

SUBJECTIVE COSTS OF TAX COMPLIANCE

*Jonathan H. Choi and Ariel Jurow Kleiman**

Abstract

This Article introduces and estimates “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 by 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, eliminating tax compliance time for high-income taxpayers and taxpayers with complex returns is not worth much more than eliminating tax compliance time for low-income taxpayers with simple returns.

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, focus on simplifications that reduce mistakes rather than merely saving time, and prioritize reforms that affect low-income taxpayers.

* Jonathan H. Choi is an Associate Professor of Law at the University of Minnesota School of Law; Ariel Jurow Kleiman is an Associate Professor of Law at Loyola Law School, Los Angeles. Acknowledgments: Trudy Cameron, June Carbone, Chris Evans, Jacob Goldin, Kristin Hickman, Claire Hill, Zachary Liscow, Joe Mitchell-Nelson, Jeessoo Nam.

Table of Contents

<i>Introduction</i>	3
<i>I. What We Do, Don't, and Should Know</i>	8
A. What We Know: Objective Compliance Costs	8
B. What We Don't Know: Subjective Compliance Costs.....	10
1. Empirical Research on Subjective Costs	11
2. Potential Subjective Benefits of Taxation.....	13
C. What We Should Know: Monetizing Subjective Costs.....	14
<i>II. Survey and Results</i>	17
A. Methodology	17
1. Discrete Choice Surveys Generally	17
2. Survey Details and Implicit Prices.....	19
3. Survey Sample	21
B. Results	23
1. Discrete Choice Experiment Results.....	23
2. Monetizing the Benefits of Tax Reform	30
3. Qualitative Context	35
C. Potential Objections and Robustness Checks	40
1. External Validity	40
2. Cheap Talk	42
3. Statistical Validity and Survey Fatigue.....	44
<i>III. Implications for Theory and Policy</i>	45
A. Implications for Theory	45
B. Implications for Policy	46
1. Think Big	47
2. Minimize Mistakes.....	49
3. Focus on Simple, Low-Income Returns.....	51
4. Public or Private, Taxpayers Don't Care	53
<i>Conclusion</i>	55
<i>Appendix</i>	56
A. Survey Contents	56
1. Respondent-Specific Data.....	56
2. Disclosure of Possible Choice Sets.....	57
3. Choice Card Sets.....	57
4. Fractional Factorial Design.....	58
5. Power Analysis and Sample Size.....	58
6. Qualitative Questions.....	59
B. Attrition and Rejected Respondents	60
1. Comprehension Check	60
2. Attention Check	61
3. Protest Responses.....	62
4. Attrition Statistics	62
C. Discrete Choice Model	63
D. Regression Table	65
E. Confirming Independence of Irrelevant Alternatives	66

INTRODUCTION

The income tax return inspires aggravation, confusion, boredom, and even anger.¹ Tax season is a time of hand-wringing foot dragging that climaxes in a flurry of online forms submitted at 11:59pm on April 15th.² Scholarship on tax compliance costs largely echoes this narrative. Commentators 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 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.

Congress has recently turned to the task of reducing the compliance costs of taxation for ordinary Americans. The Inflation Reduction Act of 2022

¹ See, e.g., CBS, *Confusion Reigns Supreme as Americans Wait Until the Last Minute to File Their Tax Returns*, CBS NEW YORK (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> (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?*, THE CONVERSATION (Mar. 22, 2021), <https://theconversation.com/why-cant-the-irs-just-send-americans-a-refund-or-a-bill-156733> (describing tax filing as “unpleasant,” “onerous,” and “tedious”); Ryan Lasker, *Why Are Taxes So Confusing?*, MONEY SCOOP (Jan. 3, 2022), <https://www.morningbrew.com/money-scoop/stories/2022/01/03/why-are-taxes-so-confusing> (last visited Dec. 14, 2022); Brooke Gladstone, *Why So Tedious, Taxes?*, WNYC (Apr. 6, 2017), <https://www.wnycstudios.org/podcasts/otm/segments/taxes-tedious>.

² See The Simpsons, *Homer on Tax Day*, YOUTUBE, <https://www.youtube.com/watch?v=ZnJcZ-5P8hE> (last visited Dec. 14, 2022).

³ E.g., Joseph Bankman, *Who Should Bear Tax Compliance Costs?* 2 (John M. Olin Program in Law and Economics Working Paper 279, 2004), <http://ssrn.com/abstract=519783> (“Compliance costs substantially reduce the social gains from taxation”); Scott A. Hodge, *The Tax Compliance Costs of IRS Regulation*, TAX FOUND. (Aug. 23, 2022), <https://taxfoundation.org/tax-compliance-costs-irs-regulations> (“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. (2013), <https://www.mercatus.org/research/research-papers/hidden-costs-tax-compliance> (tallying and critiquing the “hidden costs” of taxes, which include compliance costs).

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*, 65 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 (2006) (finding U.S. tax morale to be high relative to peer European nations); Org. Econ. Coop. & Dev., *Annex A: Data Tables*, in TAX ADMINISTRATION 2021: COMPARATIVE INFORMATION ON OECD AND OTHER ADVANCED AND EMERGING ECONOMIES 199 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%)).

directs the Internal Revenue Service (IRS) to research the possibility of creating a government-run “direct e-file” system.⁵ The Act allocates \$15 million to the research effort,⁶ which is surely a miniscule fraction of the total eventual cost of adopting a public tax-filing system. Depending on its design, a direct e-file system could dramatically reduce the time and stress associated with tax compliance by pre-populating tax returns with known data, rather than forcing taxpayers to submit that information themselves.⁷

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,⁸ we know almost nothing about how unpleasant or burdensome people find these activities to be.⁹

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.¹⁰ By asking taxpayers whether they would be willing to pay to reduce various tax compliance burdens, we can impute a monetary value to those burdens.

Based on existing literature, we hypothesize that people might in particular find tax compliance activities to be aggravating, tedious, and generally unpleasant. Separately, people might worry about making mistakes on their return or being audited. Therefore, we asked respondents whether they would be willing to purchase tax services that would reduce or eliminate

⁵ Inflation Reduction Act of 2022, P.L. 117-169 § 10301(1)(B).

⁶ *Id.*

⁷ See *infra* Part III.B.1 for further discussion of how such a system might work and how it would reduce tax compliance costs.

⁸ *E.g.*, Joshua D. McCarthy, *The Cost of Tax Compliance*, TAX FOUND. (Sept. 11, 2014), <https://doi.org/10.1787/cef472b9-en> (providing tax compliance time and expenditure estimates).

⁹ 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”).

¹⁰ 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-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.

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.¹¹ Prominent simplification reforms often seek to reduce the time people spend on their return—for example, by reducing paperwork or calculations necessary to claim itemized deductions via increasing the standard deduction.¹² But we found that taxpayers are more bothered by possible mistakes than by any 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 eliminate all the time they spend on tax compliance activities (about \$53, on average).

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.¹³ They ascribed only a slightly higher value—but still not much—to a service that would cut their tax compliance time in half.¹⁴ Respondents only seemed to meaningfully value a service that would eliminate *all* tax compliance work, with a willingness to pay of \$53 total, which comes out to about \$10.40 per hour. When it comes to saving time on their taxes, it's an all or nothing calculus for taxpayers.

By aggregating individual willingness to pay over the entire population, we were also able to estimate the dollar-value benefits of large-scale government reforms.¹⁵ We estimate, for example, that eliminating *all* of the time that people spend on tax compliance activities across the entire population would generate benefits equal to \$8.68 billion.¹⁶ Put differently, all taxpayers in the U.S., taken together, would be willing to pay \$8.68 billion

¹¹ See Hodge, *supra* note 3; Gladstone, *supra* note 1.

