tran-Nam & John Glover, Estimating the Transitional Compliance Costs of the GST in Australia: A Case Study Approach, 17 AUSTRALIAN TAX F. 499, 518 (2002) (confirming "convincingly" that "psychological costs were indeed considerable to small business during the transitional period of tax reform").

45. Relatedly, some might feel negative (or positive) emotions from paying the taxes they owe. Although these emotions are separate from those arising from preparing a tax return, the two feelings are surely related. For instance, if

^{40.} See infra Part I.B.1.

^{41.} Donald Moynihan et al., Administrative Burden: Learning, Psychological, and Compliance Costs in Citizen-State Interactions, 25 J. PUB. ADMIN. RSCH. & THEORY 43, 46 tbl.1 (2015) (identifying psychological costs as stemming from "participating in an unpopular program, as well as the loss of autonomy and increase in stress arising from program processes"); Cass R. Sunstein, *Sludge and Ordeals*, 68 DUKE L.J. 1843, 1853 (2019) (arguing that administrative compliance burdens, so-called "sludge," impose psychological costs in the form of "frustration, stigma, and humiliation").

To illustrate the difference between subjective and objective costs, imagine two employees with identical jobs and salaries attempting to file their annual tax returns. One employee, Alex, enjoys paperwork and doesn't mind filing taxes. She enters her information as instructed, checks the necessary boxes, and signs the form without further thought. A second employee, Bernie, hates paperwork. Bernie completes her returns reluctantly, finding the process tedious and unpleasant. She does her best, but later worries that she made a mistake on her forms. Bernie has experienced some meaningful amount of unpleasantness from filing her taxes. If Alex and Bernie spend identical amounts of time filing their taxes and face the same market wage rate, their objective compliance costs will be identical. However, because Bernie disliked the process of completing taxes so much more, her *subjective* costs are much higher than Alex's. Measuring subjective costs thus offers an alternative way to assess and understand tax compliance costs, one based on a taxpayer's personal experience rather than her market wage.

Researchers typically ignore subjective costs in compliance costs estimates because they are difficult to measure and compare across individuals.⁴⁶ As a result, while we know a fair amount about objective tax compliance costs, we know far less about subjective tax compliance costs, including their magnitude, distribution, and whether they increase or decrease at the margin.

someone is angry about having to pay taxes, they may feel more aggravated when preparing their tax return. Regarding positive emotions and tax payments, see Listokin & Schizer, supra note 4, at 185–86.

^{46.} See, e.g., Binh Tran-Nam et al., Personal Taxpayer Compliance Costs: Recent Evidence from Australia, 29 AUSTRALIAN TAX F. 137, 142–43 (2014) (noting that it is difficult to add monetary and psychological costs together, so most studies exclude psychological costs); John L. Guyton et al., Estimating the Compliance Costs of the U.S. Individual Income Tax, 56 NAT'L TAX J. 673, 675 (2003) (identifying, but not measuring, psychological costs of taxation, defined to include the "dissatisfaction, frustration, and anxiety" taxpayers feel due to "interaction with the tax system"); Youssef Benzarti, How Taxing Is Tax Filing? Using Revealed Preferences to Estimate Compliance Costs, 12 AM. ECON. J.: ECON. POL'Y 38, 38 (2020) (inferring compliance costs from the decision to itemize deductions, but only studying compliance costs in this context and not attempting to infer subjective costs).

1. Empirical Research on Subjective Costs

Much of the compliance costs literature simply assumes that subjective costs impose large burdens on taxpayers.⁴⁷ This assumption is partly based on the intuition that complexity generates frustration and the fact that inputs to the tax filing process are inordinately complicated.⁴⁸

There is some limited qualitative data about the subjective costs of taxation. For instance, public opinion polling unsurprisingly confirms that many are averse to tax compliance work, suggesting that some people experience high subjective costs of taxation. In national public opinion polls from 1990–2013, approximately half to two-thirds of respondents reported that they "dislike" or "hate" doing their taxes.⁴⁹ In 2013, just under

^{47.} Eighteenth-century economist Adam Smith noted the likely difference between subjective and objective costs. ADAM SMITH, AN INQUIRY INTO THE NA-TURE AND CAUSES OF THE WEALTH OF NATIONS 678 (Elec. Classics Series Publ'n 2005) (1776) ("[B]y subjecting the people to the frequent visits and the odious examination of the tax-gatherers, it may expose them to much unnecessary trouble, vexation, and oppression").

^{48.} For instance, the rules for claiming a qualifying child for purposes of the EITC, Child Tax Credit, Head of Household filing status, and Child and Dependent Care Tax Credit all involve slightly different qualifying characteristics. See Qualifying Child and Qualifying Relative Flow Chart, AARP FOUND. TAXAIDE PROGRAM, (Nov. 14, 2011), http://nytaxaide.org/wp-content/uploads/2013/06/qualifyingchildflowchart11-11-14.pdf [https://perma.cc/5R76-JJKQ]. To know which benefits might apply, a claimant must work through separate (but confusingly similar) rules for each program and in some cases compare their situation to those of other people in the child's household. See *id.*; I.R.C. § 152(c)(4)(C) ("If the parents of an individual may claim such individual as a qualifying child but no parent so claims the individual, such individual may be claimed as the qualifying child of another taxpayer but only if the adjusted gross income of such taxpayer is higher than the highest adjusted gross income of any parent of the individual.").

^{49.} Public Opinion on Taxes, AM. ENTER. INST. 79 (Apr. 10, 2009), https://www.aei.org/wp-content/uploads/2011/10/AEI-Public-Opinion-Studies -Taxes-2009.pdf?x91208 [https://perma.cc/M8DS-GUFH] (compiling historical public opinion polls regarding feelings toward doing taxes); A Third of Americans Say They Like Doing Their Income Taxes, PEW RSCH. CTR. (Apr. 11, 2013), https://www.pewresearch.org/politics/2013/04/11/a-third-of-americans-say-they -like-doing-their-income-taxes [https://perma.cc/S3JX-D2YJ] ("[A] majority of Americans (56%) have a negative reaction to doing their income taxes, with 26% saying they hate doing them.").

one-third of respondents stated that they dislike tax filing because it is "complicated" and involves "too much paperwork."⁵⁰

Despite scholarly awareness of the subjective costs of taxation, in-depth empirical research on the topic is vanishingly scant and almost entirely based outside of the United States.⁵¹ Apart from confirming the existence of some unknown amount of subjective costs, these studies fail to arrive at a consensus on the nature or magnitude of such costs. Many of the results affirm common sense.⁵² For instance, one study from Australia found that the stress of tax compliance is higher in the wake of major tax policy change and that "worry" is a major cause of such stress.⁵³ Another reported that survey participants exhibited physical signs of psychological costs, including "biting lips and wringing hands," while solving a hypothetical tax problem.⁵⁴ Researchers in Spain found that keeping financial records is among the "most disappointing aspects of tax compliance."⁵⁵

These studies provide useful confirmation of the existence of heightened subjective costs for at least certain taxpayers. However, past research is of limited general use for several reasons. For one, much of the research focuses on Value-Added Taxes (VAT) imposed on business taxpayers.⁵⁶ The United States has no VAT, and findings about a VAT may not be generalizable to

^{50.} See A Third of Americans Say They Like Doing Their Income Taxes, supra note 49. In addition, a portion of this aversion can be attributed to the tax payment itself. See *id.* (reporting certain taxpayers dislike doing their taxes because they "owe the [government] money," receive "no refund," or feel they "pay too much").

^{51.} See Evans, *supra* note 29, at 451 (stating, as of 2008, that "no studies have yet managed to successfully quantify these psychological costs"); Woellner et al., *supra* note 42, at 35 ("[T]he issue of psychological costs has been a largely neglected area of tax compliance costs work.").

^{52.} See, e.g., John Hasseldine & Ann Hansford, The Compliance Burden of VAT: Further Evidence from the U.K., 17 AUSTRALIAN TAX F. 369, 380, 383 (2002) (finding that taxpayers who report higher compliance costs are more likely to report that there is some amount of psychological cost—defined as "stress/anxiety/sleepless nights"—associated with tax compliance).

^{53.} Tran-Nam & Glover, supra note 44, at 519-20.

^{54.} Woellner et al., *supra* note 42, at 44.

^{55.} M. Luisa Delgado Lobo et al., *Hidden Tax Burden of the Personal Income Tax: Evidence from the Recent Tax Reform in Spain*, 16 AUSTRALIAN. TAX F. 463, 475 (2001).

^{56.} *E.g.*, Hasseldine & Hansford, *supra* note 52 (evaluating the compliance costs of VAT in the U.K.); Tran-Nam & Glover, *supra* note 44 (examining the transitional compliance costs of implementing a Goods and Services Tax).

other types of taxes. For instance, a VAT may be more or less complicated than other types of taxes. Additionally, nearly all of these studies gather only *qualitative* data on subjective costs.⁵⁷ While qualitative research can provide valuable insights into taxpayers' perceptions about the tax filing process, it can't reveal certain insights about the general nature of subjective costs. For instance, although someone may report feeling stressed or anxious about tax filing, it's difficult to know how troublesome such stress or anxiety was to the person. Was it debilitating, or merely a minor inconvenience? When assessing the value of policy proposals, information about the magnitude of costs is particularly important.⁵⁸

2. Potential Subjective Benefits of Taxation

On the other hand, the tax filing process could conceivably generate subjective benefits that offset subjective costs.⁵⁹ For instance, some people might feel satisfaction from filing a tax return, similar to the satisfaction of completing a necessary chore. Some taxpayers may also enjoy the process of reviewing and

^{57.} One exception on this front is an interesting study from Bangladesh that attempts to monetize psychological costs associated with VAT compliance based on the "average annual cost per taxpayer of sleeping pills, tobacco, consulting psychologists or psychiatrists or similar medication used to relieve the symptoms of anxiety or stress connected with such compliance." Nahida Faridy et al., *The Devil Is in the Detail: An Analysis of VAT Compliance Costs for SMEs in a Developing Nation*, 23 N.Z. J. TAX'N L. & POL'Y 176, 186 (2017).

^{58.} Regarding the magnitude of costs, certain tax policy puzzles might suggest that some people experience above-average subjective costs. In particular, high subjective costs might partly explain why some people fail to claim refundable tax credits to which they are entitled. For some taxpayers, the fear of making a mistake and the stress of engaging with the government might be large enough to drive them not to file a return that claims refundable tax credits. *See Why Don't Americans Claim Their Earned Income Tax Credit?*, 90.5 WESA (Jan. 28, 2016), https://www.wesa.fm/archives/2016-01-28/why-dont-americans-claim -their-earned-income-tax-credit [https://perma.cc/9KRN-QDJN] (explaining that some fail to claim the EITC because they fear making a mistake and being audited).

^{59.} The psychological benefits of tax compliance should be distinguished from the psychological benefits of tax payment. For instance, some research finds that people experience pride from contributing to shared coffers. *See* VANESSA S. WILLIAMSON, READ MY LIPS: WHY AMERICANS ARE PROUD TO PAY TAXES 32–38 (2017) (describing interviews in which respondents expressed "commitment to the civic and moral responsibility of taxpaying"); Listokin & Schizer, *supra* note 4, at 185–88 (exploring literature about pro-social behavior in the context of tax payments).

organizing financial records or may find this organization useful for business activities unrelated to tax planning. Others might find tax law interesting and may find the learning process to be edifying. Still others might look forward to filing if they expect to receive a tax refund. These subjective benefits may offset aggravation or anxiety for individual taxpayers and among the population as a whole. Public opinion polling since the 1990s reveals that between 18–34% of respondents "like" or "love" doing their taxes.⁶⁰ When asked why, people report that they are "good at it" and that the process provides them a "good overview of personal finances."⁶¹

There may also be society-wide benefits from tax filing that outweigh the unpleasantness that individuals experience. Lawrence Zelenak has argued that the process of filing an income tax return promotes tax consciousness and fiscal citizenship.⁶² This is in part because filing a tax return makes people aware of their contribution to shared social goods.⁶³

C. WHAT WE SHOULD KNOW: MONETIZING SUBJECTIVE COSTS

Aside from the fact that subjective costs of tax compliance exist, and that they might differ across individuals—for instance, by education or income—the nature and scope of these costs are poorly understood. Moreover, because nearly all research on subjective costs is qualitative, it's difficult to know how such costs compare with objective costs estimates. This Section briefly explains why we need a better understanding of subjective costs as well as why data about the monetary value of such costs would be useful.

For one, it's difficult to know which parts of tax compliance taxpayers find to be most unpleasant (or pleasant). Without this information, reformers are left guessing where to invest

^{60.} Public Opinion on Taxes, supra note 49, at 79; A Third of Americans Say They Like Doing Their Income Taxes, supra note 49.

^{61.} A Third of Americans Say They Like Doing Their Income Taxes, supra note 49.

^{62.} LAWRENCE ZELENAK, LEARNING TO LOVE FORM 1040, at 4 (2013) ("[T]ax returns have a crucial role to play in the promotion of what might be fiscal citizenship.").

^{63.} *Id.* at 111–15. Professor Zelenak also argues that simplifying the filing process—thereby making it less unpleasant—would enhance this positive consequence of mass tax return filing.

resources to simplify and streamline the process.⁶⁴ If one goal of tax simplification is to reduce taxpayers' actual disutility from tax compliance, information about the distribution of disutility is important.

While researchers could simply ask taxpayers to ordinally rank preferences, a ranking survey would fail to account for the intensity of preferences.⁶⁵ A respondent may mildly dislike task a, slightly more strongly dislike task b, but absolutely detest task c, and a survey based on ranking couldn't register the difference between a-b and b-c.

Additionally, it's difficult to know to what extent psychological *benefits* of taxation might offset, or even exceed, psychological costs. Surveys might ask taxpayers whether they "like" or "dislike" filing taxes overall, which suggests whether benefits exceed costs for individual taxpayers.⁶⁶ But we can't compare these relative costs and benefits between individuals. Does the magnitude of someone's "like" exceed someone else's "dislike"?

Relatedly, while detailed subjective cost surveys often ask taxpayers to assess the intensity of stress or worry,⁶⁷ these studies often fail to ask whether there are benefits that offset the costs (such as satisfaction or finding the process to be enjoyable).⁶⁸ Thus, even with detailed information about specific elements of subjective costs, it's difficult to know how costs and benefits offset each other for any one individual and across

^{64.} This is not to say that reformers have no idea how to effectively simplify the tax system. Many reformers are extremely knowledgeable about how the tax filing process works and about which parts cause most trouble for taxpayers. Even so, such a perspective might be skewed relative to the average taxpayer's perception of the tax filing process. For instance, tax experts might tend to see cases that are more complex than average. Tax experts might also focus on provisions that cause tax professionals a great deal of trouble but have little effect on the average taxpayer—for instance, complicated international tax or business tax provisions.

^{65.} Additionally, taxpayers may have a difficult time assessing which tasks are most pleasant/unpleasant when asked directly. How does one assess the relative tedium of gathering tax forms from employers versus entering information into TurboTax?

^{66.} See, e.g., A Third of Americans Say They Like Doing Their Income Taxes, supra note 49 (summarizing a national survey of taxpayers).

^{67.} *See, e.g.*, Tran-Nam & Glover, *supra* note 44, at 518–20 (having participants assess their level of stress as a response to tax reform on a scale from one to five).

^{68.} E.g., id. at 517–20 (failing to consider any positive feeling toward the newly implemented tax).

individuals. A taxpayer might find the process of gathering tax information to be unpleasant but might simultaneously enjoy reviewing their finances. Even with this information, we do not know if this taxpayer experiences net subjective costs of taxation or whether the experience is a net benefit to them.