¹² See, e.g., ERICA YORK & ALEX MURESIANU, TAX FOUND., THE TAX CUTS AND JOBS ACT SIMPLIFIED THE TAX FILING PROCESS FOR MILLIONS OF HOUSEHOLDS 4 (2018), <https://taxfoundation.org/the-tax-cuts-and-jobs-act-simplified-the-tax-filing-process-for-millions-of-americans> (explaining how increasing the standard deduction simplifies tax filing).

¹³ See *infra* Section II.B.1.b.

¹⁴ *Id.*

¹⁵ To arrive at these estimates, we calculated implicit prices for identifiable subgroups and then multiplied these prices by the number of taxpayers in each subgroup across the whole population based on IRS data. For a more detailed explanation, see *infra* Section II.B.2.

¹⁶ The figures in this paragraph, and *infra* Part II.B.2, ignore taxpayers' current out-of-pocket spending on tax preparation services. See *infra* note 117.

annually to eliminate the time they currently spend filing their tax returns. Additionally, they would be willing to pay \$11.98 billion annually to eliminate their risk of making a mistake or being audited, reflecting the relatively greater disutility caused by anxiety over aggravation.

Conducting aggregate calculations for specific subgroups led to our third main finding, which is that the benefits of tax simplification and automation do not vary substantially by income or by the complexity of taxpayers' returns. In other words, eliminating tax compliance time for high-income taxpayers and taxpayers with complex returns is not worth much more than eliminating tax compliance time for low-income taxpayers with simple returns.¹⁷

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.¹⁸ As mentioned above, researchers have a good sense of the amount of time that people spend on tax compliance.¹⁹ To monetize this time, analysts typically multiply it by some ascribed hourly wage.²⁰ This method is used to value people's time for nearly all cost-benefit analysis across government agencies.²¹ Our findings suggest that these market-wage based estimates do *not* reflect the welfare burden that people experience.

Our findings have important policy implications as well. Given Congress's recent directive for the IRS to research creating a direct e-file system, perhaps the most important takeaway is that policymakers should think big when it comes to reducing tax compliance costs. We found that taxpayers ascribe little value to marginal time savings; thus policymakers

¹⁷ As explained below, we generated this result by conducting a regression with interaction terms for a wide variety of respondent characteristics. We then multiplied per-person willingness to pay by the number of people with each set of characteristics, according to published IRS data, to estimate the benefits of tax reforms for specific subgroups, as well as for the entire population. *See infra*, Part II.B.2, for additional explanation.

¹⁸ *See generally* ERICA YORK & ALEX MURESIANU, TAX FOUND., REVIEWING DIFFERENT METHODS OF CALCULATING TAX COMPLIANCE COSTS (2018), <https://taxfoundation.org/different-methods-calculating-tax-compliance-costs> (describing different hourly rates used to monetize time costs).

¹⁹ *See* McCarthy, *supra* note 8.

²⁰ *Id.*

²¹ 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 & Douglas Shaw, *Estimating the Cost of Leisure Time for Recreation Demand Models*, 38 J. ENV. ECON. & MGMT. 49 (1999); Kenneth E. McConnell & Ivar Strand, *Measuring Cost of Time in Recreation Demand Analysis: An Application to Sportfishing*, 63 AM. J. AGRIC. ECON. 153 (1981) (arguing that time cost should be valued at some fraction of the wage rate less than one).

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.

Speaking directly to the question of a government-run e-file program: The IRS should adopt the most expansive version of the program, one that includes the maximum amount of taxpayer information and requires the least amount of taxpayer input for each individual taxpayer. A partially populated government return would still require a significant investment of resources by the IRS, while also still requiring taxpayers to spend a fair amount of time inputting information. The resulting social welfare improvement will be comparatively small. Instead, the IRS will get more bang for its buck by providing taxpayers a fully prepared tax return.

Our survey results also suggest that policymakers should focus on reforms that reduce concerns about inadvertent mistakes in tax filings. Congress could, for instance, simplify complex rules and eliminate “traps for the unwary” in the tax code.²² The IRS could reduce taxpayer anxiety by better publicizing remediation programs for taxpayers who make a (good faith) mistake on their tax returns as well as programs for those who can’t afford to pay their taxes right away.²³

Finally, policymakers should design simplification reforms to target low-income taxpayers with simple return. We find that taxpayers’ willingness to pay to reduce aggravation and anxiety is relatively constant across income levels.²⁴ Given declining marginal utility of income—the idea that additional income is worth more to a poor person than a rich one²⁵—this finding suggests that the actual disutility of aggravation and anxiety may be higher for low-income individuals. Moreover, while the monetary benefits are relatively constant, the *administrative costs* of filing simplification are likely much lower for low-income taxpayers with simple returns. Therefore, policymakers ought to focus resources on reducing subjective costs of tax compliance for low-income taxpayers with simple returns. Prioritizing return-free filing for these taxpayers as well as simplifying eligibility rules

²² See, e.g., ROBERT GREENSTEIN, JOHN WANCHECK & CHUCK MARR, CTR. FOR BUDGET & POL’Y PRIORITIES, REDUCING OVERPAYMENTS IN THE EARNED INCOME TAX CREDIT 1 (2019) (“EITC errors occur primarily because of the complexity of the rules surrounding the credit. Most of them reflect unintentional errors, not fraud.”).

²³ See *infra* notes 185-193 and accompanying text.

²⁴ See *infra* Section II.B.2.

²⁵ See Joseph Bankman & Thomas Griffith, *Social Welfare and the Rate Structure: A New Look at Progressive Taxation*, 75 CAL. L. REV. 1905, 1946-48 (discussing declining marginal utility).

for the earned income tax credit (EITC) and Child Tax Credit would both serve this goal.²⁶

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 “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 implications and potential objections to our methods and findings. Part III offers implications for theory and policy, including how our survey results should inform the design of a possible direct e-file program.

I. WHAT WE DO, DON’T, AND SHOULD KNOW

This Part describes the current state of the research on 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. These estimates 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.²⁷ Compliance costs are one of the three main costs that the tax system imposes on taxpayers, alongside the taxes themselves and efficiency costs incurred when taxpayers change their behavior to avoid taxes.²⁸ Across tax types, methodologies, and countries, researchers have

²⁶ See *infra* Part III.B.3 for additional discussion.

²⁷ The Paperwork Reduction Act of 1980 (PRA) requires the IRS to track the time that taxpayers spent on “recordkeeping,” “tax planning,” “form completion and submission,” and “all other,” as well as how much money they spend to have third parties perform these tasks. I.R.S., INSTRUCTIONS TO FORM 1040 108 (2022), <https://www.irs.gov/pub/irs-pdf/i1040gi.pdf>.

²⁸ See Chris Evans, *Taxation Compliance and Administrative Costs: An Overview*, in TAX COMPLIANCE COSTS FOR COMPANIES IN AN ENLARGED EUROPEAN COMMUNITY 449

found that compliance costs are “high and significant.”²⁹

As the work “cost” implies, measuring compliance costs requires monetizing the time and effort expended in complying with tax laws.³⁰ Thus, the vast majority of the research in this area focuses on ascribing a monetary value (sometimes expressed as a percentage of GDP) to tax compliance activities,³¹ including time spent as well as money spent on software, tax preparation assistance, and tax expertise.³² Tabulating expenditures is straightforward; monetizing time is somewhat more complicated.³³ To convert time spent on tax compliance into a dollar amount, policy analysts most often multiply the hours spent by some market-determined hourly rate—typically the average hourly wage for all workers or for professional workers.³⁴ 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 total objective costs of tax compliance.³⁵ Current estimates suggest that these

(Michael Lang et al. ed., 2008).

²⁹ *Id.* at 455 (surveying tax compliance research).

³⁰ *Id.*; Off. Mgmt. & Budget (OMB), *Estimating Paperwork Burden* (1999), https://obamawhitehouse.archives.gov/omb/fedreg_5cfr1320 (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”).

³¹ Evans, *supra* note 28, at 456-57.

³² *Id.* at 451 (referring to these as the “hard core” of compliance costs).

³³ *See, e.g.*, OMB, *supra* note 30 (“Monetizing burden hours would present a daunting methodological challenge and raises issues concerning certainty and ease of administration by agencies.”); Samaha, *supra* note 21, at 328-36 (challenging the government’s prevailing method of monetizing time).

³⁴ *E.g.*, YORK & MURESIANU, *supra* note 18, at 3 (describing different hourly rates to monetize time costs and using Bureau of Labor Statistics averages for full-time private sector workers (\$37.28) for individual income taxes and professional workers (\$52.05) for business income taxes); Samaha, *supra* note 21, 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. SERV., FOOD & DRUG ADMIN., TOBACCO PRODUCT STANDARD FOR N-NITROSONORNICOTINE LEVEL IN FINISHED SMOKELESS TOBACCO PRODUCTS 78 (2017), <https://www.fda.gov/about-fda/economic-impact-analyses-fda-regulations/tobacco-product-standard-n-nitrosonornicotine-level-finished-smokeless-tobacco-products-proposed> (“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 the value of individuals’ “shadow time”).

³⁵ *See, e.g.*, ARTHUR B. LAFFER, WAYNE H. WINEGARDEN & JOHN CHILDS, *THE ECONOMIC BURDEN CAUSED BY TAX CODE COMPLEXITY* 20 (2016) (monetizing individual taxpayer time at \$68.42 per hour and business taxpayer time at \$55 per hour); J. SCOTT MOODY, WENDY P. WARCHOLIK & SCOTT A. HODGE, *THE RISING COST OF COMPLYING WITH*

objective costs are huge, annually totaling hundreds of billions of dollars.³⁶

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.

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.³⁷ 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 psychological costs of taxation), it often does so via qualitative data that are difficult to compare with the quantitative compliance cost measures just described.³⁸

Many experts believe that conventional calculations of objective costs underestimate the true cost of tax compliance because they exclude important *psychological* costs of taxation.³⁹ These scholars argue that tax compliance is uniquely unpleasant along two dimensions. First, tax compliance causes *aggravation* because it’s tedious, difficult, and stressful.⁴⁰ Second, tax compliance causes *anxiety* because taxpayers fear repercussions from making

THE FEDERAL INCOME TAX 9 (2005) (monetizing individual taxpayer time at \$39.18 per hour and business taxpayer time at \$47.96 per hour); JOEL SLEMROD, WRITTEN TESTIMONY SUBMITTED TO THE COMMITTEE ON WAYS AND MEANS, SUBCOMMITTEE ON OVERSIGHT, HEARING ON TAX SIMPLIFICATION (2004) (monetizing taxpayer time at \$20 per hour).

³⁶ Hodge, *supra* note 3, at 1 (estimating total compliance costs at \$409 billion); Fichtner & Feldman, *supra* note 3 (summarizing empirical estimates of compliance costs at between \$67 billion and \$378 billion).

³⁷ See Bankman, *supra* note 3, at 2 (noting this gap in the research).

³⁸ See *infra* Part I.B.1.

³⁹ Donald Moynihan, Pamela Herd & Hope Harvey, *Administrative Burden: Learning, Psychological, and Compliance Costs in Citizen-State Interactions*, 25 J. PUB. ADMIN. RSCH & THEORY 43, 46 (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, 1852 (2019) (arguing that administrative compliance burdens, so-called “sludge,” imposes psychological costs in the form of “frustration, stigma, and humiliation”).

⁴⁰ See Robin Woellner, Cynthia Coleman, Margaret McKerchar, Michael Walpole & Julie Zetler, *Taxation or Vexation – Measuring the Psychological Costs of Tax Compliance*, in TAX COMPLIANCE COSTS: A Festschrift for Cedric Sandford 37 (Chris Evans, Jeff Pope, & John Hasseldine eds., 2001) (defining psychological costs to include “anxiety and frustration caused by complying with complicated revenue legislation”).

a mistake.⁴¹ Researchers agree that psychological costs exist and are significant.⁴² Even so, researchers typically ignore such costs in compliance costs estimates because they are difficult to measure and compare across individuals.⁴³

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.

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.

1. Empirical Research on Subjective Costs

Much of the compliance costs literature simply assumes that subjective costs impose large burdens on taxpayers.⁴⁴ This assumption is partly based

⁴¹ See *id.*; Evans, *supra* note 28, at 451-52. A related but distinct literature explores the psychological determinants of evading tax laws. See generally, Ken Devos, *Tax Compliance Theory and the Literature*, in FACTORS INFLUENCING INDIVIDUAL TAXPAYER COMPLIANCE BEHAVIOUR (2014).

⁴² See, e.g., Bankman, *supra* note 3, at 2; Binh Tran-Nam & John Glover, *Estimating the Transitional Compliance Costs of the GST in Australia: A Case Study Approach*, 17 AUSTL. TAX F. 499, 518 (2002).

⁴³ See, e.g., Binh Tran-Nam, Chris Evans & Phil Lignier, *Personal Taxpayer Compliance Costs: Recent Evidence from Australia*, 29 AUS. TAX FOR. 135, 141 (2014) (noting only one study that aims to measure the psychological costs of tax compliance); John L. Guyton, John F. O'Hare, Michael Stavrianos & Eric J. Toder, *Estimating the Compliance Costs of the U.S. Individual Income Tax*, 3 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 (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).

⁴⁴ Eighteenth-century economist Adam Smith noted the likely difference between subjective and objective costs. ADAM SMITH, AN INQUIRY INTO THE NATURE AND CAUSES

on the intuition that complexity generates frustration and the fact that inputs to the tax filing process are inordinately complicated.⁴⁵

Some limited research gathers qualitative data about 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 collected between 1990-2013, approximately half to two-thirds of respondents reported that they “dislike” or “hate” doing their taxes.⁴⁶ In 2013, just under one-third of respondents stated that they dislike tax filing because it is “complicated” and involves “too much paperwork.”⁴⁷

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.⁴⁸ Apart from confirming the existence of some 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.⁴⁹ For instance, one study from Australia found that the stress of tax compliance is higher during periods of policy change and that “worry” is a major cause of such stress.⁵⁰ Another reported that survey participants

OF THE WEALTH OF NATIONS 678 (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.”).

⁴⁵ 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* AARP FOUND. TAXAIDE PROGRAM, QUALIFYING CHILD AND QUALIFYING RELATIVE FLOW CHART (2014), <http://nytaxaide.org/wp-content/uploads/2013/06/qualifyingchildflowchart11-11-14.pdf>. 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.*; IRC § 152(c)(4)(C).