Another challenge with qualitative surveys is the opportunity for "cheap talk"⁶⁹ or strategic behavior. Especially because aversion to tax compliance is a well-known cultural trope, survey respondents may reflexively state a strong dislike of filing taxes while giving little thought to the question. Attaching dollar amounts and asking about willingness to pay forces respondents to be more reflective in considering actual tradeoffs between time, psychological stress, and money.

Qualitative data also tell us little about the magnitude of subjective costs for any given individual. For instance, someone may indicate in a qualitative survey that filing taxes causes them extreme stress and worry. That same person might also indicate that they are not willing to pay more than \$20 to reduce the time or risk associated with filing a tax return. Even if this person truthfully and reflectively perceives themself to experience high subjective costs of taxation, they would not allocate much of their own resources to reduce those costs. This willingness-to-pay information provides a useful metric to evaluate the magnitude of subjective costs, as well as to compare costs across individuals.

The stakes are high. Many provisions that increase complexity also increase efficiency or improve distributive justice.⁷⁰ The Earned Income Tax Credit (EITC), the Child and Dependent

^{69.} See Vincent Crawford, A Survey of Experiments on Communication via Cheap Talk, 78 J. ECON. THEORY 286, 286 (1998) (describing experiments in which "talk is cheap," meaning that respondents' answers have no effect on their lives). See also *infra* Part II.C.2 for further discussion of the "cheap talk" phenomenon.

^{70.} See Emily Cauble, Superficial Proxies for Simplicity in Tax Law, 53 U. RICH. L. REV. 329, 333 (2019) ("Pursuit of the goal of simplicity . . . can sometimes sacrifice other goals"); Samuel A. Donaldson, The Easy Case Against Tax Simplification, 22 VA. TAX REV. 645, 650–53 (2003) (presenting tax complexity as a guard against "unfair, inefficient laws"); William G. Gale, Tax Simplification: Issues and Options, 92 TAX NOTES 1463, 1463 (2001) ("[Simpler taxes] . . . reduce the ability of policy makers to achieve other goals of tax policy."). But see Edward J. McCaffery, The Holy Grail of Tax Simplification, 1990 WIS. L. REV. 1267, 1284–91 (challenging the equity-simplicity and efficiencysimplicity trade-offs).

Care Tax Credit, and the medical expense deduction all increase the complexity of a tax return.⁷¹ But the presence of these tax benefits also increases the distributional fairness of the tax system.⁷² Other provisions increase complexity to increase efficiency, like the preferential rate for capital gains.⁷³ Ascribing a value to subjective costs allows us to weigh such costs against these other important values.

Finally, cost-benefit analysis has been the bedrock of regulatory design and reform for several decades.⁷⁴ And within the literature on cost-benefit analysis, willingness to pay is the coin of the realm. It's frequently used in federal regulatory estimates, including the valuation of human life.⁷⁵ Monetizing subjective costs based on willingness to pay therefore allows us to engage with agency cost-benefit analysis. It also allows us to compare our results to the objective costs estimates that agencies and researchers most commonly provide. Gathering qualitative data about tax compliance is useful, but it can't speak directly to costbenefit analysis that relies on monetary values.

II. SURVEY AND RESULTS

To estimate the subjective costs of tax compliance, we use a discrete choice survey experiment to measure taxpayers' willingness to pay to reduce or eliminate the burdens (or benefits) of tax compliance activities. This Part describes our survey

^{71.} I.R.C. §§ 21, 32, 213.

^{72.} E.g., Margot L. Crandall-Hollick & Joseph S. Hughes, *The Earned Income Tax Credit (EITC): An Economic Analysis*, CONG. RSCH. SERV. 13–23 (Aug. 13, 2018), https://crsreports.congress.gov/product/pdf/R/R44057 [https://perma.cc/7H3F-69KT] (discussing effects of the EITC on poverty, health, education, and distribution of tax burdens).

^{73.} *E.g.*, Noël B. Cunningham & Deborah H. Schenk, *The Case for a Capital Gains Preference*, 48 TAX L. REV. 319, 350–53 (1993) (discussing how a preferential rate for capital gains may ameliorate inefficient "lock-in" by encouraging sales of capital assets); *id.* at 358 ("[T]here seems to be almost universal agreement that the capital gains rules account for a significant portion of the Code's complexity").

^{74.} See, e.g., CASS R. SUNSTEIN, THE COST-BENEFIT REVOLUTION (2018) (describing the rise and modern ubiquity of cost-benefit analysis); Jonathan S. Masur & Eric A. Posner, Cost-Benefit Analysis and the Judicial Role, 85 U. CHI. L. REV. 935 (2018) (describing judicial interventions in cost-benefit analysis).

^{75.} For seminal work on the subject, see Daniel A. Graham, *Cost-Benefit* Analysis Under Uncertainty, 71 AM. ECON. REV. 715 (1981).

methodology and results. The Appendix provides additional details about both.

A. Methodology

1. Discrete Choice Surveys Generally

Environmental and health economists commonly use discrete choice experiments to solicit valuations of public goods for which there is no market.⁷⁶ Discrete choice experiments ask survey respondents to choose between hypothetical services with varying attributes, one of which is typically price.⁷⁷ By presenting respondents with a selection of options, including the option *not* to purchase (called the status quo option), the experiments mimic decisions that people make when facing real-life market choices. By varying the attributes of the services offered and then seeing how the variation affects respondents' likelihood of purchasing a service, we can determine how they value each attribute.

To use a simple example, say that we asked respondents to choose whether they would pay for a single attribute—for example, saving x hours on tax compliance at y price. We could vary x, the number of hours saved, and see how that affects the amount a given respondent would be willing to pay for the service. This would reveal how much respondents value that particular attribute. We could also vary the price to see how that affects respondents' willingness to pay. This would reveal how much respondents value money.

Rather than explicitly asking respondents how much they would pay to save x hours (the "explicit price"), the discrete choice survey allows us to calculate what's known as the "implicit price" of each attribute. By using a regression model (described at length in Section C of the Appendix), we can calculate

^{76.} See generally Robert J. Johnston et al., Contemporary Guidance for Stated Preference Studies, 4 J. ASS'N ENV'T & RES. ECONOMISTS 319 (2017) (providing best-practices recommendations for stated-preference studies, including discrete choice experiments); Giles Atkinson et al., Cost-Benefit Analysis and the Environment: Further Developments and Policy Use, ORG. FOR ECON. COOP. & DEV. 130–31 (2018) https://doi.org/10.1787/9789264085169-en (select "Discrete choice experiments" in "Methods of environmental valuation" section of drop-down menu) (describing various trade-offs in the design of discrete choice experiments).

^{77.} Johnston et al., supra note 76, at 320.

what implicit values attached to each service best explain the respondents' choices of which services to buy. For example, imagine that respondents are offered two nearly identical services, one of which includes risk elimination and the other of which does not. Imagine further that respondents select the service without risk elimination if and only if it is at least \$50 cheaper than the service that includes risk elimination. This behavior would imply that the value of risk elimination in this context is \$50. The discrete choice experiment thus allows us to infer how much respondents value particular tax services in monetary terms, even though we never directly ask respondents how much they value these services.

While our discrete choice experiment included various refinements to optimize statistical power,⁷⁸ it essentially followed the logic above. We varied attributes and prices in order to elicit respondents' implicit prices for attributes related to saving time on tax compliance and reducing the risk of errors.

This method has several important advantages. Because the attributes and prices were varied randomly in an experimental setting, we didn't need to worry about confounding respondent characteristics. We could study the *average* change in willingness to pay depending on the attributes offered, which we calculated regardless of the specific motivations for willingness to pay. (We separately asked respondents to explain their willingness to pay, and discuss those results in Subsection II.B.2.) If desired, we could see how specific respondent characteristics influenced willingness to pay by explicitly adding them to the regression model. This would allow us to dig deeper into the determinants of willingness to pay while still excluding unobservable confounding characteristics.

An obvious alternative to discrete choice experiments would be simply to ask respondents for their dollar willingness to pay to avoid tax compliance obligations. However, surveys of this type encounter many difficulties, including respondents' tendency to "provide either unrealistically high or zero [dollar] responses."⁷⁹ In contrast, discrete choice survey designs have been found to outperform surveys based on open-ended questions in predicting real-world behavior, potentially because they "lead to higher engagement, increase immersion, and reduce

^{78.} See infra Appendix Part A.4-5.

^{79.} Johnston et al., supra note 76, at 346.

satisficing."⁸⁰ Empirical scholars in subfields like environmental economics have moved toward discrete choice experiments and away from alternative methodologies in recent decades, largely because of these advantages.⁸¹

Discrete choice experiments also avoid certain forms of bias that commonly appear in surveys. Open-ended willingness-topay surveys may suffer from status quo bias, where respondents underestimate their willingness to pay for a service that departs from the status quo.⁸² Status quo bias may be caused by loss aversion or cognitive dissonance because respondents are attached to the basket of services that they currently have.⁸³ Discrete choice experiments allow us to control for status quo bias in the regression model, quantifying the extent of this bias and excluding it from the calculation of implicit prices.⁸⁴

Another concern in surveys is the presence of demand characteristics, where survey respondents infer the surveyor's desired outcome and shape their answers to meet that outcome.⁸⁵ In an open-ended willingness-to-pay survey, respondents might infer that surveyors desire high or low responses and respond accordingly, masking their true preferences. In contrast, because discrete choices are complex and the statistical implications of any particular choice are relatively unclear, the experimenter's desired response is also unclear.⁸⁶ Because respondents can't adapt their behavior based on the surveyor, demand

^{80.} Stefanie Stantcheva, *How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible* 45 (Nat'l Bureau of Econ. Rsch., Working Paper No. 30527, 2022).

^{81.} See Johnston et al., *supra* note 76, at 346 (noting that the use of "openended questions has decreased in recent year relative to other [survey] formats").

^{82.} See Raymond S. Hartman et al., Status Quo Bias in the Measurement of Value of Service, 12 RES. & ENERGY 197, 197–98 (1990) (discussing status quo bias).

^{83.} See id. at 198 ("[A] survey [can become] contaminated by [a] respondent's current situation \dots ").

^{84.} See infra Appendix Part D.

^{85.} See Jim McCambridge et al., The Effects of Demand Characteristics on Research Participant Behaviours in Non-Laboratory Settings: A Systematic Review, 7 PLOS ONE 1, 1 (2012), https://doi.org/10.1371/journal.pone.0039116 (defining demand characteristics).

^{86.} Stantcheva, *supra* note 80, at 44 ("When several (sometimes many) dimensions vary, it may be harder for respondents to know what is being sought after.").

characteristics, as well as bias from subtle differences in survey framing, are less likely.

2. Survey Details and Implicit Prices

We surveyed respondents using Prolific, an online provider of surveys for academic research. Before designing the discrete choice survey, we conducted an initial pilot survey of 200 respondents. In the pilot survey, respondents answered simple, open-ended questions about their willingness to pay to eliminate various aspects of tax compliance burdens. For the reasons discussed above, we do not consider these responses reliable, but they provided rough intuitions about taxpayers' willingness to pay that we used to generate the various features and appropriate price levels in the discrete choice experiment.⁸⁷

After answering a series of questions about their employment and tax-filing situation, respondents were presented with several "choice sets." Each choice set contained two hypothetical tax services that differed along four key attributes. Based on existing literature as well as pilot survey responses, we hypothesized that people might in particular find it *aggravating* to spend time on taxes and that people might separately feel *anxious* about making mistakes on their returns or being audited.⁸⁸ Although taxpayers surely experience other emotions apart from aggravation and anxiety, discrete choice experiments cannot capture much detail about respondents' emotions. Qualitative questions helped provide additional nuance in this regard.⁸⁹

To measure aggravation and anxiety, we asked respondents whether they would be willing to purchase a tax service that would reduce or eliminate time spent on taxes, eliminate risk, or both. Time reduction was further broken down into four different levels: no time savings, one hour of time savings, half of their tax compliance time eliminated, or all of their tax compliance time eliminated.⁹⁰ Based on early-stage survey feedback, we also hypothesized that respondents might value a service differently depending on whether it was provided by a private third-party or

^{87.} It's interesting to note that the pilot results generally aligned with results in the full survey, in the sense that willingness to pay was consistently lower than post-tax hourly wage.

^{88.} See supra notes 41-44 and accompanying text.

^{89.} See infra Part II.B.2.

^{90.} See Appendix Part A.3.

1281

by the government. Finally, each service was offered at a different price.

In each choice set, respondents were asked whether they would purchase one of the two presented tax services, or whether they would decline both services. Figure 1, below, shows a representative choice set for a survey participant who currently spends ten hours a year on tax compliance:

Figure 1: Example of Choice Set Card

Which of the following tax services would you prefer?

	Service 1	Service 2	Neither Service
	Reduce by Half	Eliminate All	
Time Savings	(spend 5 hours	(spend no time or	
	instead of 10 hours)	effort on taxes)	
	No Reduction	Eliminate All	
Risk	(no change in	(taxes are 100%	No Change
Elimination	accuracy or audit	accurate, no risk of	(continue to spend
	risk)	future audit)	10 hours, no
Provider	Private Company	Government	change in accuracy or audit risk)
Fee	\$25	\$40	
	Service 1	Service 2	Neither
Your choice:	0	0	0

Unlike a survey simply asking respondents how much they would pay to eliminate various tax compliance burdens, a single choice in a discrete choice experiment reveals relatively little. In the example above, a respondent might choose Service 2 because they highly value time savings, because they highly value risk elimination, or because they would rather trust the government with sensitive tax information than a private party. But the beauty of the discrete choice experiment is that we vary the attribute levels in each choice set in a way that ultimately reveals

respondents' preferences in general for each attribute, including the implicit prices they attach to the attributes.

3. Survey Sample

We screened 1,000 respondents to restrict the survey to people who would have meaningful opinions about the tax filing process. In particular, we screened for respondents who have filed a tax return for the past three years. Additionally, because our survey asked respondents how much they would pay to eliminate one hour, half, or all of their tax compliance time, we limited our survey to those taxpayers who expect to spend at least three hours on tax compliance, to make these comparisons meaningful. 475 respondents qualified in the pre-screening and completed the survey.

Of the 475 people who completed the survey, we discarded 231 answers according to various preset criteria, either because they failed an attention or comprehension check or because their written comments indicated they rejected the premise of the survey. Section B of the Appendix discusses the portion of respondents who were excluded on these or any other grounds. This left 244 respondents for the full analysis, each of whom received seven discrete choice questions (excluding one question that served as an attention check).⁹¹

Table 1 provides brief descriptive statistics for our survey sample. 92

^{91.} Based on our analysis, provided *infra* Appendix Part A.5, our sample was approximately two-thirds larger than the minimum sample necessary to ensure reliable results.

^{92.} We discuss sample representativeness and external validity *infra* Part II.C.1. To compare with demographic statistics for the general U.S. population, see *QuickFacts*, U.S. CENSUS BUREAU, https://www.census.gov/quickfacts/fact/table/US/PST045221 [https://perma.cc/V8PB-JEP5].

m 11	1	a	1	D	1 •
Table	1.	Sami	nle	Demo	ranhics
I UDIO	-	Num	DIC.	Domoy	arapinos

Gender	Percentage of Sample		
Male	59.4%		
Female	40.3%		
Race	Percentage of Sample		
White	64.0%		
Black	13.7%		
Hispanic	12.6%		
Asian	5.5%		
Mixed-Rac	e 3.2%		

Education	Percentage of Sample
Highest Degree = High School	20.6%
Highest Degree = College	43.9%
Highest Degree = Master's	16.0%
Highest Degree = Doctorate	2.2%

Income	Median
Individual Income	\$45,625.04
Household Income	\$75,000

To contextualize our findings, we also collected data on the number of hours that respondents spent on tax compliance. The mean number of hours spent was 8.32 hours, and the median was 5 hours. As noted above, the minimum number of hours spent for a participant to be included in our study was 3 hours. The number of reported hours roughly followed an exponential distribution, with 3 hours being the most common number reported.