⁴⁶ AM. ENTER. INST., PUBLIC OPINION ON TAXES 79 (2009), <https://www.aei.org/wp-content/uploads/2011/10/AEI-Public-Opinion-Studies-Taxes-2009.pdf?x91208>; Pew Rsch. Ctr., *A Third of Americans Say They Like Doing Their Income Taxes*, (Apr. 11, 2013), <https://www.pewresearch.org/politics/2013/04/11/a-third-of-americans-say-they-like-doing-their-income-taxes>.

⁴⁷ *Id.* In addition, a large portion of this aversion can be attributed to the tax payment itself. *Id.*

⁴⁸ *See* Evans, *supra* note 28, at 451 (stating, as of 2008, that “no studies have yet managed to successfully quantify these psychological costs”); Woellner et al., *supra* note 40, at 35 (“[T]he issue of psychological costs has been a largely neglected area of tax compliance costs work.”).

⁴⁹ *See, e.g.*, John Hasseldine & Ann Hansford, *The Compliance Burden of VAT: Further Evidence from the U.K.*, 17 AUSTL. TAX F. 369, 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).

⁵⁰ Tran-Nam & Glover, *supra* note 42, at 519-20.

exhibited physical signs of psychological costs, including “biting lips and wringing hands,” while solving a hypothetical tax problem.⁵¹ Researchers in Spain found that keeping financial records is among the most “disappointing aspects of tax compliance.”⁵²

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.⁵³ The United States has no VAT, and findings about a VAT may not be generalizable to 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.⁵⁴ 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.⁵⁵

2. Potential Subjective Benefits of Taxation

On the other hand, the tax filing process could conceivably generate some

⁵¹ Woellner et al., *supra* note 40, at 44.

⁵² M. Luisa Delgado Lobo, Javier Salinas-Jimenez, José Félix Sanz Sanz, *Hidden Tax Burden of the Personal Income Tax Evidence from the Recent Tax Reform in Spain*, 16 AUST. TAX FORUM 463, 475 (2001).

⁵³ *E.g.*, Hasseldine & Hansford, *supra* note 49; Tran-Nam & Glover, *supra* note 42.

⁵⁴ 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, Brett Freudenberg & Tapan Sarker, *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).

⁵⁵ 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*, WESA (Jan. 28, 2016), <https://www.wesa.fm/archives/2016-01-28/why-dont-americans-claim-their-earned-income-tax-credit> (explaining that some fail to claim the EITC because they fear making a mistake and being audited).

subjective benefits that offset subjective costs.⁵⁶ For instance, some people might feel satisfaction from filing a tax return, similar to the satisfaction of having completed a necessary chore. Some taxpayers may also enjoy the process of reviewing and organizing financial records or may find this organization useful for their business activities unrelated to tax planning. Others might find tax law interesting and may find the process of learning about it 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.⁵⁷ When asked why, people report that they are “good at it” and that the process provides them a “good overview of personal finances.”⁵⁸

There may also be society-wide benefits that accrue 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.⁵⁹ This is in part because filing a tax return makes people aware of their contribution to shared social goods.⁶⁰

C. *What We Should Know: Monetizing Subjective Costs*

Aside from the fact that subjective costs of tax compliance exist, and that they might differ in some systematic way—for instance, by education or income—the nature and scope of these costs is 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

⁵⁶ 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 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).

⁵⁷ AM. ENTER. INST., *supra* note 46, at 79; Pew Rsch. Ctr., *supra* note 46.

⁵⁸ *Id.*

⁵⁹ LAWRENCE ZELENAK, LEARNING TO LOVE FORM 1040, at 4 (2013).

⁶⁰ *Id.* at 111-15. Prof. Zelenak also argues that simplifying the filing process—thereby making it less unpleasant—would enhance this positive consequence of mass tax return filing.

are left guessing where to invest resources to simplify and streamline the process.⁶¹ 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.⁶² 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 the psychological costs. Surveys might ask taxpayers whether they "like" or "dislike" filing taxes overall, which perhaps reveals whether benefits exceed costs for individual taxpayers.⁶³ But we can't compare these relative costs and benefits between individuals or across the population. Does the magnitude of someone's "like" exceeds someone else's "dislike"?

Relatedly, while detailed subjective cost surveys often ask taxpayers to assess the intensity of stress or worry,⁶⁴ these studies often fail to ask whether there are benefits that offset the costs (such as satisfaction or finding the process to be enjoyable).⁶⁵ 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 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"⁶⁶ or strategic behavior. Especially because aversion to tax compliance

⁶¹ 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.

⁶² 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?

⁶³ See Pew Rsch. Ctr., *supra* note 46.

⁶⁴ See, e.g., Tran-Nam & Glover, *supra* note 42, at 520.

⁶⁵ *Id.*

⁶⁶ 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,"

is a well-known cultural trope, survey respondents may reflexively state a strong dislike of filing taxes when 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 themselves 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.⁶⁷ The EITC, the child and dependent care tax credit, and the medical expense deduction all increase the complexity of a tax return.⁶⁸ But the presence of these tax benefits also increases the distributional fairness of the tax system.⁶⁹ Other provisions increase complexity to increase efficiency, like the preferential rate for capital gains.⁷⁰ 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

meaning that respondents' answers have no effect on their lives). See also *infra* Section II.C. 2 for further discussion of the "cheap talk" phenomenon.

⁶⁷ See Emily Cauble, *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, 650-52 (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) ("[S]impler 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 WISC. L. REV. 1267, 1284-91 (1990) (challenging the equity-simplicity and efficiency-simplicity trade-offs).

⁶⁸ I.R.C. §§ 21, 32, 213 (West 2022).

⁶⁹ E.g., MARGOT CRANDALL-HOLLICK & JOSEPH S. HUGHES, CONG. RSCH. SERV., 4RR057, THE EARNED INCOME TAX CREDIT (EITC): AN ECONOMIC ANALYSIS 13-23 (2018) (discussing effects of the EITC on poverty, health, education, and distribution of tax burdens).

⁷⁰ E.g., Noel 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. . . .").

and reform for several decades.⁷¹ 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.⁷² 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 cost-benefit 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 methodology and results. An 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.⁷³ Discrete choice experiments ask survey respondents to choose between hypothetical services with varying attributes, one of which is typically price.⁷⁴ By presenting respondents with a selection of options, including the option *not* to purchase (called the status quo option), the experiments mimic the decisions that people make when facing real-life market choices. By varying the attributes of the services offered and then

⁷¹ 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).

⁷² For seminal work on the subject, see Daniel A. Graham, *Cost-Benefit Analysis Under Uncertainty*, 71 AM. ECON. REV. 715 (1981).

⁷³ See generally Robert J. Johnston et al., *Contemporary Guidance for State Preference Studies*, 4 J. ASS'N ENV. & RESOURCE ECONOMISTS 319 (2017) (providing best-practices recommendations for stated-preference studies, including discrete choice experiments); Giles Atkinson & Susana Mourato, *Discrete Choice Experiments*, in *COST BENEFIT ANALYSIS AND THE ENVIRONMENT* § 5.1 (2018) (describing various trade-offs in the design of discrete choice experiments).

⁷⁴ Johnston et al., *supra* note 73, at 320.

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 likelihood that 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' likelihood of being willing to pay. This would reveal how much respondents value money. By using a regression model (described at length in Section C of the Appendix), we can then calculate what's known as the "implicit price" of the attribute. That is, we can combine our (non-monetary) estimates of how much respondents value the attribute with how much they value money to produce a monetary estimate of how much they value the attribute.

While our discrete choice experiment includes various refinements to optimize statistical power,⁷⁵ it essentially follows the logic above. We vary 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 are varied randomly in an experimental setting, we don't need to worry about confounding respondent characteristics. We can study the *average* change in willingness to pay depending on the attributes offered, which we calculate regardless of the specific motivations for willingness to pay. (We separately ask respondents to explain their willingness to pay and discuss those results in Section II.B.3.) When desired, we can see how specific respondent characteristics influence willingness to pay by explicitly adding them to the regression model.⁷⁶ This allows 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 a tendency by respondents to "provide either unrealistically high or zero [dollar] responses."⁷⁷ In contrast, discrete choice survey designs have been found to outperform surveys based on open-ended questions in

⁷⁵ *Infra* Appendix Section A.4-5.

⁷⁶ *See infra* Section II.B.1, Appendix Section C (discussing how we use interaction terms in the model to calculate differential willingness to pay depending on respondent characteristics).

⁷⁷ Johnston et al., *supra* note 73, at 346.

predicting real-world behavior, potentially because they “lead to higher engagement, increase immersion, and reduce satisficing.”⁷⁸ 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.⁷⁹

Discrete choice experiments also avoid certain forms of bias that can otherwise appear in surveys. Open-ended willingness-to-pay surveys may suffer from status quo bias, where respondents underestimate their willingness to pay for a service that departs from the status quo.⁸⁰ 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.⁸¹ 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.⁸²

Another form of bias in surveys is social desirability bias, where survey respondents infer the surveyor’s desired outcome and shape their answers to meet that outcome.⁸³ 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.⁸⁴ 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.⁸⁵ Because respondents can’t adapt their behavior based on the surveyor, social desirability bias, 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

⁷⁸ 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. 30,527, 2022).

⁷⁹ Johnston et al, *supra* note 73.

⁸⁰ See Raymond S. Hartman, Michael & Donae & Chi-Keung Woo, *Status Quo Bias in The Measurement of Value of Service*, 12 RES. & ENERGY 197, 197-98 (1990).

⁸¹ *Id.*

⁸² See *infra* Appendix Part D.

⁸³ See Sebastian Linzen et al., *The State of The Art of Discrete Choice Experiments in Food Research*, 102 FOOD QUAL. & PREF. 104677, 8 (2022).

⁸⁴ *Id.*

⁸⁵ Stantcheva, *supra* note 78, at 44.

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.⁸⁶

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 people might separately feel *anxious* about making mistakes on their return or being audited.⁸⁷ Therefore, 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 their tax compliance time eliminated.⁸⁸ 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 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:

⁸⁶ It's interesting to note that the pilot results generally aligned with results in the full survey.

⁸⁷ See *supra*, notes 39-42 and accompanying text.

⁸⁸ See Appendix Part A.3.

Figure 1: Example of Choice Set Card

Which of the following tax services would you prefer?

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

Your choice: Service 1 Service 2 Neither

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 1000 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 the 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).⁸⁹

The following table provides brief descriptive statistics for our survey sample. As the results show, the survey sample is relatively representative of the general U.S. population in terms of gender, age (restricted to adults), race, income, and education.⁹⁰

Table 1: Sample Demographics

<i>Gender</i>	<i>Percentage of Sample</i>
Male	59.4%
Female	40.3%

<i>Race</i>	<i>Percentage of Sample</i>
White	64.0%
Black	13.7%
Hispanic	12.6%
Asian	5.5%
Mixed-Race	3.2%

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

⁸⁹ 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.

⁹⁰ 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 BUR., <https://www.census.gov/quickfacts/fact/table/US/PST045221> (last visited Dec. 18, 2022).

<i>Income</i>	<i>Median</i>
Individual Income	\$45,625.04
Household Income	\$75,000

B. Results

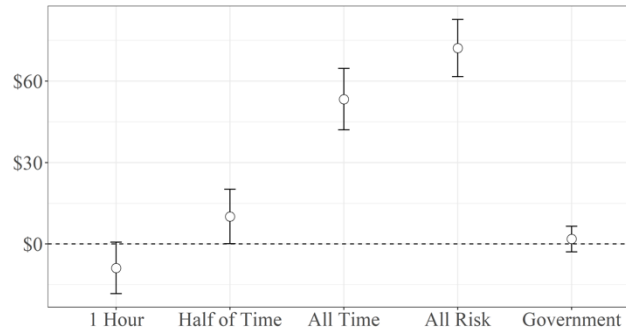
1. Discrete Choice Experiment Results

Calculating willingness to pay across the entire sample of respondents reveals several interesting findings. First, people are more concerned by risk than by 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 find 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 \$3.97 per hour (about \$10 in total, on average) to eliminate half of their tax compliance time. Both figures are significantly lower than the \$10.40 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.

Figure 2 and Table 2 provide implicit prices and confidence intervals for our full-sample model.⁹¹ Both provide results for all three non-price attributes in the choice sets: time savings, risk elimination, and government provision of the service. Based on these results, 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.⁹² In contrast, the attributes related to time savings and risk elimination did affect respondents' choices.

⁹¹ Table 12 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 & Mourato, *supra* note 73, § 5.2 (explaining the conceptual foundation for the calculation of implicit prices in discrete choice experiments).

⁹² We discuss policy implications of this indifference *infra*, Part III.B.4.

Figure 2: Implicit Prices of Attributes**Table 2: Implicit Prices**

<i>Attribute</i>	<i>Implicit Price</i>	<i>95% CI</i>
Risk Elimination	\$72.17	\$61.68 – \$82.66
Time Elimination: 1 Hour	-\$8.79	-\$18.23 – \$0.64
Time Elimination: Half	\$10.13	\$0.11 – \$20.15
Time Elimination: All	\$53.35	\$42.06 – \$64.63
Government	\$1.85	-\$2.87 – \$6.57

In addition to implicit prices calculated using the full sample of respondents, we calculated implicit prices for particular subgroups using the methods described in Section C of the Appendix. (Extensive tables of implicit prices are available in the Online Appendix.⁹³) We find large and statistically significant differences in willingness to pay based on a few respondent characteristics.

Women had much lower willingness to pay to eliminate all time spent on taxes than men—\$34.94 versus \$67.98—but not to eliminate the risk of errors. Respondents who spent more than five hours on tax compliance per year (the median time spent) were willing to pay more to eliminate risk compared to those who spend five hours or fewer—\$81.97 versus \$61.70—and more to eliminate all time spent on taxes—\$68.25 versus \$37.55. Respondents who reported disliking tax compliance activities were willing to pay more to eliminate all time spent on taxes compared to those who did not report disliking tax compliance—\$59.99 versus \$35.84—but not to eliminate risk. Finally, respondents with household annual incomes above the median were willing to pay more to eliminate all time spent compared to those with

⁹³ Online Appendix, <https://jonathanhchoi.squarespace.com/s/Subjective-Costs-of-Taxation-Online-Appendix.pdf>.

household annual incomes at or below the median⁹⁴—\$61.84 versus \$44.10.⁹⁵

Note that each of these subgroup results comes from a regression that includes interaction terms only with a single demographic variable (described in more detail in Section C of the Appendix). Consequently, we should treat these results as descriptive rather than causal. For example, it might (hypothetically) be that men have higher willingness to pay to eliminate all time spent on tax compliance because they take more complicated deductions and credits; in that case, the complexity of their returns would drive their increased willingness to pay, not their gender. Section II.B.2 monetizes welfare benefits of tax reforms using a more sophisticated methodology that includes all variables for which the IRS publishes population-level data.