B. Results

1. Discrete Choice Experiment Results

Calculating willingness to pay revealed several interesting findings. First, people are more concerned with risk than with aggravation or tedium. That is, people are willing to pay more to eliminate the risk of making a mistake or being audited (\$72.17) than they are willing to pay to eliminate all the time they spend on tax compliance activities (\$53.35).

Second, we found that there is declining marginal cost of time spent on tax compliance. People are willing to pay essentially nothing to eliminate one hour of tax compliance time; they are willing to pay around \$2.44 per hour (about \$10 in total, on average) to eliminate half of their tax compliance time. Both figures are significantly lower than the \$6.40 per-hour amount they would pay to eliminate all tax compliance time. This suggests that the first hours of tax compliance are the most painful and costly, whereas the last hour is hardly worth paying to eliminate at all.

Third, respondents valued government provision of tax services near zero. In other words, respondents did not seem to care whether their chosen tax service was provided by the government or by a private third party.⁹³ In contrast, the attributes related to time savings and risk elimination did affect respondents' choices.

Figure 2 and Table 2 provide implicit prices and confidence intervals for our full-sample model.⁹⁴ Both provide results for all three non-price attributes in the choice sets: time savings, risk elimination, and government provision of the service.

^{93.} We discuss policy implications of this indifference infra Part III.B.3.

^{94.} Table 6 in the Appendix provides the regression results for our main effects equation. We do not provide the regression results in the main text because regression results on their own carry limited interpretive weight in a discrete choice experiment. Rather, the coefficients are used to calculate the implicit prices reported here. *See generally* Atkinson et al., *supra* note 76, at 131–33 (explaining the conceptual foundation for the calculation of implicit prices in discrete choice experiments).

SUBJECTIVE COSTS



Figure 2: Implicit Prices of Attributes

Table 2: Implicit Prices

Attribute	Implicit Price	95% CI
Risk Elimination	\$72.17	64.55 - 81.83
Time Elimination: 1	-\$8.79	-\$15.11\$1.80
Hour		
Time Elimination: Half	\$10.13	\$1.84 - \$19.10
Time Elimination: All	\$53.35	45.41 - 62.61
Government	\$1.85	-\$2.22 - \$6.18

a. Mistakes Versus Aggravation

Respondents were willing to pay more (\$72 on average) to reduce the risk of error or audit than they were willing to pay to eliminate all the time they spend on tax compliance (\$53 on average). This result suggests that taxpayers are more concerned about making mistakes in their tax filings than by any unpleasantness associated with the task itself, such as aggravation or tedium.

The average difference in willingness to pay for risk elimination versus time elimination was \$18.83, a statistically significant difference.⁹⁵ Another way to understand this result is that people's willingness to pay for risk elimination is, on average,

^{95.} The 95% confidence interval for this difference is \$11.01 to \$26.65.

35% greater than their willingness to pay to eliminate all time spent on tax compliance.

Anecdotally, this finding is consistent with the marketing that accompanies many tax filing products. A taxpayer visiting the TurboTax or H&R Block websites sees nothing about time savings or efficiency gains, but they see a great deal about TurboTax's "100% Accurate Calculations" or H&R Block's "100% Accuracy Guarantee."⁹⁶ Insofar as marketing copy is a proxy for revealed consumer preferences, commercial emphasis on accuracy, rather than time savings, supports our findings.

What might be driving this result? One possibility is that people find the anxiety and anticipation of future cost and hassle to be more unpleasant than the actual task itself. This sentiment would make sense because such anxiety is purely negative to the person experiencing it. That is, there are no psychological benefits that arise from anxiety or the anticipation of a stressful or costly event in the future. In contrast, the act of filing a tax return might entail some aggravation and tedium, but these negative feelings might be offset by some positive ones—for instance, a sense of control, getting an overview of finances, and so forth. Respondents' qualitative responses support this explanation and provide additional context for understanding the result.⁹⁷

It's also likely that the willingness to pay to reduce risk captures something else, in addition to anxiety and anticipation of future unpleasantness. We asked respondents how much they would pay for a service that would eliminate the risk of errors in filing their taxes and eliminate the possibility of future audits. Such a service would do more than simply remove the psychological discomfort associated with worrying about incorrect filing it would substantively improve the accuracy of the respondent's tax filing. For many taxpayers, such a result would be worth paying for, apart from the peace of mind it would bestow on the taxpayer. On the other hand, for those taxpayers who would prefer

^{96.} Introducing TurboTax's New Lifetime Guarantee on Tax Returns, IN-TUIT TURBOTAX: BLOG (Apr. 13, 2023), https://blog.turbotax.intuit.com/turbotax -news/turbotax-new-lifetime-guarantee-on-tax-returns-54673 [https://perma.cc/ YB3M-QACR]; We Guarantee Your Maximum Refund. And So Much More., H&R BLOCK, https://www.hrblock.com/guarantees [https://perma.cc/ZBA5 -B97W].

^{97.} See infra Part II.B.2.

to submit an inaccurate return, improved accuracy would reduce willingness to pay.

We chose to frame the discrete choice language around accuracy and risk rather than anxiety because improved accuracy and risk elimination would be primary outcomes of adopting a "return-free filing" system, a prominent type of tax simplification proposal.⁹⁸ In a return-free system, the government completes returns on taxpayers' behalf.⁹⁹ For example, California's ReadyReturn program pre-populated taxpayers' returns with information that the state government already had, significantly simplifying taxpayers' filing process.¹⁰⁰ Many advanced countries have return-free tax systems.¹⁰¹ The federal government could plausibly eliminate much of the risk of error for a wide swath of taxpayers by using such a system.

In addition to the fact that this framing helps us to draw more useful policy conclusions, we feared that respondents would reject the premise of a service that magically purported to change their mental state and lead to inaccurate survey responses.¹⁰² We prioritized offering a plausible service over hyperspecific framing, based on best-practices advice to this effect.¹⁰³

101. See What Other Countries Use Return-Free Filing?, TAX POL'Y CTR.: BRIEFING BOOK, https://www.taxpolicycenter.org/briefing-book/what-other -countries-use-return-free-filing (last updated May 2020), https://www.tax policycenter.org/briefing-book/what-other-countries-use-return-free-filing [https://perma.cc/4HGG-HU5S] ("At last count, 36 countries, including Germany, Japan, and the United Kingdom, permit return-free filing for some taxpayers."); cf. François Vaillancourt, Prefilled Personal Income Tax Returns: A Comparative Analysis of Australia, Belgium, California, Québec, and Spain, FRASER INST. (June 2011), https://www.fraserinstitute.org/sites/default/files/ prefilled-personal-income-tax-returns.pdf [https://perma.cc/MN9H-MWJC] (comparing various countries pre-populated return systems).

102. Even as framed, many respondents refused to believe that a service could eliminate the risk of error or audit. The following qualitative response reflects such a viewpoint: "I do not believe that paying money will do this so I think it's wasteful to pay money towards this effort."

103. Best-practices guidance for discrete choice experiments stresses the importance of offering respondents plausible choice scenarios. The choices offered

^{98.} See infra notes 170–79 and accompanying text (discussing return-free filing); Joseph Bankman et al., *Why Filing Taxes Isn't Easy*, POLITICO (July 18, 2018), https://www.politico.com/agenda/story/2018/07/18/tax-filing-congress-irs -000683 [https://perma.cc/SD54-T33A] (discussing the possibility of "pre-population" of tax return forms by the IRS).

^{99.} See Bankman et al., supra note 98.

^{100.} Id. (describing ReadyReturn).

It's impossible to tease out what portion of someone's willingness to pay to reduce risk reflects a desire to reduce anxiety alone versus a desire to reduce risk alone. Suffice to say that the implicit price for risk reduction reflects just that: risk reduction. While respondents' qualitative explanations¹⁰⁴ help to illuminate sources of aversion to mistakes and audits, we can't precisely circumscribe the factors contributing to that aversion based on our experiment.

One final factor might drive down average willingness to pay to reduce anxiety. Some taxpayers intentionally take positions on their tax returns that are not supported by tax law or by financial reality. These taxpayers would likely pay less, perhaps nothing, for a service that guarantees accuracy. Even so, they might experience anxiety-perhaps even above-average anxiety—from preparing and filing a tax return. Our survey ignores this type of anxiety. From a normative perspective, this omission might be a feature, rather than a bug, of our survey since policymakers need not care about making life easier for tax cheats. Perhaps more concerning, if the zero-value responses drive down average willingness to pay for the risk reduction service, a high proportion of such respondents might make anxiety appear lower than it is. However, it's worth noting that none of the qualitative answers reflected a desire to intentionally submit an inaccurate return. While this might reflect social desirability bias¹⁰⁵ (even though the responses were anonymous), it at least provides some evidence suggesting that a desire to evade taxes didn't skew survey responses.

b. Declining Marginal Cost of Tax Compliance Time

Our results also suggest that people have declining marginal cost of time spent on tax compliance. Respondents were

in a discrete choice experiment "must be described in a way that is understood and viewed as credible by respondents and that enables respondents to anticipate accurately the likely effects on their welfare." Johnston et al., *supra* note 76, at 326.

^{104.} See infra Part II.B.2.

^{105.} See Sebastien Lizin et al., The State of the Art of Discrete Choice Experiments in Food Research, 102 FOOD QUALITY & PREFERENCE 1, 7 (2022), https:// doi.org/10.1016/j.foodqual.2022.104678 (describing social desirability bias as respondents "behav[ing] to please the researcher, avoid embarrassment, or 'look good," which leads them to "misrepresent their true preferences and ... systematically misreport socially sensitive behavior or attitudes").

willing to pay more per hour to eliminate all tax compliance time compared to half of their tax compliance time, and more per hour to eliminate half compared to one hour. Respondents were essentially not willing to pay anything to eliminate one hour of tax compliance time.

In fact, the implicit price for one hour of time savings was slightly negative. A negative implicit price suggests that, all else equal, a respondent would be less likely to choose a service that saves one hour of time than a service that saves no time. While this result seems to defy common sense and rational thinking, we can offer a few thoughts.

First, the negative implicit price is only very slightly statistically significant at a 95% confidence level, and not significant at the 99% level. As such, we can't confidently conclude that the true implicit price of this service is negative, rather than zero. Second, respondents may have felt that an hour of time savings was so small relative to their overall tax compliance time that the suggestion offended them, causing them to reflexively opt against a choice with that attribute even though doing so did not reflect their true preferences. Third, respondents might have believed that one hour of time savings was so minimal that it wouldn't merit the effort required even to purchase such a service. That is, it might reflect some presumption of transaction costs that our model ignores.

Fourth and finally, this finding might reflect the "peak-end rule" from behavioral psychology, under which one's perception of an unpleasant experience is primarily determined by the maximum unpleasantness of the entire experience and the unpleasantness at the very end. In one famous study, participants either (1) held a hand in painfully cold water for a set amount of time or (2) held a hand in the painfully cold water for the same amount of time and *then* held the hand in slightly warmer (but still painfully cold) water. Participants remembered the latter experience as less painful because it was slightly less unpleasant at the very end, even though the sum total of pain seems strictly greater.¹⁰⁶

^{106.} Daniel Kahneman et al., When More Pain Is Preferred to Less: Adding a Better End, 4 PSYCH. SCI. 401 (1993) (summarizing the research study and discussing the "peak-and-end pattern"). This discussion relates to our observations regarding the magnitude of compliance costs as well. As we explain below, infra Part II.B.1.d, our survey results suggest that tax compliance costs may be

If the last marginal hour of taxpaying is the least painful, it might be that respondents are correctly optimizing their remembered utility by prolonging the least painful part rather than

[108:1255

While intriguing, this last explanation requires fairly strong assumptions about our participants' affective forecasting that may be debatable in practice. Ultimately, the true explanation is likely some combination of the above. However, we at least find no evidence of a significant *positive* willingness to pay to eliminate the last hour of tax compliance.

A finding of declining marginal cost of tax compliance time is an interesting behavioral result. Conventional economic theory suggests that unpleasant activities have an *increasing* marginal cost.¹⁰⁷ In other words, under standard assumptions, the last hour spent on unpleasant work should be the *most* unpleasant hour—intuitively, one hour spent cleaning gutters might be mildly enjoyable, but by the fifth hour it becomes pure torture. Our findings contradict this intuition.

Nineteenth-century economist Stanley Jevons had a more complex economic theory about the (dis)utility of work, which better aligns with our findings. Jevons hypothesized that the disutility of work is positive for the first hour—that is, getting started is hard.¹⁰⁸ Once begun, however, the disutility (unpleasantness) of working drops significantly, and in some cases may even be overtaken by utility ("an excess of satisfaction").¹⁰⁹ At some point, after too many hours of work, disutility will once again overtake utility.¹¹⁰

1290

eliminating it.

lower than much of the literature asserts. Declining marginal compliance costs, further contextualized by the "peak-end rule" from behavioral psychology, might partly explain this result. If the last hour of tax compliance is relatively painless, taxpayers might remember the entire process to be less unpleasant than it was in reality.

^{107.} See, e.g., Pak-wai Liu, Monitoring Cost, Disutility of Effort and the Forcing Employment Contract, 8 ECON. LETTERS 187, 188 (1981) (assuming increasing marginal disutility of effort).

^{108.} David A. Spencer, *Love's Labor's Lost? The Disutility of Work and Work Avoidance in the Economic Analysis of Labor Supply*, 61 REV. SOC. ECON. 235, 237 (2003) ("At the moment of commencing labor it is usually more irksome than when the mind and body are well bent to the work." (quoting STANLEY JEVONS, THE THEORY OF POLITICAL ECONOMY 191 (1871))).

^{109.} Id.

^{110.} Id. at 237-38.
It's possible that the slope of the disutility curve for tax compliance work follows the same pattern that Jevons proposed for labor, increasing at the very start, then decreasing for some amount of time before eventually increasing again.¹¹¹ Among our respondents, the median annual time spent on tax compliance work was five hours total. It seems likely that, at this moderate number, taxpayers are still on the decreasing portion of their disutility curves. At that point, each additional hour of work is *less* unpleasant, not more.

It's also possible that aggravation is front-loaded because of the nature of tax preparation work. Perhaps the first hour of tax compliance time is the most substantively unpleasant, and each successive hour is relatively less unpleasant. It might be that sitting down to do taxes in the first place is painful, but after that a taxpayer settles into a comfortable rhythm. Moreover, if we consider the types of tasks that might make up individual tax compliance work, this result isn't so counterintuitive. For instance, the first several hours of tax preparation might entail tracking down difficult-to-find tax documents and conducting research about tax inputs like filing status, credit entitlements, new tax benefits, and so forth. Such work might be frustrating or confusing. Subsequent hours, in contrast, merely require inputting numbers into tax software or sending information to a tax preparer. The last hour likely entails double-checking work that has already been done, whether the return was self-prepared or done by a professional.

Even if tax compliance work entails the same level of unpleasantness throughout, it's possible that people adapt to that unpleasantness over time, consistent with the theory of hedonic adaptation.¹¹² Even if someone dislikes working on their tax return, once they've begun, they accept their fate and make the best of it.

c. Indifference Between Government Versus a Private Provider

Our results reveal that taxpayers ascribe no value to whether tax simplification services are offered by the government or a private company. Put slightly differently, at all price

^{111.} See id. at 238 fig.1 (graphing the disutility curve).