Other characteristics did not affect willingness to pay, with small and statistically insignificant differences between subgroups. The other characteristics we tested were: level of education,⁹⁶ whether the respondent's state of residence has an income tax, whether the respondent claims the EITC, whether the respondent claims the Child Tax Credit, whether the respondent owns a small business, whether the respondent uses a paid preparer, and whether the respondent expects to receive a tax refund. Finally, different races in some cases had large differences in willingness to pay, but these differences were not statistically significant because there were too few non-white respondents in our sample.

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 of 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 is \$18.83, a statistically significant difference.⁹⁷ Another way to understand this result is that people's willingness to pay for risk elimination is, on average, 35% greater than their willingness to pay to

⁹⁴ This result may seem to contradict the finding, presented in the next section, that the benefits of simplification do not vary substantially between income subgroups. It does not. Note 120 explains how including other variables in the regression and mapping characteristics to the population rather than our sample slightly decreases the magnitudes of the differences, which are modest to start with.

⁹⁵ Respondents with household incomes above the median were also more willing to pay for risk elimination, \$77.18 versus \$66.96, but this difference was not statistically significant at the 95% level.

⁹⁶ We compared people with no college degree versus a college or higher.

⁹⁷ The 95% confidence interval for this difference is \$11.01 to \$26.65.

eliminate all of their time spent on tax compliance.

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.⁹⁸

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. Such a result would be worth paying for apart from the peace of mind it would bestow on the taxpayer.

We chose to frame the discrete choice language around risk rather than anxiety because risk elimination would be a primary outcome of adopting a “return-free filing” system, a prominent type of tax simplification proposal.⁹⁹ In a return-free system, the government completes returns on taxpayers’ behalf.¹⁰⁰ For example, California’s Ready Return program pre-populated taxpayers’ returns with information that the state government already had, significantly simplifying taxpayers’ filing process.¹⁰¹ Many advanced countries have return-free tax systems.¹⁰² The federal government could plausibly eliminate almost all of the risk of error for a wide swath of taxpayers by using such a system.

⁹⁸ See *infra*, Part II.B.3.

⁹⁹ See *infra* notes 171-175 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>.

¹⁰⁰ *Id.*

¹⁰¹ *Id.* (describing Ready Return).

¹⁰² See, e.g., PREFILLED PERSONAL INCOME TAX RETURNS: A COMPARATIVE ANALYSIS OF AUSTRALIA, BELGIUM, CALIFORNIA, QUÉBEC, AND SPAIN (François Vaillancourt ed., 2011) (comparing systems to pre-populate returns in various countries); *Auto-Fill My Return*, GOV’T OF CAN. (May 12, 2022), <https://www.canada.ca/en/revenue-agency/services/e-services/about-auto-fill-return.html> (describing Canada’s current federal system for pre-populating returns).

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, leading to inaccurate survey responses.¹⁰³ We prioritized offering a plausible service over hyper-specific framing, based on best-practices advice to this effect.¹⁰⁴

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 explanations¹⁰⁵ 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.

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 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 is 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 not statistically significant, meaning that we can't reject the possibility at conventional levels of statistical confidence that the true implicit price is zero. Second, if the implicit price is truly negative, a negative value may suggest that respondents rejected the premise. Perhaps an hour of time savings was so small relative to their overall tax compliance time that the suggestion of it offended them. Alternatively, respondents might have believed that one hour of time savings was so minimal that it wouldn't merit the effort required even

¹⁰³ 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."

¹⁰⁴ Best-practices guidance for discrete choice experiments stressed the importance of offering respondents plausible choice scenarios. As Johnston et al. explain, the choices offered 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 73, at 326.

¹⁰⁵ See *infra* Part II.B.3.

to purchase such a service. That is, it might reflect some presumption of transaction costs that our model ignores.

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.¹⁰⁶ 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 to do.¹⁰⁷ Once begun, however, the disutility (unpleasantness) of working drops significantly, and in some cases may even be overtaken by utility (“an excess of satisfaction”).¹⁰⁸ Of course at some point, after too many hours of work, disutility will once again overtake utility.¹⁰⁹

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. Among our respondents, the average annual time spent on tax compliance work is 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. If we consider the types of tasks that might make up individual tax compliance work, this result may not seem 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

¹⁰⁶ See 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).

¹⁰⁷ David 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) (quoting STANLEY JEVONS, *THE THEORY OF POLITICAL ECONOMY* 191 (1871)) (“At the moment of commencing labor it is usually more irksome than when the mind and body are well bent to the work.”).

¹⁰⁸ *Id.*

¹⁰⁹ *Id.*

require inputting numbers into tax software or sending information to a tax preparer. The last hour likely entails double checking the 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.¹¹⁰ Even if someone dislikes working on their tax return, once they've begun, they accept their fate and make the best of it.

c. Subjective Versus Objective Costs

Our study finds 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 \$27.50 but were only willing to pay \$10.40 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 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.¹¹¹ 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.¹¹²

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

Third, responses may reflect mental accounting. For instance, some people may "bucket" their spending into predetermined categories, like food,

¹¹⁰ Ed Diener, Richard E. Lucas & Christie Napa Scollon, *Beyond the Hedonic Treadmill*, 61 AM. PSYCH. 305, 305 (2006) (explaining hedonic adaptation).

¹¹¹ 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.3.

¹¹² See *supra* Part I.B.2.

rent, and entertainment. If respondents don't have an existing bucket for "reducing tax compliance time," they may be unwilling to add it.¹¹³ 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.¹¹⁴

2. Monetizing the Benefits of Tax Reform

In addition to individual estimates of implicit prices, the discrete choice model allows us to estimate the dollar benefits of large-scale government reforms, by aggregating individual willingness to pay over the entire population. To do this, we multiply the number of individuals with each set of characteristics by the implicit prices for that individual.

For example, imagine a female taxpayer with household income between \$30,000 and \$40,000 per year. She files a separate tax return and claims the Child Tax Credit, but not the EITC. She takes the standard deduction and has no small business income. Using a discrete choice model that takes all of these characteristics into account, we estimate that this taxpayer would be willing to pay \$47.59 to eliminate all the time she spends on tax compliance, \$10.14 to eliminate half the time she spends on tax compliance, and \$70.55 to eliminate all risk of error in her tax returns. We also estimate that there are 145,552 taxpayers with this exact set of characteristics.¹¹⁵ Therefore, we

¹¹³ 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. DEC. 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).

¹¹⁴ 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*, 105 ORG. 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").

¹¹⁵ We estimate this number by using the IRS's Statistics of Income. *SOI Tax Stats - Statistics of Income*, I.R.S. (Dec. 8, 2022), <https://www.irs.gov/statistics/soi-tax-stats-statistics-of-income>. The discrete choice model includes all available taxpayer characteristics that the IRS includes in Statistics of Income. Certain demographic characteristics, like race and education, are excluded because the IRS doesn't compile information on them. Because the IRS generally reports the number of returns filed by each return characteristic and income (for example, the number of joint returns filed within each income group), we impute more

estimate that taxpayers like this one would be willing to pay \$6.9 million in total each year to eliminate all time spent on taxes, \$1.5 million in total each year to eliminate half of the time spent on taxes, and \$10.3 million in total each year to eliminate all risk of error. By conducting the same analysis for every possible combination of characteristics, we can produce the following estimate of dollar benefits of tax reform (with 95% confidence intervals in parentheses¹¹⁶):

Table 3: Monetary Benefits – All Taxpayers¹¹⁷

Percentage of total population	100.00%
Benefits from total time elimination	\$8.68 billion (\$6.02b – \$13.33b)
Benefits from risk elimination	\$11.98 billion (\$8.7b – \$16.33b)

We can also produce estimates for any subgroup of taxpayers based on their specific characteristics. Helpfully, this allows us to restrict our analysis to taxpayers whose returns are the most realistic objects of reform. These might include taxpayers who take the standard deduction rather than itemizing, because they likely have simpler returns that might be more easily pre-populated by the IRS. They also might include low-income taxpayers and taxpayers without small business income, for the same reasons—the government could easily pre-populate a tax return for a relatively low-income taxpayer who takes the standard deduction and receives only wage income.¹¹⁸

elaborate combinations assuming that characteristics are independent of each other when their joint distribution is unknown. While this is unlikely to be strictly true, it seems to be an appropriate working assumption and doesn't substantially affect the results.