^{112.} Ed Diener et al., *Beyond the Hedonic Treadmill: Revising the Adaptation Theory of Well-Being*, 61 AM. PSYCH. 305, 305 (2006) (explaining hedonic adaptation).

levels, our respondents were not more likely to select a service merely because it was offered by the government or by a private company. This feature was essentially valueless to respondents. This finding is one of the most reliable in our survey, in the sense that we estimate a tight confidence interval around zero effect. This seems to suggest that our survey respondents are indifferent between services being offered by a private company or by the government.

A few things are worth noting about this result. For one, people were indifferent between government and private provision of *equivalent* services. But public and private services might not be equivalent. Private services might offer higher levels of customer support or wraparound financial services. Public tax services might offer customers greater data privacy, since their information would not pass through an additional third party. Public services might also appear to be free—although, of course, they are not free; they are merely funded by taxpayer dollars rather than private fees. These pros and cons of public versus private options were not highlighted in our survey, and so we cannot assume that people considered them. Faced with real-world options, it's possible that people would consistently place a higher value on either private services or public services.

This finding is especially interesting because it goes against conventional wisdom that people mistrust government.¹¹³ If people mistrusted government, you would expect to see a premium placed on the private service. We did not see that. It's possible that reporting on government mistrust often reflects the exact kind of survey bias that discrete choice experiments are designed to ameliorate. That is, if you ask someone directly whether she trusts the government, she might say no. If you offer her a functioning government service that she otherwise wants, she might have no problem accepting the service. Our findings support this reasoning.

^{113.} See Would You Trust the IRS to Do Your Taxes?, supra note 16 ("A Pew Research poll last year found that two-in-ten Americans say they trust the government in Washington to do what is right just about always (2%) or most of the time (19%)."); see also INTERNAL REVENUE SERV., IRS REPORT TO CON-GRESS: INFLATION REDUCTION ACT § 10301(1)(B) IRS-RUN DIRECT E-FILE TAX RETURN SYSTEM 11–12 (2023) [hereinafter IRS REPORT: DIRECT E-FILE SYS-TEM], https://www.irs.gov/pub/irs-pdf/p5788.pdf [https://perma.cc/8VBB -HBRU] (discussing taxpayer trust in an IRS-run filing system).

d. Magnitude: How Our Results Differ from Other Estimates

Our survey results differ from other prevailing tax compliance cost estimates in two ways. First, on an individual level, people seem to value their tax compliance time at something less than the marginal wage rate that other estimates tend to use. Second, relatedly, our results complicate attempts to monetize the benefits of particular reforms by aggregating compliance costs across the population. In many cases, the value of simplification reforms may be lower than other estimates suggest. Although magnitude estimates are necessarily imprecise, our cost estimates are sufficiently different from other estimates to merit discussion.

Our study found that people are willing to pay *less* to reduce time spent on tax compliance than market-wage-based estimates would assume. The respondents in our sample reported an average post-tax hourly wage of \$26.60 but were only willing to pay \$6.40 per hour to eliminate all tax compliance time. This comparison is important because, as explained above, researchers use market wages to calculate the monetary costs of tax compliance. Our survey results suggest that people value their tax compliance time at a lower rate than the rate at which the labor market values their time. This in turn suggests that researchers are overestimating the true cost of tax compliance.

Several factors might be driving this result. First, it's possible that tax compliance work simply isn't all that unpleasant for many taxpayers. Many qualitative answers reflected this perspective, stating that taxes aren't that complicated, that tax preparation is simple, and so forth.¹¹⁴ Perhaps people find tax compliance to be more pleasant than working; or perhaps the offsetting benefits make tax compliance time less burdensome on net compared to working.¹¹⁵

Second, respondents' *marginal* wage likely differs from their *average* wage in ways that conventional calculations of objective costs fail to account for. Many people might correctly value their marginal cost of time at \$0. Perhaps they can't earn more income, either because they are on salary or because they can't obtain additional work hours. If so, their opportunity cost of

^{114.} For instance, one respondent stated, "I do not mind taking the time to do my taxes myself." For further detail about respondents' qualitative answers, see *infra* Part II.B.2.

^{115.} See supra Part I.B.2.

forgone work is \$0, in which case spending time on tax preparation saves them money and doesn't cost them anything in forgone work.

Third, responses may reflect mental accounting. For instance, some people may "bucket" their spending into predetermined categories, like food, rent, and entertainment. If respondents don't have an existing bucket for "reducing tax compliance time," they may be unwilling to add it.¹¹⁶ Even if they have an existing bucket for tax preparation expenses, they may be unwilling to expand that bucket for some hypothetical service. Without such a bucket, or with a static bucket, their willingness to pay to reduce tax compliance activities will be low, even if they experience some subjective costs.¹¹⁷ Relatedly, people's willingness to pay might anchor on existing prices for tax preparation services. If current prices serve as an anchor, results will converge around values that are similar to current tax preparation fees, even for people who find tax compliance to be extremely unpleasant.

Our findings also complicate attempts to monetize the benefits of particular reforms by aggregating compliance costs across the population. Simply multiplying compliance-time savings by an average hourly wage, as estimators often do, may overstate the total benefits. For reforms that only cut an hour or two of tax compliance time, such estimates completely overlook possible declining marginal costs of tax compliance time. For example, the Tax Foundation has estimated that the Tax Cuts and Jobs Act (TCJA) led to "compliance savings" worth \$5.4 billion

^{116.} In other mental accounting contexts, there is evidence that people do not account for time and money in the same way. See Dilip Soman, The Mental Accounting of Sunk Time Costs: Why Time Is Not Like Money, 14 J. BEHAV. DECISION MAKING 169, 182 (2001) (finding that the "sunk-cost effect"—the tendency to focus on past costs when making current decisions—is weaker for past time investments than for past money investments).

^{117.} Somewhat related to mental accounting, some research finds that people are simply more willing to spend time than money toward a given task. See Tore Ellingsen & Magnus Johannesson, Time Is Not Money, 72 J. ECON. BEHAV. & ORG. 96, 101 (2009) (concluding, based on their findings, that "subjects are generally more prone to make non-monetary sacrifices than to make equivalent monetary sacrifices"). But see Sanford E. DeVoe & Jeffrey Pfeffer, When Time Is Money: The Effect of Hourly Payment on the Evaluation of Time, 104 ORGAN-IZATIONAL BEHAV. & HUM. DECISION PROCESSES 1, 4 (2007) (finding that people who are paid hourly are more likely "to think about their time in the same way they thought about money").

by saving taxpayers one hour of tax compliance time.¹¹⁸ We estimate the benefit from saving an hour of time or less to be much lower—indeed, to be approximately zero.

In general, these findings suggest that marginal simplification efforts may not be worthwhile if they come at great administrative cost or impugn other values, like fairness or equity. For example, the Tax Foundation used its TCJA estimates to defend a piece of legislation that benefited high-income taxpayers at the expense of low- and middle-income taxpayers.¹¹⁹ Based on our study, we should view arguments like these with greater skepticism.

These figures should be taken as a kind of rough, middle-ofthe-road estimate of tax compliance costs. Because of certain limitations with our sample, they do not reflect all subgroups of the population. For instance, very high-income and very-low-income taxpayers are not included in our sample, or at least not in high enough proportions to offer convincing estimates of costs for these groups.¹²⁰ We discuss these external validity issues further in Subsection II.C.1.

2. Qualitative Context

To provide qualitative context for respondents' choices, we asked each respondent to provide reasons for their willingness or unwillingness to pay for time savings or risk reduction.¹²¹ Each respondent was randomly asked either about time savings or risk reduction, but not both. The resulting answers can be divided into four groups: (1) those who think it's worthwhile to pay money to reduce time spent on tax filing; (2) those who *do not*

^{118.} York & Muresianu, *supra* note 10, at 1, 8. Because this is simply the *average* predicted reduction in tax compliance time, the reduction might still be dramatic for some taxpayers and minimal for others in a way that would increase the estimated monetized benefit under our rubric. However, the Tax Foundation did not engage in this analysis, assuming a uniform benefit from time savings based on average wage rates.

^{119.} William Gale et al., *A Preliminary Assessment of the Tax Cuts and Jobs Act of 2017*, 71 NAT'L TAX J. 589, 589 (2018) (estimating that the TCJA would "make the distribution of after-tax income less equal").

^{120.} Moreover, we assume that non-filers incur no subjective costs of tax compliance. But this is not necessarily accurate—it could be that non-filers must exert some effort to verify their status as non-filers and that they would pay some amount to alleviate this burden.

^{121.} For more information about the qualitative questions, see infra Appendix Part A.6.

think it's worthwhile to pay money to reduce time spent on tax filing; (3) those who think it's worthwhile to pay money to reduce *risk* of error and audit; and (4) those who *do not* think it's worthwhile to pay money to reduce risk of error and audit. Figure 3 provides the proportion of yes/no responses to each initial question.



Figure 3: Responses to Initial Qualitative Questions

Both Authors read every qualitative answer and coded each answer according to a list of common answer features.¹²² Answers could be coded with multiple features.

a. Time Savings

Half of respondents were asked whether they felt it was worthwhile to pay money to reduce their time spent on tax compliance. Among those who answered yes to this question, responses can be divided into two common themes.¹²³ Most of these

^{122.} Both Authors initially coded the results separately after agreeing on categories but without discussing any specific responses. The initial round of coding produced an intercoder agreement rate of 88.5%. *Cf.* Young Ik Cho, *Intercoder Reliability, in* ENCYCLOPEDIA OF SURVEY RESEARCH METHODS 344, 345 (2008) ("[C]oefficients .90 or greater are considered highly reliable, and .80 or greater may be acceptable in most studies."). After the first round of coding, we discussed each response on which we disagreed and produced final codings.

^{123.} Among these respondents, 7% seemed to misunderstand the question, rejected the premise of the question, or provided answers that didn't clearly explain the respondent's choice. For instance, "I wish the government would do this automatically \ldots ."

answers (69%) reflected the view that respondents' time could be better spent in some way other than preparing their tax returns. Many of these answers specified that they could instead be working or spending time with family. Some answers stated something simple like, "My time is valuable," or described tax compliance work as "a waste of time."

The following two answers reflected this opportunity-cost viewpoint:

That time I could spend doing something fun like taking my dogs for a walk or something.

Because the time you spend on tax activities could be better spent on other things, including earning money.

A significant portion (37%) of respondents who were willing to pay money to reduce time spent on tax filing said that they would be willing to do so because they found tax compliance work unpleasant or stressful. For instance:

I tend to procrastinate every year because it feels like it takes forever, and it's really stressful doing it. I would pay some money just to take the stress and procrastination away.

Paying taxes is a pain in the butt.

Doing taxes is a stressor to me, and I'm stressed while preparing them and stressed before I do them like I feel like the task is kind of hanging over me . . .

Among those who stated they were *not* willing to pay money to reduce time spent on taxes, responses reflected more diverse viewpoints.¹²⁴ The most common answer in this group (56%) indicated that respondents wouldn't pay to reduce time spent on taxes because their tax return is simple, the work is easy, or they simply don't mind doing it. For instance:

^{124.} Among respondents who said they were not willing to pay money to reduce time spent on taxes, 10% seemed to misunderstand the question, rejected the premise of the question, or provided answers that didn't clearly explain the respondent's choice. For instance, "[I] should not have to pay any money to get my taxes done it is such a disgusting exploitation of citizens to make them pay money to pay taxes."

I do not mind taking the time to do my taxes myself.

My taxes are not that complicated.

The next most common answer category, comprising 22% of those who were unwilling to pay to reduce time spent on taxes, stated that they couldn't afford to do so or that it wasn't worth the cost. For instance:

My money is hard earned and I'd rather save it to feed my family.

16% of these respondents stated that they prefer to complete their tax return themselves in order to control their tax preparation, to better understand their finances, or some other nontax benefit. For instance:

Because I have a high need for control and I really need to know what is going on and if it is accurate.

The time I spend educating myself improves my net returns.

Finally, 16% of these respondents stated that they would not pay to reduce time spent on taxes because they would prefer to reduce their risk of error instead.¹²⁵ These responses were notable because we didn't explicitly prompt respondents to suggest an alternative service; rather, they likely reflect the overall framing of the discrete choice experiment, in which time savings and risk reduction were the two service attributes respondents were previously asked to consider before the qualitative questions. This explanation is consistent with our finding that respondents would generally pay more to eliminate risk than to reduce time spent on taxes. It's also worth noting that no respondents indicated the opposite, that they would rather spend money to save time than to eliminate risk.

b. Risk Reduction

Half of respondents were asked whether they felt it was worthwhile to pay money to reduce their risk of error or audit.

^{125.} The following is an example of this type of response: "To me, time spent doesn't matter as much as accuracy when it comes to filing taxes."

Those who answered yes to this question provided diverse reasons for their willingness to pay.¹²⁶ The three most common answer categories were each present in approximately 30% of these answers.

31% stated that they would be willing to pay to reduce risk in order to reduce current worry or to provide "peace of mind." For instance:

I feel it is worth paying for greater peace of mind.

It's worth it so you don't have to stress out wondering if you made an error or will get audited.

30% stated that they would pay to reduce risk in order to reduce the likelihood of owing additional taxes, penalties, or interest upon a potential future audit. A similar but slightly smaller portion, 28%, stated that they would be willing to pay in order to minimize future effort or hassle associated with an audit. In contrast to the first response category, these two categories reflect a concern about *future* expenses or hassle rather than current anxiety.¹²⁷ For instance:

Tax law is complex and errors can be very costly and time consuming.

No one wants to go through the hassle of being audited.

Because the audit in itself would trigger a lot of surplus activity and extra work—such as finding paperwork, re-calculating figures, and reviewing tax tables—that you wouldn't have had to do, if you'd just paid the fee to eliminate the risk.

^{126.} Among respondents who stated that they were willing to pay money to reduce risk, 4% seemed to misunderstand the question, rejected the premise of the question, or provided answers that didn't clearly explain the respondent's choice. For instance, "Trust a tax preparer more than any federal employee. Want to know why...look at how poorly they do work now."

^{127.} Many answers reflected concern about both current worry as well as future hassle. These answers were coded with both categories.

A smaller portion of responses in this group (7%) expressed a specific fear of legal repercussions or a desire to shift legal liability to a third party.¹²⁸ For instance:

I don't care for saving money or time as much as I care about staying out of trouble.

[If I paid for a risk reduction service] I would not be held accountable if something went wrong with my taxes.

Although many respondents said they would pay to reduce risk, some didn't provide a specific rationale for their willingness to pay. These answers instead expressed a general sentiment that the service was worth the money, or a general dislike of risk, errors, or audits. 23% of responses fell into this general bucket, defying more specific classification.

The final, and smallest, group of respondents stated that it was *not* worthwhile to pay money to reduce risk of error or audit. These answers reflected two viewpoints.¹²⁹ Most (63%) stated that their taxes are simple or that risk of error or audit is relatively low. For instance:

My taxes are not complicated. It's very easy for me to do the prep work and complete and file my taxes using free online software.

My taxes are very simple, I don't think with my current method of filing (H&R Block) I am at any meaningful risk of error or audit.

The risk is minimal. Remedy would be readily available.

These responses are especially notable because federal audit rates *are* extremely low for ordinary taxpayers—the overall

^{128.} Tax preparers do not assume liability for underpaid taxes discovered upon audit. Such answers were either referring to a desire to reduce certain tax penalties or reflected a misunderstanding of the law.

^{129.} Among respondents who stated that they were not willing to pay money to reduce risk, 29% seemed to misunderstand the question, rejected the premise of the question, or provided answers that didn't clearly explain the respondent's choice. For instance, "I do not believe that paying money will do this so I think it's wasteful to pay money towards this effort." The relatively high percentage of such answers reflects the fact that there were fewer answers in this sub-category compared to the other three sub-categories. The total number is only marginally higher than for the other sub-categories.

audit rate in 2019 was 0.25%.¹³⁰ So although a larger number of respondents expressed concern about audits than those who didn't, the worry-free members of the latter camp were arguably better informed.

A minority (7%) of respondents stated that they could not afford to pay to reduce risk or that they would rather save the money. For instance:

I am very cheap, and prefer to keep as much money invested or in my pocket as possible.

I don't generally think it's worth it to spend money unless it's very affordable for me as a low-income person

I see the benefits, but when you know how to do it, and money is tight, it makes more sense just to do the work and save money.

In addition to providing interesting context, these qualitative answers reiterated the need for quantitative data about the subjective costs of taxation. Viewed in isolation, we can draw very little in the way of conclusions or policy implications from them. Some people dislike doing their taxes; others don't mind it. Some people worry about making a mistake; others don't. However, when considered together with the results of the discrete choice experiment, the qualitative answers offer important explanations that inform our interpretation of the discrete choice survey results.¹³¹

C. POTENTIAL OBJECTIONS AND ROBUSTNESS CHECKS

1. External Validity

While online survey platform users may be representative of an "average" middle-income person, they may not reflect the full U.S. population.¹³² Moreover, our pre-screening procedure

132. We used Prolific, which takes greater care to gather a more representative sample of higher-quality users than some other survey services, but the same problems with those services may apply to Prolific as well. *See* Scott M.

^{130.} U.S. GOV'T ACCOUNTABILITY OFF., GAO-22-104960, TAX COMPLIANCE: TRENDS OF IRS AUDIT RATES AND RESULTS FOR INDIVIDUAL TAXPAYERS BY INCOME 6 (2022), https://www.gao.gov/assets/gao-22-104960.pdf [https://perma.cc/8Q6S-3FSF].

^{131.} See supra text accompanying note 97.

explicitly selected for taxpayers with more complex returns by limiting the sample to those who had filed tax returns for at least three years and expected to spend at least three hours on their returns.¹³³

Consequently, low-income taxpayers were significantly underrepresented in our sample compared to the overall population of tax filers: only 6.2% of our respondents reported household incomes below \$20,000 per year, compared to 28.2% of federal tax filers.¹³⁴ On the other hand, higher-income taxpayers were overrepresented, with 34.7% of our respondents reporting household incomes above \$100,000 per year, as opposed to 19.4% of federal tax filers.¹³⁵

The low-income and high-income respondents in our sample might also be non-representative. For low-income respondents, the three-hour tax compliance threshold means that our low-income taxpayers likely spend more time on their taxes than the average low-income person with a simple tax return. Low-income respondents in our sample might therefore have a higher-than-normal willingness to pay (since time spent on taxes is correlated with willingness to pay).¹³⁶

There may be certain biases related to higher-income respondents as well. Generally speaking, because higher-income taxpayers tend to have higher willingness to pay in general (without controlling for other variables) and because higher-

133. The survey also excludes non-filers, since, by definition, they spend less than three hours on tax compliance. Our findings therefore ignore any tax compliance costs that non-filers experience—e.g., from structuring transaction to avoid taxes or researching tax filing requirements.

134. INTERNAL REVENUE SERV., INDIVIDUAL INCOME TAX RETURNS: COM-PLETE REPORT 2020, at 49 tbl.1.1 (2022), https://www.irs.gov/pub/irs-pdf/p1304 .pdf [https://perma.cc/6BCM-FAZR] (subtracting the percentage of tax returns reporting a household income above \$20,000 (71.8%) from the total percentage of tax returns (100%)).

135. *Id*.

136. It's also possible that non-filers could have *higher* willingness to pay, since taxpayers might choose not to file because of high subjective costs of filing.

Smith et al., A Multi-Group Analysis of Online Survey Respondent Data Quality: Comparing a Regular USA Consumer Panel to MTurk Samples, 69 J. BUS. RSCH. 3139, 3140, 3142 (2016) (finding that a Mechanical Turk survey panel in the United States had lower income and lower average education levels, as well as a higher proportions of non-white survey takers, compared to a sample drawn from a "reputable commercially maintained Internet survey panel"). We did not attempt to balance the sample geographically, and our sample necessarily excludes respondents who do not read or write English.

income taxpayers were overrepresented in our sample, our estimates of individual willingness to pay may be excessively high.¹³⁷ However, the high-income respondents in our sample were those who had time to take online surveys. Common sense might suggest that such people do not have particularly complex business affairs or taxes.¹³⁸ This might drive estimates lower than they otherwise would be.

It's also possible that online survey respondents are unusually comfortable with administrative busywork or technologically savvy, as evidenced by their willingness to participate in online surveys. Online survey takers could therefore plausibly have unusually low subjective costs. Our results offer some evidence against this concern. When asked to rate how pleasant or unpleasant they find tax compliance work, 68% of respondents selected that they find tax compliance work to be "somewhat unpleasant" or "extremely unpleasant." This proportion matches, and in fact slightly exceeds, the proportion of respondents who reported negative feelings about tax compliance in nationwide polls.¹³⁹

Although we attempted to make our sample broadly representative, there are specific subpopulations that were not represented sufficiently for us to make statements about them with confidence. For example, small-business owners and self-employed taxpayers may have substantially higher subjectively measured compliance costs, but we do not address them separately.

The timing of the survey might have also affected the results. We ran the survey during the fall, several months after the standard tax filing deadline of April 15th, and several months before the filing season start in January. It's therefore possible our respondents had forgotten the full weight of the aggravation and anxiety that they feel during tax season. Alternatively, the passage of time may have caused some to inflate the unpleasantness of tax filing or underweight the psychological benefits. It's

^{137.} *Cf.* DeVoe & Pfeffer, *supra* note 117, at 8 ("Beyond idiosyncratic preferences, the primary determinants of a person's willingness to trade more of their time for money should be how much money they earn and how much time they spend working.").

^{138.} In other words, individuals with particularly complex business affairs or taxes may be less likely to have time to take online surveys.

^{139.} *See supra* notes 49–50 and accompanying text (discussing negative public feelings toward filing taxes).

not clear as a conceptual matter whether one should focus on "instant utility" measured during an experience or "remembered utility" measured in hindsight.¹⁴⁰ Both are valid and might lead to potentially different results, although evidence suggests that remembered utility (which is what we study here) affects actual decision-making more than instant utility.¹⁴¹

A final concern related to external validity is the possibility of framing effects: respondents' answers may have depended on arbitrary survey characteristics, like question order or wording.¹⁴² In our case, for instance, asking respondents about their tax filing experiences at the start of the survey might have focused their attention on the negative aspects of tax compliance, leading them to overestimate how much they would be willing to pay to reduce the time and risk associated with filing taxes. While no survey is entirely free of framing effects,¹⁴³ we tried to minimize this problem by phrasing questions in as neutral a manner as possible.

2. Cheap Talk

We tested respondents' willingness to pay for tax services in part to reduce the problem of "cheap talk," which plagues survey research that relies on stated preferences.¹⁴⁴ Where a respondent's expression of dislike is qualitative and vague, it may reflect a superficial attitude that would change upon further inspection.

^{140.} See generally Daniel Kahneman et al., Back to Bentham? Explorations of Experienced Utility, 112 Q.J. ECON. 375, 375 (1997) (introducing and discussing the distinction between instant and remembered utility); Adam Oliver, Distinguishing Between Experienced Utility and Remembered Utility, 10 PUB. HEALTH ETHICS 122 (2017) (discussing instant and remembered utility).

^{141.} Daniel Kahneman & Alan B. Krueger, *Developments in the Measurement of Subjective Well-Being*, 20 J. ECON. PERSPS. 3, 6 (2006) ("Further evidence suggests that individuals' choices are affected by their remembered utility... not the profile of their experiences."). This is completely intuitive since consumer decisions are made on the basis of remembered utility and without direct access to measurements of instant utility.

^{142.} Jacob Goldin & Daniel Reck, *The Analysis of Survey Data with Framing Effects*, 73 AM. STATISTICIAN 264, 264 (2019) ("A well-known difficulty in survey research is that how survey-takers respond to a question may depend on seemingly arbitrary details about how the question is asked.").

^{143.} See *id.* ("Such *framing effects* arise in many contexts; large literatures in psychology, political science, communications, and marketing are devoted to documenting and explaining their presence.").

^{144.} See Crawford, *supra* note 69, at 286 (describing experiments in which "talk is cheap," meaning that respondents' answers have no effect on their lives).

The discrete choice experiment design reduces this problem by presenting respondents with choices that mimic those they make in a real-world marketplace.

However, there's a further cheap talk problem. Because our survey was hypothetical and respondents never actually paid for the tax filing service we described, their choices may still have been reflexive and unconsidered. This problem is generally known as "hypothetical bias."¹⁴⁵ A substantial literature suggests that hypothetical bias generally leads to overestimation that is, respondents generally give excessively high willingnessto-pay estimates in hypotheticals compared to their willingness to pay in real life.¹⁴⁶ Thus, hypothetical bias would have caused our estimates of subjective costs to be too high.

Together with framing effects, the problem of cheap talk reveals a broader difficulty with all so-called "stated preference" surveys: respondents' stated preferences may not reflect their true preferences.¹⁴⁷ Many empiricists prefer to use revealed preferences where possible, for example, by looking at consumers' willingness to pay for an accountant to save time. In the economics literature, the debate mostly concerns contingent valuation surveys that ask individuals how they value inherently hard-to-value public goods, such as clean river water or saving the California condor from extinction.¹⁴⁸ The "embedding effect" leads people to offer very similar values for very different interventions.¹⁴⁹ For instance, they might provide "willingness-to-pay to clean up one lake roughly equal to that for cleaning up five

^{145.} E.g., David A. Hensher, *Hypothetical Bias, Choice Experiments and Willingness to Pay*, 44 TRANSP. RSCH. PART B 735, 735 (2010) ("The extent to which individuals might behave inconsistently, when they do not have to back up their choices with real commitments, is linked to the notion of *hypothetical bias*...." (emphasis added)).

^{146.} *Id.* at 739 ("The accumulating evidence . . . suggests that individuals in . . . hypothetical [contingent valuation] studies exaggerate their [willingness to pay] for private and public goods.").

^{147.} *See* Johnston et al., *supra* note 76, at 321 (discussing this problem with stated preference surveys).

^{148.} See Daniel Kahneman & Jack L. Knetsch, Valuing Public Goods: The Purchase of Moral Satisfaction, 22 J. ENV'T ECON. & MGMT. 57, 58–59 (1992) (describing issues associated with the contingent valuation method).

^{149.} *Id.* (describing the embedding effect to occur when "the same good is assigned a lower value if [willingness to pay] for it is inferred from [willingness to pay] for a more inclusive good rather than if the particular good is evaluated on its own").

lakes."^{150} Such illogical responses produce implausible per-unit variable estimates."^{151} $\,$

However, these criticisms tend to deal narrowly with surveys that ask respondents to value public goods.¹⁵² In contrast, we asked individuals to value *private goods*. This distinction matters for at least two reasons. First, the embedding effect applies specifically to public goods.¹⁵³ One prominent hypothesis for its mechanism is that respondents are putting a dollar figure on the "warm glow" from prosocial behavior, like protecting the environment.¹⁵⁴ This warm glow is constant whether one lake or five is protected. In contrast, respondents should (and in our survey did) ascribe a larger value to more, rather than less, time savings. Second, with private goods, respondents likely have more personal experiences to draw from when arriving at a willingness-to-pay estimate. They have almost certainly previously considered how much they would pay for a service that saves them time or reduces their anxiety. In contrast, many respondents may have little personal experience to guide them in how to value public goods that have little direct observable effect on their lives.

More broadly, while studies of revealed preferences can be valuable, in the context of tax regulation, stated preferences can be more powerful in exploring the attitudes of vulnerable populations. Revealed preference studies often focus on individuals at the margin, for design reasons.¹⁵⁵ In the context of tax

^{150.} Peter A. Diamond & Jerry A. Hausman, *Contingent Valuation: Is Some Number Better than No Number*?, 8 J. ECON. PERSPS. 45, 46 (1994).

^{151.} A related problem is "scope sensitivity," which we addressed by providing respondents with the full range of attributes and prices in advance of the choice cards. For further explanation, see infra Appendix Part A.2.

^{152.} Public goods are generally non-rivalrous, meaning one person's use doesn't diminish another's, and non-excludable, meaning that people can't be excluded from using the good. Military protection is a classic example of a public good.

^{153.} See Kahneman & Knetsch, supra note 148, at 58–59 (identifying the embedding effect only in the context of valuation of public goods); Diamond & Hausman, supra note 150, at 46 ("The embedding effect is usually thought to arise from the nonexistence of individual preferences for the public good in question").

^{154.} Diamond & Hausman, *supra* note 150, at 47 ("Individuals may receive a 'warm glow' from expressing support for good causes").

^{155.} *E.g.*, Benzarti, *supra* note 46, at 38 (observing the revealed preferences of taxpayers deciding between itemizing deductions or claiming the standard deduction to estimate certain specific hassle costs of tax compliance).

compliance, a revealed preference study might observe taxpayers in deciding whether to pay for an accountant, or whether to claim a certain tax credit.¹⁵⁶ But many taxpayers might be far from these margins—for example, because they have too little income or such simple tax returns that it would be pointless to pay for assistance. Stated preference surveys allow us to consider these inframarginal individuals as well. Because tax simplification reforms would have major implications for most American households, we found the greater inclusivity of stated preference surveys to be appealing, despite their limitations.

3. Statistical Validity, Survey Fatigue, and Attenuation Bias

Consistent with best research practices,¹⁵⁷ we pre-registered our experimental design and statistical models with the Open Science Framework.¹⁵⁸ We also conducted two robustness checks to confirm the validity of our experimental design. Section E of the Appendix describes how our study passed tests for the Independence of Irrelevant Alternatives, an important statistical prerequisite for the validity of our discrete choice model.¹⁵⁹

An additional concern was that respondents may have experienced survey fatigue, decreasing the quality of their responses as the survey went on. We tested this by re-estimating implicit prices based only on the responses in our sample to the first six choice cards.

^{156.} See, e.g., id.

^{157.} See generally John P.A. Ioannidis, *Why Most Published Research Findings Are False*, 2 PLOS MED. 696 (2005) (describing how many published findings are false due to publication bias and p-hacking, and suggesting pre-registration as one remedy).

^{158.} Subjective Costs of Taxation, OPEN SCI. FRAMEWORK REGISTRIES (Oct. 18, 2022), https://osf.io/s8q7h [https://perma.cc/G4QT-KTVG].

^{159.} Atkinson et al., *supra* note 76, at 131 ("[S]elections from the choice set must obey the Independence from Irrelevant Alternatives (IIA) property, \ldots which states that the relative probabilities of two options being selected are unaffected by the introduction or removal of other alternatives." (citation omitted)).

Table 3 shows the results from this analysis.

Attribute	Implicit Price	95% CI
Risk Elimination	\$72.36	\$59.35 - \$85.38
Time Elimination: 1 Hour	-\$7.09	-\$17.31 - \$3.12
Time Elimination: Half	\$11.40	0.31 - 22.48
Time Elimination: All	\$54.13	\$39.73 - \$68.53
Government	\$2.64	-\$3.05 - \$8.35

Table 3: Main Effects Implicit Prices

These implicit prices are almost identical to the implicit prices generated from the full sample, suggesting that survey fatigue didn't substantially affect the later responses.