¹¹⁶ We use a method called “bootstrapping” to generate confidence intervals. First, we randomly draw discrete choices from our sample (with replacement) to re-construct 100 samples of identical size. We conduct the same multinomial logistic regression and calculate the relevant monetary benefit estimates using each reconstructed sample. We generate “empirical” bootstrap intervals by taking the 2.5th-percentile and 97.5th-percentile estimates. See A. C. DAVISON & D. V. HINKLEY, *BOOTSTRAP METHODS AND THEIR APPLICATION* 194 (1997). Because bootstrapping is a non-parametric method to calculate standard errors, it avoids potential statistical problems that could arise using standard parametric techniques. *Id.*

¹¹⁷ The estimates in this section ignore taxpayers' out-of-pocket spending on tax preparation services. The full monetary benefit of a reform that eliminated taxpayers' need to file a tax return would ideally account for such out-of-pocket expenses—while also perhaps acknowledging that such fees represent income to the service providers.

¹¹⁸ This statement sets aside possible complexities related to claiming children for the EITC and Child Tax Credit. For most families, the IRS could presume that the same adult will claim the same child for each year that the child is eligible. For certain families, however, care arrangements change over time in such a way that families would need to provide additional information about family structure to the IRS to ensure that a pre-populated tax

The following tables give dollar estimates of tax reform for these subgroups.

Table 4: Monetary Benefits – Taxpayers Who Take the Standard Deduction

Percentage of total population	90.52%
Benefits from total time elimination	\$7.80 billion (\$5.55b – \$10.33b)
Benefits from risk elimination	\$10.41 billion (\$8.41b – \$13.54b)

Table 5: Monetary Benefits – Taxpayers with No Small Business

Percentage of total population	73.23%
Benefits from total time elimination	\$5.88 billion (\$4.34b – \$7.93b)
Benefits from risk elimination	\$8.23 billion (\$6.61b – \$10.6b)

Table 6: Monetary Benefits – Taxpayers with Household Income ≤ \$50,000

Percentage of total population	58.10%
Benefits from total time elimination	\$4.46 billion (\$2.84b – \$5.76b)
Benefits from risk elimination	\$5.82 billion (\$4.52b – \$7.37b)

As explained above,¹¹⁹ respondents from higher-income households were generally willing to pay more for total time elimination and risk elimination. But this difference becomes much smaller once we include other variables in the regression and correct for sampling bias by weighting our coefficients to reflect the full population of filers. For example, taxpayers with household incomes less than \$50,000 make up 58.1% of all taxpayers, but their total willingness to pay for complete time elimination and risk elimination are 51.4% and 48.6% of the overall total.¹²⁰

return is accurate. We ignore these complexities in our estimates, presuming that pre-populated tax returns would still *dramatically* simplify tax filing for nearly all low-income taxpayers, even if for those who must report some small amount of information to the IRS each year.

¹¹⁹ See *supra* Part II.B.1.

¹²⁰ Estimates of per-person willingness to pay between high- and low-income taxpayers differ from those discussed in Section II.B.1 for two reasons. First, the difference in respondents' willingness to pay between income groups may be driven by other non-income characteristics, such as other return attributes (itemizing deductions, types of income) and gender. When those other characteristics are included as interaction terms in our regressions,

Advocates who focus on low-income taxpayers have argued for simplifying EITC compliance by automatically sending the credit to presumptively qualified households.¹²¹ Restricting the estimate only to recipients of the EITC produces the following estimate:

Table 7: Monetary Benefits – Taxpayers Who File for the EITC

Percentage of total population	16.17%
Benefits from total time elimination	\$1.38 billion (\$0.71b – \$1.66b)
Benefits from risk elimination	\$1.77 billion (\$1.25b – \$2.77b)

What can we learn from these estimates? First, we find that the benefits of eliminating all tax compliance burdens for all individual taxpayers are large, but much smaller than prior commentators have claimed.¹²² As explained above, standard analyses monetize time spent on tax compliance using some wage rate. For example, the Tax Foundation monetizes individual taxpayers' time spent on tax compliance by multiplying IRS estimates of hours spent by \$35.98 per hour, the average wage plus benefits for professional tax preparers in the United States.¹²³ Using these figures, they estimate that time spent on individual income taxes costs American households \$73.7 billion per year.¹²⁴ In line with our finding that individual willingness to pay is lower than the average wage rate, we estimate that the total subjective cost of all time spent on individual tax returns is \$8.68 billion,

the systematic difference in willingness to pay by income group drops away. Second, Section II.B.1 calculates average differences in willingness to pay *among respondents in our sample*. In contrast, this Section calculates aggregate differences based on subgroups *in the actual population*. Because our sample slightly differs from the population due to our pre-screening (most importantly, we only include taxpayers who spent at least three hours on tax compliance), the estimates in this Section correct for sampling bias by attempting to simulate the differences between subgroups in the real world.

¹²¹ See, e.g., Nina Olson, *We Applaud Your Work and We Feel Your Pain, But We Need You to Do More to Get Dollars Out to Vulnerable Taxpayers*, PROCEDURALLY TAXING (Aug. 4, 2020), <https://procedurallytaxing.com/irs-we-applaud-your-work-and-we-feel-your-pain-but-we-need-you-to-do-more-to-get-dollars-out-to-vulnerable-taxpayers> (discussing automatic payment of the EITC and citing earlier proposals for the same); CASSANDRA ROBERTSON, GABRIEL ZUCKER & NINA OLSON, NEW AM., STRATEGIES FOR INCREASING UPTAKE OF THE EARNED INCOME TAX CREDIT (2020), <https://www.newamerica.org/pit/reports/strategies-increasing-uptake-earned-income-tax-credit> (“The best option is to automate the [EITC] payments, regardless of whether they are explicitly claimed.”).

¹²² See *supra* notes 34-36 and accompanying text.

¹²³ Hodge, *supra* note 3.

¹²⁴ *Id.*

with a 97.5% chance of being less than \$13.33 billion—significantly lower than figures calculated using average wage rates.

Our findings also complicate attempts to monetize the benefits of particular reforms. For example, the Tax Foundation has estimated that the Tax Cuts and Jobs Act (TCJA) led to “compliance savings” worth \$5.4 billion by saving taxpayers one hour of tax compliance time.¹²⁵ 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.¹²⁶ Based on our study, we should view arguments like these with greater skepticism.

Nonetheless, our findings still leave a significant role for tax simplification. In particular, we find that per-individual benefit estimates do not vary substantially between income subgroups, implying that we should focus on lessening burdens for low-income taxpayers with simple returns. Table 8 provides benefit estimates for this specific subgroup, with both relatively low income and without itemized deductions or business income that might require more complex administrative work for the IRS.