Finally, it's possible that respondents may have failed to understand the survey, which would introduce noise into our estimates and systematically bias implicit prices toward zero—a phenomenon known as attenuation bias. Attenuation bias is a widespread concern in empirical legal studies, and we tried to ameliorate it by using an instrument that was as easy to understand as possible. We also filtered aggressively using comprehension and attention checks, to ensure that participants understood the frame and paid appropriate attention.

III. IMPLICATIONS FOR THEORY AND POLICY

A. IMPLICATIONS FOR THEORY

These survey results have several important implications for economic theory relating to the estimation of tax compliance costs, the valuation of forgone time, and the marginal disutility of work.

These findings challenge a tacit assumption prevalent in much of the tax compliance cost literature: that psychological costs impose a burden on taxpayers *in addition to* their time costs of tax compliance.¹⁶⁰ However, our discrete choice experiment revealed implicit prices for time savings that were on average below respondents' hourly wages.¹⁶¹ For their last hour of tax compliance work, taxpayers were willing to pay *far* less than

^{160.} *E.g.*, Evans, *supra* note 29, at 451 (stating that psychological costs occur "[i]n addition to" the time costs of tax compliance).

^{161.} See supra Part II.B.1.d (discussing this result).

their hourly wage, perhaps nothing. This finding undermines any straightforward, arithmetic combination of time costs and psychological costs. Indeed, there may not be a credible way to combine objective and subjective tax compliance costs. The two concepts measure distinct things—one tallies the market cost of lost labor time, whereas the other measures the personal burden of the activity. Perhaps a full accounting of compliance costs requires calculating and providing separate estimates for both.

Our findings also suggest that there may indeed exist psychological benefits of tax compliance work that offset the psychological costs. Scholarship on tax compliance has been somewhat circumspect on this point. Among our respondents, 12% reported that they find tax compliance work to be "somewhat pleasant" or "extremely pleasant."¹⁶² For these taxpayers, it may be that the subjective benefits of tax compliance outweigh any subjective costs they experience. Several of the qualitative answers suggested the existence of psychological benefits as well. Some respondents reported that they "enjoy" doing their taxes, that they learn from the process, or that doing tax compliance work helps them in other ways.¹⁶³ These findings support scholarship on the diverse benefits of the tax filing process.¹⁶⁴ While much of this scholarship relates to society-wide benefits from tax filing, our results suggest that there are individual-level benefits as well.¹⁶⁵

Our survey results also have implications for economic theory outside the study of taxation. Perhaps most importantly, our results call into question the nearly universal practice of using market wages to value the time that people spend on regulatory compliance.¹⁶⁶ Analysts use market wages to value people's time for nearly all cost-benefit analysis across government agencies, not just for tax policy analysis.¹⁶⁷ Our findings suggest that these

^{162.} Another 20% reported that they find tax compliance work to be neither pleasant nor unpleasant.

^{163.} See supra Part II.B.2.

^{164.} See supra Part I.B.2.

^{165.} E.g., ZELENAK, supra note 62, at 4 ("The return-preparation process serves—or at least has the potential to serve—the important civic purpose of recognizing and formalizing the financial responsibilities of citizenship."); cf. WILLIAMSON, supra note 59, at 180–82 (describing taxpaying as a source of pride among survey respondents and Americans generally).

^{166.} See supra notes 35-37 and accompanying text.

^{167.} This valuation method has been repeatedly and persuasively challenged elsewhere. *See supra* note 19 (compiling articles that challenge this valuation method).

estimates should be interpreted more narrowly than they have been. Using market wages to value time might tell us, for instance, how much it would cost to pay a third party to do the task, or how much value would be added to the national economy if a person were working instead of filling out paperwork.¹⁶⁸ However, as our results suggest, wage-based estimates do *not* tell us the welfare burden that people personally experience. Thus, wage-based paperwork burden calculations that purport to measure the total cost that a regulation imposes on society should be taken with a grain of salt.

Finally, our findings challenge the assumption that people consistently experience increasing marginal disutility of effort.¹⁶⁹ Contrary to some conventional economic theory, our respondents experienced *decreasing* marginal disutility of tax effort. Further research is necessary to ascertain how broadly this finding applies across taxpayers and across other tasks. For one thing, it's likely that marginal disutility would increase once taxpayers spend more than a certain number of hours on their taxes. Moreover, other types of regulatory compliance might exhibit more traditional increasing disutility of effort. For instance, surely waiting in line at the DMV becomes more unpleasant the longer one does it.

B. IMPLICATIONS FOR POLICY

Our findings have important policy implications for lawmakers and agency officials. First, analysts should reevaluate how they presently calculate compliance costs. Compliance cost estimates based on market wages are at best oversimplified and at worst significantly overestimated. Policymakers, analysts, and agency staff should consider using alternative methods either instead of or in addition to the current methods. If researchers determine that market wages remain the best method to estimate the time costs of tax compliance, they should consider offering additional context in all publications that use such

^{168.} The latter interpretation relies on questionable assumptions, including that a person can easily increase their hours of labor.

^{169.} *Cf.* Liu, *supra* note 107, at 188 ("In addition we assume that the more the effort, the greater will be the marginal disutility. That is, there is a diminishing marginal utility of effort or equivalently, an increasing marginal disutility of effort.").

estimates. Such context might explain, for instance, how to narrowly interpret market-based cost estimates.

Our specific findings also counsel policymakers to think big when it comes to reducing tax compliance work and to focus on anxiety over aggravation. Our final Subsection addresses considerations regarding the public or private nature of tax simplification and automation services.

1. Think Big

Given the declining marginal cost of tax compliance work, policymakers should not waste time on small reforms that only slightly reduce taxpayers' time or aggravation. Reforms that would eliminate all tax compliance activities will be disproportionately more valuable to taxpayers than reforms that merely shave an hour or two off their total tax preparation time.

These findings lend some support to proposals for returnfree filing, in which many (most) individual taxpayers do not need to file a tax return.¹⁷⁰ There are various ways such a system could work. For instance, under an "exact-withholding" system, like that used in the United Kingdom and Germany, the IRS would attempt to withhold the exact amount from taxpayers' paychecks during the year, obviating the need to file a return at the end of the year.¹⁷¹ Alternatively, with "tax agency reconciliation," taxpayers could voluntarily provide the IRS with information that would allow the IRS to calculate their tax liability based on income reported by employers and other third parties.¹⁷² The taxpayer could then review this government-populated tax return before filing it. Either system would dramatically reduce the time that taxpayers spend on tax compliance. Both would also likely reduce the risk of error and audit for most

^{170.} See supra notes 99–101 and accompanying text (discussing return-free filing); Thomas J. Healy, *Return-Free Filing Would Revolutionize Americans' Taxes*, BARRON'S (Dec. 13, 2022), https://www.barrons.com/articles/return-free -filing-would-revolutionize-americans-taxes-irs-51670884852 [https://perma.cc/ Y8M8-46QM].

^{171.} See What Is Return-Free Filing and How Would It Work?, TAX POL'Y CTR.: BRIEFING BOOK, https://www.taxpolicycenter.org/briefing-book/what -return-free-filing-and-how-would-it-work [https://perma.cc/XN4N-ME94] (last updated May 2020) ("If an income tax system were simple enough, the government could withhold taxes owed and do its own accounting at the end of the year without much help from taxpayers.").

^{172.} Id. (describing a tax agency reconciliation system).

taxpayers, especially audits triggered when taxpayers' reported information conflicts with third-party reported information.¹⁷³

A fully return-free system in the U.S. would require significant policy changes. Several features of our tax code make it difficult or impossible to implement the exact-withholding system described above. Because the United States allows joint filing, it's difficult to calculate any individual's specific tax liability without information about their partner that may not be available at the time of withholding.¹⁷⁴ Moreover, the United States offers many tax credits and deductions whose amount depends on conduct that the government does not directly track (for example, whether an individual will buy an electric car or experience a casualty loss).¹⁷⁵ Other countries with exact-withholding regimes utilize individual filing and don't offer these sorts of credits and deductions.¹⁷⁶ It's difficult to imagine how the United States could eliminate all taxpayer involvement without reforming our tax code along the same lines.

Deciding whether to enact a return-free tax system requires a careful weighing of all the costs and benefits of such a system. Eliminating subjective costs is merely one benefit. Moreover, as discussed above, the value of this benefit may be significantly lower than previous estimates have suggested.¹⁷⁷ Other costs and benefits may be difficult to value. For instance, eliminating tax filing would mean eliminating the primary way that U.S.

^{173.} An open question in the design of a return-free system is who would bear the responsibility for mistakes made on government-prepared returns. *Id.* Certainly, if the government bore the risk of error, it would dramatically reduce taxpayers' anxiety. It would also presumably incentivize the IRS to take special care in preparing taxpayers' returns.

^{174.} See U.S. DEP'T OF THE TREASURY, REPORT TO THE CONGRESS ON RE-TURN-FREE TAX SYSTEMS: TAX SIMPLIFICATION IS A PREREQUISITE 7 (2003), https://home.treasury.gov/system/files/131/Report-Return-Free-2003.pdf

[[]https://perma.cc/W5MC-DJTM] ("Tax systems that rely on exact withholding often have structural features that facilitate taxation at source. For example, the individual is generally the unit of taxation (unlike the U.S. system in which married couples are taxed as a unit).").

^{175.} See *id.* ("Relative to the U.S. income tax system, [exact withholding] systems are also characterized by fewer rates, fewer deductions, and fewer tax credits.").

^{176.} Id.

^{177.} See supra Part II.B.1.d.

residents interact with the federal government.¹⁷⁸ A return-free system might therefore lead to less political engagement or lower levels of government accountability.¹⁷⁹

As a less drastic alternative, the IRS could allow taxpayers to pre-fill their returns with all third-party information the IRS has received, as well as information from prior years' tax returns. According to recent analysis, for over 60 million taxpayers, no additional information would be required.¹⁸⁰ A pre-filled tax return would nearly eliminate aggravation costs for these people. And, by including all the information the IRS currently has, government-prepared returns would eliminate the audits that currently occur when taxpayers neglect to include thirdparty reported information. Again, deciding whether to pursue such a policy requires weighing all the costs and benefits. Reducing subjective costs is merely one of the benefits. It's also worth noting that for those taxpayers who still have to provide significant information on their tax return-such as small business owners-pre-filled tax returns would offer only marginal subjective cost savings.

Notwithstanding declining marginal costs of time, policymakers should continue to pursue reforms that simplify the taxfiling process. Many reforms, such as attempts to modernize tax filing with new technology, have low costs or could even decrease net administrative costs.¹⁸¹

Policymakers should also continue to pursue simplification reforms that offer additional benefits aside from merely reducing compliance costs. As one example, policymakers could reform child-claiming rules in the EITC and Child Tax Credit to expand

^{178.} ZELENAK, *supra* note 62, at 4 ("For most of the people most of the time, the most prominent and meaningful connection with the federal government is through the income tax.").

^{179.} *Cf. id.* at 35–38 (discussing how the April 15th tax filing deadline spurs a national conversation about the nation's fiscal policies).

^{180.} Lucas Goodman et al., Automatic Tax Filing: Simulating a Pre-Populated Form 1040, at 1 (Nat'l Bureau Econ. Rsch., Working Paper No. 30008, 2022), http://www.nber.org/papers/w30008 [https://perma.cc/KJ8T-KHEJ].

^{181.} See generally INTERNAL REVENUE SERV., IRS INTEGRATED MODERNIZA-TION BUSINESS PLAN (2019), https://www.irs.gov/pub/irs-pdf/p5336.pdf [https:// perma.cc/Z7CY-F2N9] (describing proposals to improve tax administration through new technology).

and simplify eligibility standards for claimants.¹⁸² Although such reforms would likely simplify how taxpayers evaluate their eligibility for child-based credits, the time saved per household would be miniscule. But, in addition to saving time, better-designed rules would likely reduce improper claims by aligning with families' child-care realities. Such a reform would be eminently worthwhile, despite the small amount of time saved.

2. Minimize Mistakes

Our survey suggests that taxpayers are bothered more by the risk of accidental errors on their tax returns than by aggravation. Policymakers should therefore focus on reforms that reduce the risk of mistakes, rather than prioritizing simplification reforms that aim to reduce time spent on tax compliance.

One of the most important developments in tax compliance over recent decades has been the dramatic increase in the "computational complexity" of tax filing.¹⁸³ Tax filing software like TurboTax largely automates the calculations required to file taxes, which taxpayers previously had to compute by hand. Legislators have responded by massively increasing the amount of computation required to file a tax return.¹⁸⁴ While scholars have pointed out both costs and benefits to this increase in computational complexity,¹⁸⁵ these scholars generally assume that computational complexity is essentially costless to taxpayers.¹⁸⁶ Some provisions that increase computational complexity—like

184. See *id.* at 99 ("[I]t is no accident that the increase in tax return complexity has coincided with the triumph of return preparation software.").

185. Compare id. at 93–94 (arguing that computational complexity makes the tax system less democratically responsive and reduces the impact of tax incentives), with id. at 93 (arguing that desirable equitable programs can sometimes be computationally complex), and David I. Walker, Tax Complexity and Technology, 97 IND. L.J. 1095, 1105–17 (2022) (describing the benefits of computational complexity, including a reduction in inefficient tax planning).

186. *E.g.*, Walker, *supra* note 185, at 1097 (describing computational complexity as "trivial and acceptable" given tools like TurboTax); Zelenak, *supra* note 183, at 92 ("[T]he practicalities of return preparation impose virtually no limitations on the computational complexity to which Congress may subject the average taxpayer.").

^{182.} See Jacob Goldin & Ariel Jurow Kleiman, Whose Child Is This? Improving Child-Claiming Rules in Safety-Net Programs, 131 YALE L.J. 1719, 1764–82 (2022) (suggesting and evaluating such reforms to child-claiming rules).

^{183.} See generally Lawrence Zelenak, Complex Tax Legislation in the TurboTax Era, 1 COLUM. J. TAX L. 91 (2010) (coining the phrase "computational complexity" and describing this phenomenon).

the Alternative Minimum Tax, for example, a parallel tax system that essentially requires taxpayers to compute two sets of tax liabilities and use the higher one¹⁸⁷—may not require taxpayers to laboriously commit pen to paper, but they surely increase the likelihood that taxpayers will make mistakes. And, just as importantly, they may also increase taxpayers' *perception* that they are likely to make mistakes.

More generally, our results suggest that Congress should attempt to close "traps for the unwary," areas of the tax code that are surprising or unintuitive in ways that invite taxpayer errors.¹⁸⁸ While traps for the unwary are unpopular in theory, advocates for simplification tend to neglect them in favor of more measurable reductions in time spent on tax compliance.¹⁸⁹ Our findings suggest that eliminating these traps is as important, indeed *more* important, than simply reducing time spent on taxes.

Finally, our findings offer a counterpoint to the popular argument that the IRS should promote fear of audits to motivate tax compliance.¹⁹⁰ While anxiety may motivate compliance, it also imposes psychological costs on individuals in the form of discomfort and stress.¹⁹¹ Moreover, in a system where third-party reporting confirms most of the information on the tax return, anxiety plays a limited role in encouraging compliance for most

^{187.} I.R.C. § 55; Zelenak, *supra* note 183, at 99–104 (discussing the computational complexity of the Alternative Minimum Tax).

^{188.} See, e.g., Greenstein et al., supra note 26, at 3–4 (explaining how the complexity of the EITC rules accounts for most erroneous claims); Goldin & Jurow Kleiman, supra note 182, at 1759 (explaining how current child-claiming rules in tax credits often misalign with taxpayers' intuitions about caregiving).

^{189.} *E.g.*, York & Muresianu, *supra* note 10, at 3–5 (describing how the TCJA simplified tax filing by increasing the standard deduction).

^{190.} *Cf.* Joshua D. Blank & Daniel Z. Levin, *When Is Tax Enforcement Publicized?*, 30 VA. TAX REV. 1, 31–33 (2010) (considering arguments for and against publicizing IRS audit activity to encourage tax compliance).

^{191.} Researchers have found that stress and anxiety lead people to procrastinate, causing them to put off important tasks until the last minute. See Charlotte Lieberman, Why You Procrastinate (It Has Nothing to Do with Self-Control), N.Y. TIMES (Mar. 25, 2019), https://www.nytimes.com/2019/03/25/smarter -living/why-you-procrastinate-it-has-nothing-to-do-with-self-control.html [https://perma.cc/G7UX-JRNR] ("Procrastination . . . [is] a way of coping with challenging emotions and negative moods induced by certain tasks — boredom, anxiety, insecurity, frustration, resentment, self-doubt and beyond.").

taxpayers.¹⁹² These considerations suggest that fomenting anxiety¹⁹³ may do more harm than good.

The IRS could reduce taxpayers' anxiety by better publicizing the various ways that taxpayers can work with the IRS if they make a mistake on their tax return or can't afford to pay their taxes right away. For instance, the IRS offers qualifying taxpayers installment payment plans for those who can't afford to pay their full tax balance right away.¹⁹⁴ The IRS could publicize these payment programs in the lead-up to the April 15th tax-filing deadline.

Additionally, penalties for underreporting are quite unlikely as long as taxpayers file their tax returns on time and in good faith.¹⁹⁵ For most taxpayers, a mistake on a tax return simply means paying any additional taxes plus interest.¹⁹⁶ And, convictions for criminal tax evasion or tax fraud require an extremely high evidentiary bar and do not apply to taxpayers who made an honest mistake in their tax filings.¹⁹⁷ Our qualitative findings suggest that most taxpayers don't understand this, since many more expressed concerns about the risk of audits and penalties than (accurately) noted that their risk of facing audit or penalties was low.¹⁹⁸ Publicizing the difference between good-faith tax

^{192.} Leandra Lederman & Joseph C. Dugan, *Information Matters in Tax Enforcement*, 2020 BYU L. REV. 145, 160–65 (emphasizing the importance of third-party reporting to tax compliance and enforcement).

^{193.} The IRS may increase taxpayers' anxiety by, for instance, publicizing high-profile tax enforcement actions in the days leading up to the tax filing deadline. *See* Blank & Levin, *supra* note 190, at 17 tbl.1 (comparing the number of tax enforcement press releases in time windows leading up to tax day versus the rest of the year).

^{194.} Additional Information on Payment Plans, INTERNAL REVENUE SERV., https://www.irs.gov/payments/payment-plans-installment-agreements#plandef [https://perma.cc/6S8Z-G57H].

^{195.} The tax code imposes an accuracy-related penalty only if specific conditions are met, including negligence, disregard of rules, or substantial understatement of tax. I.R.C. § 6662. A substantial understatement means an understatement that exceeds the higher of 10% of the amount of tax required or \$5,000. *Id.* Moreover, a taxpayer can avoid such a penalty if they have reasonable cause for their position and engaged in good faith. I.R.C. § 6664(c).

^{196.} See generally I.R.C. § 6601(a) (providing that interest will accrue on unpaid taxes during the period they remain unpaid).

^{197.} *See* I.R.C. § 6664(c) (providing good faith exception to fraud penalties); *United States v. Moran*, 493 F.3d 1002, 1013 (9th Cir. 2007) (defining the government's burden of proof in a criminal tax case).

^{198.} See supra Part II.B.2.

filing and abusive tax behavior, as well as the remedies available to taxpayers who make an honest mistake, might help to reduce taxpayer anxiety.

Minimizing compliant taxpayers' fear of audits need not always reduce deterrence for noncompliant taxpayers. Offering a more seamless audit process would reduce audit costs for most taxpayers without appreciably affecting the decision-making of tax evaders. For instance, the IRS could strive to speed up audit timelines¹⁹⁹ or simply adopt a more "customer-friendly" attitude during audits. Any reform that decreases audit costs across the board could be paired with higher penalties for bad actors in a manner that cancels out any reduction in deterrent effect. In light of our findings that stress and worry about audits are a major source of disutility for ordinary taxpayers, the IRS should do more to separate those good actors who worry unnecessarily about audit risk from those bad actors who should be made to worry more.

3. Public or Private, Taxpayers Don't Care

Respondents in our survey did not place a premium on the public or private nature of the tax service provider.²⁰⁰ In their view, it didn't matter whether tax simplification was provided directly by the government or outsourced to a private third party. We offer two policy implications from this finding. First, policymakers should decide on the public/private nature of tax services based on factors other than taxpayers' preferences. Second, mistrust of government is unlikely to dissuade taxpayers from using an otherwise-desirable government-provided service.

Many factors aside from taxpayer preference can bear on governments' decision of whether to provide a service directly or outsource to a private company. Cost is foremost among them. For instance, private companies may benefit from existing proprietary technology or economies of scale. On the other hand, government agencies might be able to integrate certain services more easily with existing government systems, which could

^{199.} INTERNAL REVENUE SERV., INTERNAL REVENUE MANUAL § 4.10.2.2.2 (2019), https://www.irs.gov/irm/part4/irm_04-010-002r#idm140555766782096 [https://perma.cc/59S2-C9P4] (providing that "the audit and disposition of an income tax return should be completed" within a 26-month or 27-month time-line).

^{200.} See infra Appendix Part D for the regression table with survey results.

reduce the cost of public services.²⁰¹ The government might also be able to take advantage of its dominant position in the market to obviate the need for expensive advertising.

Certain factors might also affect the relative quality of public or private services. Because budget appropriations are somewhat beyond agencies' control, government agencies like the IRS might have cost constraints that limit long-term technological capacity, make it difficult for them to upgrade technology at pace with the private sector, or undermine their ability to provide robust customer support. On the other hand, because private companies are driven by profit motives, they might face incentives to upsell users or mislead users into paying higher fees than necessary.²⁰²

Certain services may be considered too important to outsource entirely to the private sector. Schools, public transportation, and postal services are examples of such core functions. Public provision ensures that these services exist even without sufficient profit. Easy tax filing has never previously been considered a core public service. But norms can change; public schooling wasn't broadly provided in the United States until the second half of the nineteenth century.²⁰³ Especially as the tax code becomes increasingly used to deliver public support, like the EITC; or essential needs, like health insurance tax credits; norms may shift to expect that the IRS provide a free, easy way for everyone to file a tax return.

Our findings also suggest that mistrust of government does not seem to be an overriding concern in the context of tax compliance. Some may find this result somewhat surprising. Many have conjectured that taxpayers would be reluctant to adopt a tax preparation service offered directly by the IRS due to

^{201.} For a discussion of such integration in the context of a public tax filing system, see *IRS Direct File: Independent Third Party Report to Congress*, NEW AMERICA 45 (2023), https://dly8sb8igg2f8e.cloudfront.net/documents/

Independent-Third-Party-Report-to-Congress-508c.pdf [https://perma.cc/VQL8-QRLH].

^{202.} See, e.g., Justin Elliott, Intuit Will Pay Millions to Customers Tricked into Paying for TurboTax, PROPUBLICA (May 4, 2022), https://www.propublica .org/article/intuit-will-pay-millions-to-customers-tricked-into-paying-for -turbotax [https://perma.cc/H4NG-3CMB] (explaining a settlement by Intuit to

compensate users who paid for TurboTax despite being eligible for free services).

^{203.} Nancy Kober & Diane Stark Rentner, *History and Evolution of Public Education in the US*, CTR. ON EDUC. POLY 4 (2020), https://files.eric.ed.gov/fulltext/ED606970.pdf [https://perma.cc/L5BA-KAYP].

mistrust of government.²⁰⁴ In challenging this assumption, our findings align with other recent survey research conducted by the IRS. In taxpayer surveys, the IRS recently found that 72% of taxpayers would be "very interested" or "somewhat interested" in using an IRS-run public filing system.²⁰⁵ Taxpayers cited trust in the IRS as one of the main reasons for their interest, specifically trust in the security and privacy of their information.²⁰⁶

Trust in government is complex and multifaceted. Taxpayers might trust the IRS to keep their information secure, but not trust it to ensure they receive the highest possible tax refund. Reported levels of trust likely depend on many factors, including who is asking the question and how the question is framed. Acknowledging these necessary caveats, our results suggest that mistrust of government may not dissuade taxpayers from using a public service they otherwise want. At a minimum, this finding should encourage policymakers to consider providing public services where other factors would support direct government provision.

CONCLUSION

This Article introduces the concept of "subjective costs" of tax compliance and offers one way to measure them, by asking taxpayers their willingness to pay to eliminate such costs. The results of our inquiry challenge many conventional beliefs about the compliance costs of income tax filing in the United States.

We found that the tedium, aggravation, and frustration of tax filing bothers taxpayers less than the risk of making a mistake, calling into question the laser focus of most tax simplification initiatives merely on saving time. We also found that the last hour of tax compliance work imposes little or no burden on taxpayers, challenging classical assumptions about increasing marginal costs of labor. And we found that taxpayers were indifferent between simplification services offered by a private company versus the government.

^{204.} See, e.g., Thomas Catenacci, Biden Admin Steamrolled Internal Study in Pursuit of Key Tax Priority, FOX NEWS (Apr. 27, 2023), https://www.foxnews .com/politics/biden-admin-steamrolled-internal-study-pursuit-key-tax-priority [https://perma.cc/Z9DL-3DD4] (reporting concerns about taxpayer distrust of IRS affecting adoption of public tax filing tool).

^{205.} IRS REPORT: DIRECT E-FILE SYSTEM, *supra* note 113, at 7.206. *Id.* at 8.

While objective compliance cost measures remain important, we suggest that there might be other things worth measuring and other ways to understand how compliance costs are borne by individual taxpayers. As policymakers continually revisit the possibility of tax automation and simplification, understanding subjective costs of taxation is extremely important. We hope that this research might inform those efforts and contribute to reforms that are valuable for taxpayers and for broader society.

A. SURVEY CONTENTS

We summarize some of the most important aspects of our survey below. The full contents of the survey are available in our Online Appendix.²⁰⁷

1. Respondent-Specific Data

Our survey began by collecting a variety of respondent-specific data. To calculate respondents' annual post-tax income, we asked for average take-home pay per pay period, as well as pay period (weekly, monthly, etc.). To calculate their average posttax hourly income, we also asked for their average hours worked per pay period. In addition, we asked about: (1) use of tax-preparation software; (2) use of a professional tax preparer; (3) whether the respondent had wage income, had small business income, took the EITC, or took the Child Tax Credit; (4) whether the respondent itemized deductions or took the standard deduction; (5) whether the respondent usually receives a tax refund or owe taxes; (6) the respondent's experience of tax compliance, ranging from extremely unpleasant to extremely pleasant on a Likert scale;²⁰⁸ (7) whether the respondent files a joint or single tax return; and (8) if the respondent filed jointly, whether the respondent or the respondent's partner spends more time on taxes.

We also asked how much time and money the respondent currently expects to spend on tax compliance this year, using the following questions. We dynamically changed the choice sets presented to each participant based on their responses to these questions.

Based on your experience during the past 3 years, approximately how many hours do you expect to personally spend on tax compliance activities for the 2022 tax year?

^{207.} The Online Appendix is available at http://minnesotalawreview.org/vol108-choi-jurow-kleiman-appendix.

^{208.} A Likert scale allows respondents to rate their "levels of agreement to statements of interest," with levels falling along a progressive range, for instance from 1 (strongly dislike) to 5 (strongly like). Che Cheng et al., *Can Likert Scales Predict Choices? Testing the Congruence Between Using Likert Scale and Comparative Judgment on Measuring Attribution*, 5 METHODS PSYCH. 1, 1 (2021).

Based on your experience during the past 3 years, approximately how much money do you expect to spend on tax compliance activities for the 2022 tax year?

In addition to these data, Prolific provided extensive demographic data on its survey participants (e.g., ethnicity, current state of residence), which we also used in our analysis and which is described at greater length in the Online Appendix.²⁰⁹

2. Disclosure of Possible Choice Sets

Prior to viewing the choice sets, each respondent received information about the range of choices that would be available to them in the discrete choice sets. Research suggests that advance disclosure regarding the full range of possible choices increases the likelihood that respondents' choices will exhibit "scope sensitivity"—that is, that their willingness to pay will increase appropriately with the scope or magnitude of the good offered—and reduces the likelihood that respondents' choices will vary according to the order in which the choices are presented, which would be an undesirable framing effect.²¹⁰

3. Choice Card Sets

Each respondent received seven choice card sets (as well as one attention check set that was excluded from the analysis). Figure 1 depicted an example choice card set.²¹¹ Each choice card contained four attributes, with the following values and descriptions:

^{209.} The Online Appendix is available at http://minnesotalawreview.org/vol108-choi-jurow-kleiman-appendix.

^{210.} See Ian J. Bateman et al., On Visible Choice Sets and Scope Sensitivity, 47 J. ENV'T ECON. & MGMT. 71, 90 (2004) (explaining scope sensitivity within contingent valuation studies).

^{211.} See supra Part II.A.2.

Time Savings	All (spend no time or effort on your taxes)	
	Half (spend [x/2] hours, instead of x hours)	
	1 Hour (spend [x-1] hours, instead of x hours)	
	No reduction (no change in time or effort spent)	
Risk Elimina-	Full elimination of risk (taxes are 100% accurate,	
tion	no risk of future audit)	
	No reduction (no change in accuracy or audit risk)	
Provider	Private company	
	The government	
Fee	\$5	
	\$15	
	\$25	
	\$40	
	\$50	
	\$100	

Table 4: Choice Attributes

The fee levels were chosen based on the pilot. Using the empirical distribution of willingness to pay to eliminate compliance costs in the pilot, we conducted a Monte Carlo simulation of willingness to pay for the combination of choice attributes in our actual discrete choice cards, then selected the fee levels based on evenly spaced percentiles in the simulated willingness-to-pay distribution. This method to generate fee levels is state of the art, but not perfect; for example, status quo bias might distort the fee levels downward (for example, if too many pilot respondents reported zero willingness to pay), or conversely hypothetical bias might distort the fee levels upward (for example, if pilot respondents reported extravagantly high willingness to pay, knowing they would not actually need to pay these amounts).

To the extent that the fee levels were too low or too high, our experiment simply would have been underpowered, and we would have been unable to recover meaningful results. For example, if our prices were too high and respondents almost always picked the status quo option, our implicit prices would be very sensitive and our confidence intervals would be excessively wide. The fact that this was not the case suggests that the fee levels we selected were appropriate.

4. Fractional Factorial Design

Each choice card set included two hypothetical services with varying attributes as well as the option to choose the status quo. Because there were 4 attributes with 4, 2, 2, and 6 levels,

respectively, there were $(4 \cdot 2 \cdot 2 \cdot 6)^2 = 9,216$ possible choice card sets. A design that presented all of these choice card sets (known as a "full factorial" design) would require an impracticably large number of respondents. Instead, we employed a "fractional factorial" design²¹² that presented a subset of the possible choice card sets. Doing so reduced the number of respondents needed and maximized statistical power.

To select the subset of choice card sets, we employed Expert Choice, a package developed for R statistical software.²¹³ Expert Choice produced a set of 21 choice card sets (42 individual choice cards) that was orthogonal and balanced with respect to each attribute, with total d-efficiency of 91.558%.²¹⁴ We divided the 21 choice card sets into three blocks of seven each, randomly assigning one block per respondent. We balanced the blocks presented to different respondents to equalize the number of each attribute presented.

5. Power Analysis and Sample Size

Among studies that conduct sample size analysis, a plurality apply simple rules of thumb to determine the appropriate sample size. Because of the expense and complexity of collecting survey responses, discrete choice experiments typically feature smaller sample sizes than other kinds of economic study; one

^{212.} In health economics, "fractional factorial designs are typically used to reduce the number of scenarios to manageable numbers that can be implemented in surveys." Rosalie Viney et al., *Discrete Choice Experiments to Measure Consumer Preferences for Health and Healthcare*, 2 EXPERT REV. PHARMACOECONOMICS & OUTCOMES RSCH. 319, 323 (2002).

^{213.} Jed Stephens, *Theoretical Introduction to ExpertChoice*, THE COMPRE-HENSIVE R ARCHIVE NETWORK (Mar. 31, 2020), https://cran.r -project.org/web/packages/ExpertChoice/vignettes/include_theory.pdf [https:// perma.cc/B56A-4ZAW]; Jed Stephens, *Expert Choice*, GITHUB (Apr. 7, 2020), https://github.com/JedStephens/ExpertChoice [https://perma.cc/L6KK-MYMP].

^{214.} We utilized the Expert Choice setting that applies the widely used Street/Burgess fractional factorial design. See generally F. Reed Johnson et al., Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force, 16 VALUE HEALTH 3 (2013) (discussing the Street/Burgess approach, among others). Fractional factorial designs aim to achieve orthogonality and level balance. Emily Lancsar & Jordan Louviere, Conducting Discrete Choice Experiments to Inform Healthcare Decision Making: A User's Guide, 26 PHARMACOECONOMICS 661, 669 (2008).

review of discrete choice experiments found that most involved fewer than 300 respondents.²¹⁵

A common rule of thumb suggested by Johnson and Orme is that, where N is the sample size, c is the number of analysis cells (equal to the largest number of levels for any one attribute when main effects are being analyzed, as here), t is the number of choice tasks per respondent, and a is the number of alternatives per choice set:

$$N > 500c/(t \cdot a) = 500 \cdot 6/(7 \cdot 3) = 142.9$$

That is, the Johnson-Orme sample sizing method suggested that we require a minimum of 143 respondents. In fact, this figure may be an overestimate, because it assumes a naïve full factorial design, whereas we use a fractional factorial design that maximizes statistical power for any given number of respondents. Because our sample size, even after all attrition, was 244 respondents, or 1,708 discrete choices, we substantially exceeded the sample size required for adequate power.

6. Qualitative Questions

To provide qualitative context for respondents' choices, we asked respondents open-ended questions about reasons for their willingness or unwillingness to pay for time savings or risk reduction. Each respondent received one of two possible questions plus a follow-up question. The question distribution was randomized. The questions were:

In general, do you think it's worthwhile to pay money to reduce time spent on taxes? [yes/no]

In general, do you think it's worthwhile to pay money to eliminate risk of error and/or audit associated with taxes? [yes/no]

Depending on their answer to the above question, they received one version of the following question:

^{215.} Esther W. de Bekker-Grob et al., Sample Size Requirements for Discrete-Choice Experiments in Healthcare: A Practical Guide, 8 PATIENT 373, 375 (2015).

Please give one reason why you think it [is/is not] worthwhile to pay money to [reduce time spent on taxes/eliminate risk of error and/or audit associated with taxes].

Subsection II.B.2 discusses the qualitative responses in more detail. Word clouds summarizing the most common phrases in the qualitative responses are available in the Online Appendix.²¹⁶ We applied other natural language processing techniques (for example, those proposed by Ferrario and Stantcheva²¹⁷) but ultimately decided to classify the responses manually, since they were insufficiently long for sophisticated text analysis to be worthwhile.

B. ATTRITION AND REJECTED RESPONDENTS

Like any survey, our discrete choice experiment relied on attentive, high-quality answers from respondents who understand the concepts being tested. We therefore rejected respondents (and dropped them from our analysis) under the following circumstances, even if the respondent completed the entire survey²¹⁸: (1) failure to pass a comprehension check, (2) failure to pass an attention check (below), and (3) protest responses, as evidenced by qualitative responses.

We reviewed each answer individually and would also have removed any nonsensical or objectively low-effort responses, as well as any responses that were completed implausibly quickly. However, after applying the exclusion criteria above, no remaining responses were sufficiently nonsensical, low-effort, or fast to warrant exclusion.

1. Comprehension Check

Immediately after reading the disclosure about the range of choices available to them in the choice sets, respondents received the following comprehension check:

^{216.} The Online Appendix is available at http://minnesotalawreview.org/vol108-choi-jurow-kleiman-appendix.

^{217.} Beatrice Ferrario & Stefanie Stantcheva, *Eliciting People's First-Order Concerns: Text Analysis of Open-Ended Survey Questions*, 112 AM. ECON. ASS'N PAPERS & PROC. 163 (2022).

^{218.} Respondents who refused to give consent to participate in the survey or who refused to commit to provide "honest, considered responses to each question" were not allowed to complete the survey.
Imagine that you select a service that eliminates **all time** spent on tax compliance activities. In the year that you use this service, how much time will you spend preparing and filing your tax return (including gathering W-2s/1099s, hiring a tax preparer, checking your return, etc.)?

- A. The same time I currently spend
- B. A bit less time than I currently spend
- C. A lot less time than I currently spend
- D. I would not spend any time or effort working on my taxes

The correct answer was D, and any respondents who did not provide this answer were excluded from the sample. 62.5% of respondents who completed the survey correctly answered the comprehension check on the first try. Those who answered incorrectly were permitted to try answering again, and 92.6% of respondents correctly answered the comprehension check on the first or second try. However, respondents who answered incorrectly on the first try and correctly on the second try were still excluded from the sample.

2. Attention Check

As an attention check, the final choice set presented to each respondent substituted for the "Risk Elimination" attribute a line saying "Choice Not Available" and instructing the respondent to "please select the 'Neither' option."²¹⁹ This choice card is presented below²²⁰:

^{219.} For a similar attention check, see Alexandre Mas & Amanda Pallais, *Valuing Alternative Work Arrangements*, 107 AM. ECON. REV. 3722, 3731–32 (2017).

^{220.} The choice cards were generated using Python code with HTML formatting, adapting code from Sylvain Weber, A Step-by-Step Procedure to Implement Discrete Choice Experiments in Qualtrics, 39 SOC. SCI. COMPUT. REV. 903 (2021).

Figure 4: Example of Attention Check Choice Card Which of the following tax services would you prefer?

	Service 1	Service 2	Neither Service
Time Savings	Reduce by 1 Hour (spend 9 hours instead of 10 hours)	Reduce by Half (spend 5 hours instead of 10 hours)	
Risk Elimination	Choice Not Available (please select the "Neither Service" option)	Choice Not Available (please select the "Neither Service" option)	No Change (continue to spend 10 hours, no change in accuracy
Provider	Private Company	Government	or audit risk)
Fee	\$15	\$25	
Your choice:	Service 1	Service 2	Neither

76.4% of respondents who completed the survey correctly answered the attention check.

3. Protest Responses

We excluded any respondents whose qualitative responses indicated that they rejected the frame of the survey experiment.²²¹ The following are examples of qualitative responses that were flagged as indicating protests:

I don't believe it's possible [to eliminate the risk of error].

^{221.} See IAN J. BATEMAN ET AL., ECONOMIC VALUATION WITH STATED PREF-ERENCE TECHNIQUES: A MANUAL 304–11 (2002) (discussing strategies for identifying protest responses through open-ended questions).

At the very core, I do not think that we should have to pay to have our taxes easily figured out. Easy means of doing so should be provided by the government if they expect us to calculate our own taxes.

However, qualitative responses were not considered protest responses if they simply indicated that the respondent did not find the service in question worthwhile. Protest responses were rare: Only 19 respondents out of 475 (4%) were coded as protesters.

4. Attrition Statistics

Because our procedures to reject responses were relatively stringent, we rejected almost half of the respondents who completed our survey. As Table 5 shows, our final sample had 244 responses out of 475 completed responses.

Statis- tic	All	Pre- Screened, Started Survey	Com- pleted Sur- vey	Passed Com- prehen- sion Check	Passed Atten- tion Check	Non- Pro- tester
Age	37.376	38.113	38.217	37.795	38.135	38.029
Female	0.408	0.383	0.387	0.397	0.402	0.399
Male	0.589	0.615	0.611	0.599	0.594	0.597
White	0.621	0.660	0.672	0.710	0.713	0.708
Black	0.145	0.119	0.109	0.091	0.100	0.103
His-	0.135	0.125	0.118	0.091	0.084	0.086
panic						
Asian	0.053	0.057	0.059	0.057	0.048	0.045
Mixed	0.031	0.028	0.029	0.037	0.040	0.041
N	1,000	506	475	297	251	244

Table 5: Attrition Statistics

C. DISCRETE CHOICE MODEL

To model respondents' discrete choices regarding whether to purchase the tax services offered to them, we specified the following indirect utility function U_{ij} , given respondent *i*, choice *j*, vector V_{ij} reflecting an index of observable characteristics of *i* and *j*, and stochastic scalar ϵ_{ij} reflecting unobservable characteristics of *i* and *j*:

$$U_{ij} = V_{ij} + \epsilon_{ij}$$

We generated a linear random utility model for U_{ij} by specifying V_{ij} as follows, where ASC is an alternative specific constant (a dummy variable that is 1 for the status quo option and 0 for the other options; the ASC controls for any status quo bias),²²² Fee is the cost in dollars of the package of services the individual buys to reduce tax compliance burdens, and Risk, Time₁, Time_{half}, Time_{all}, and Government are indicator variables corresponding with the various attributes:

$$V_{ij} = \beta_0 \cdot ASC + \beta_1 \cdot Risk_j + \beta_2 \cdot Time_{1j} + \beta_3 \cdot Time_{halfj} + \beta_4 \cdot Time_{allj} + \beta_5 \cdot Government_j + \beta_f \cdot Fee_j$$

Note that each of the models in this Article is a utility model, which explains the absence of error terms as might be present in simple linear regression. We presented each respondent with three discrete choices at once: two alternative packages of tax services, or the status quo (no service). We used the conditional logit model described in McFadden (1973) to predict the probability that respondent *i* selects any alternative *g* (including the status quo) as²²³:

$$P(U_{ig} > U_{ih}, \forall h \neq g) = \frac{exp(V_{ig})}{\sum_{j} exp(V_{ij})}$$

The coefficients in the multinomial conditional logit model were determined via softmax optimization to maximize goodness of fit given each respondent's choice among the three discrete choices offered. Note that we could not include constants, individual-level variables, or choice-set-level variables in this optimization, because they would not affect any respondent's choice of one option with respect to the others. Mathematically, because we solved the objective function by differentiating with respect

^{222.} Jürgen Meyerhoff & Ulf Liebe, Status Quo Effect in Choice Experiments: Empirical Evidence on Attitudes and Choice Task Complexity, 85 LAND ECON. 515 (2009).

^{223.} See Zsolt Sándor, Computation, Efficiency and Endogeneity in Discrete Choice Models 8, equation 2.5 (Nov. 22, 2001) (Ph.D. thesis, University of Groningen), https://pure.rug.nl/ws/portalfiles/portal/3083725/thesis.pdf [https://perma.cc/BJF3-G6SU] (showing the standard logit model specification); Atkinson et al., *supra* note 76, at 131 (describing an identical formula, assuming that the scale parameter μ is set to 1).

to the coefficients, the coefficient of any variable constant across the choice set would "disappear in the differentiation."²²⁴

We conducted our analysis in R using the mlogit package, a widely used tool for multinomial logit modeling. 225

Given this model, we calculated the implicit price of any given attribute with coefficient β_a , given β_f as the coefficient from the *Fee* variable in our regressions, using the formula:

Implicit price =
$$-\frac{\beta_a}{\beta_f}$$

All confidence intervals in this Article were calculated through bootstrapping. First, we randomly drew individuals from our sample (with replacement) to re-construct 1,000 samples of identical size. (Note that this bootstrapping did assume between-individual independence, analogous to clustering in calculating parametric standard errors.) We conducted the same multinomial logistic regression to estimate implicit prices using each reconstructed sample. We then generated "empirical" bootstrap intervals by taking the 2.5th-percentile and 97.5th-percentile estimates.²²⁶ Because bootstrapping is a non-parametric method to calculate standard errors, it avoids potential statistical problems that could arise using standard parametric techniques.²²⁷

D. REGRESSION TABLE

Table 6 below summarizes the results of the main effects regression.

^{224.} Data Management, Model Description and Testing, THE COMPREHEN-SIVE R ARCHIVE NETWORK, https://cran.r-project.org/web/packages/mlogit/ vignettes/c2.formula.data.html [https://perma.cc/54EM-QL2M].

^{225.} Yves Croissant, *mlogit: Multinomial Logit Models*, THE COMPREHEN-SIVE R ARCHIVE NETWORK, https://cran.r-project.org/web/packages/mlogit/index .html [https://perma.cc/4WV4-8ZVF].

^{226.} See A. C. DAVISON & D. V. HINKLEY, BOOTSTRAP METHODS AND THEIR APPLICATION 194 (1997) (showing how the bootstrap method is applied).

Attribute	Coefficient
Risk	2.154***
	(0.110)
$Time_1$	-0.263*
	(0.149)
$Time_{half}$	0.302**
	(0.146)
$Time_{all}$	1.592***
	(0.141)
Government	0.055
	(0.072)
Fee	-0.030***
	(0.002)
$ASC_{status\ quo}$	0.263
	(0.195)
$ASC_{option \ 2}$	0.081
	(0.075)
Observations	1694
Log Likelihood	-1308.060
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 6: Regression Table - Main Effects

E. CONFIRMING INDEPENDENCE OF IRRELEVANT ALTERNATIVES

One important assumption underlying the discrete choice model is the Independence of Irrelevant Alternatives (IIA): the model assumes that each attribute has a particular effect on the likelihood of selecting a given option (as reflected by its regression coefficient), which is independent of the other attributes included on the same choice card. For example, the fact that a choice card included risk elimination should not change the effect of time reduction on a respondent's likelihood of selecting a particular choice card.

The standard method to test whether the IIA assumption holds is the Hausman-McFadden test. To conduct a Hausman-McFadden test, we re-estimate the model on a subset of the

attributes in the full model, excluding irrelevant attributes. If the IIA assumption was correct, the subset regression would be less efficient, but the parameter estimates obtained should be consistent with the parameter estimates in the full regression. The Hausman-McFadden test produces a χ^2 statistic that can be translated into a *p*-value, estimating the probability that the null hypothesis (that the IIA assumption holds) is true.²²⁸

We conducted a series of Hausman-McFadden tests using subsets that exclude the risk reduction attribute, the time reduction attributes, and the government/private company attribute, respectively, to test whether the remaining attributes are independent of these attributes. The tests produced the following results:

> Table 7: Hausman-McFadden Test – Subset Excluding Risk

X2	-554.89
degrees of freedom	7
<i>p</i> -value	1

Table 8: Hausman-McFadden Test – Subset Excluding Time

X2	-61.613
degrees of freedom	5
<i>p</i> -value	1

Table 9: Hausman-McFadden Test – Subset Excluding Government

X2	-0.57904
degrees of freedom	7
<i>p</i> -value	1

Each of the tests produces a *p*-value of 1. Based on this *p*-value, we failed to reject the null hypothesis, suggesting that the IIA assumption holds. Each of the tests also produced a negative

2024]

^{228.} See generally Jerry Hausman & Daniel McFadden, Specification Tests for the Multinomial Logit Model, 52 ECONOMETRICA 1219 (1984) (detailing the Hausman-McFadden test).

 χ^2 value, which (according to Hausman and McFadden) is evidence that the IIA assumption has not been violated. 229

229. Id. at 1226.