Table 8: Monetary Benefits – Taxpayers with Household Income ≤ \$50,000, Filing Standard Deduction, with No Small Business Income

Percentage of total population	44.31%
Benefits from total time elimination	\$3.33 billion (\$2.14b – \$4.24b)
Benefits from risk elimination	\$4.03 billion (\$3.31b – \$5.39b)

Simplification would be relatively straightforward for the taxpayers in Table 8, who make up a substantial fraction of all taxpayers. The group’s willingness to pay as a percentage of total willingness to pay for all taxpayers

¹²⁵ Erica York & Alex Muresianu, *The Tax Cuts and Jobs Act Simplified the Tax Filing Process for Millions of Households*, TAX FOUND. (2018), <https://taxfoundation.org/the-tax-cuts-and-jobs-act-simplified-the-tax-filing-process-for-millions-of-americans>. 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.

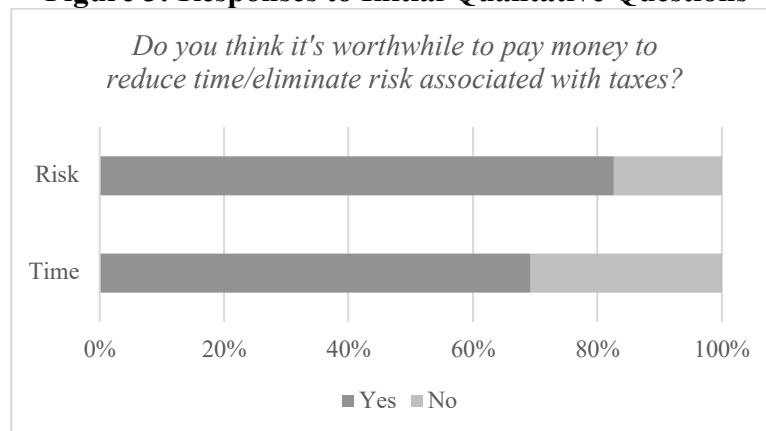
¹²⁶ William Gale et al., *A Preliminary Assessment of the Tax Cuts and Jobs Act of 2017*, 71 NAT’L TAX J. 589, 589 (estimating that the TCJA would “make the distribution of after-tax income less equal”).

is close to the group's percentage of the total population—38.4% for total time elimination and 33.6% for total risk elimination. In other words, the benefits of simplification are relatively constant between simple and complicated tax returns, even though the costs of simplification are much lower for simple returns. This implies that simplification efforts should focus on the poorest taxpayers with the simplest returns.¹²⁷ We discuss the policy implications of these findings in greater detail below, in Part III.B.3.

3. 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.¹²⁸ Each respondent was randomly asked either about time savings or risk reduction, 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* 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

¹²⁷ Declining marginal utility of income also counsels toward focusing on reforms that affect low-income taxpayers. For further explanation and discussion, see *infra* Section III.B.3.

¹²⁸ For more information about the qualitative questions, see *infra* Appendix Section A.6.

according to a list of common answer features.¹²⁹ 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.¹³⁰ Most of these answers, 69%, reflect the view that their time could be better spent in some way other than preparing their tax return. 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 reflect 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.”

¹²⁹ 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, 344 (“[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.

¹³⁰ 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. . . .”

“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. The most common answer in this group, 56%, stated that they 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:

“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 non-tax 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.¹³¹ These responses are 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 they were previously asked to consider before the qualitative questions.¹³² This category of explanation is

¹³¹ 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.”

¹³² 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

consistent with our finding that respondents would generally pay more to eliminate risk than to reduce time spent on taxes. It 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. Those who answered yes to this question provided diverse reasons for their willingness to pay.¹³³ 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 example:

“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.¹³⁴ 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

exploitation of citizens to make them pay money to pay taxes.”

¹³³ 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.”

¹³⁴ Many answers reflected concern about both current worry as well as future hassle. These answers were coded with both categories.

reviewing tax tables -- that you wouldn't have had to do, if you'd just paid the fee to eliminate the risk."

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.¹³⁵ 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.¹³⁶ 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*

¹³⁵ 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.

¹³⁶ 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.

extremely low for ordinary taxpayers—the overall audit rate from 2010 to 2019 was 0.25%.¹³⁷ 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 smaller portion, 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 reiterate 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.¹³⁸

C. Potential Objections and Robustness Checks

1. External Validity

As reported above,¹³⁹ our sample largely resembled the general population in terms of gender, race, income, and educational attainment. However, while online survey platform users may be representative of an “average” middle-income person, they may not reflect the full U.S. population.¹⁴⁰ Moreover, our pre-screening procedure explicitly selected for

¹³⁷ 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>.

¹³⁸ See *supra* text accompanying note 98.

¹³⁹ See *supra* Section II.A.3 for discussion of our sample characteristics.

¹⁴⁰ 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. Smith, Catherine A. Roster, Linda L.

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

Consequently, low-income taxpayers are 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 28.1% of federal tax filers.¹⁴¹ And higher-income taxpayers are somewhat *over*represented, with 34.7% of our respondents reporting household incomes above \$100,000 per year, as opposed to 19.4% of federal tax filers.¹⁴²

Because higher-income taxpayers tend to have higher willingness to pay in general (without controlling for other variables), our estimates of individual willingness to pay may be excessively high. On the other hand, the monetary estimates calculated in Section II.B.2 extrapolate estimates for the population using the actual distribution of household incomes, essentially overweighting our estimates for low-income households. This means that the underrepresentation of low-income households in our sample would not alone bias the estimates, *unless* the low-income respondents in our sample are unrepresentative of low-income individuals in some other way.

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 report negative feelings about tax compliance in nationwide polls.¹⁴³

A final concern related to external validity is the possibility of framing effects: Respondents' answers may depend on arbitrary survey

Golden & Gerald S. Albaum, *A Multi-Group Analysis of Online Survey Respondent Data Quality: Comparing a Regular USA Consumer Panel to MTurk Samples*, 69 J. BUS. RES. 3139, 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").

¹⁴¹ I.R.S., *INDIVIDUAL INCOME TAX RETURNS: COMPLETE REPORT 2020* 49 tbl. 1.1 (2022), <https://www.irs.gov/pub/irs-pdf/p1304.pdf>.

¹⁴² *Id.*

¹⁴³ See AM. ENTER. INST., *supra* note 46, at 79; Pew Rsch. Ctr., *supra* note 46 (reporting that 56% of Americans feel negatively about "doing their income taxes"); *supra*, text accompanying notes 46-47.

characteristics, like question order or wording.¹⁴⁴ In our case, for instance, asking respondents about their tax filing experiences at the start of the survey might focus 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, 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.¹⁴⁵ Where a respondent's expression of dislike is qualitative and vague, it may reflect a superficial attitude that would change upon further inspection. 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 be reflexive and unconsidered. This problem is generally known as "hypothetical bias."¹⁴⁶ A substantial literature suggests that hypothetical bias generally leads to overestimation—that is, respondents generally give excessively high willingness-to-pay estimates in hypotheticals compared to their willingness to pay in real life.¹⁴⁷ Thus, hypothetical bias would cause 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.¹⁴⁸ Many empiricists prefer to use revealed preferences where possible, for example, by looking at consumers' willingness to pay for an accountant to save time. Regarding the economics literature, the debate mostly concerns contingent valuation surveys that ask individuals to how they value inherently hard-to-value public goods, such as clean river water or saving the California condor from extinction.¹⁴⁹ Something called the "embedding effect" leads people to offer

¹⁴⁴ Jacob Goldin & Daniel Reck, *The Analysis of Survey Data with Framing Effects*, 73 AM. STATISTICIAN 264, 264 (2019).

¹⁴⁵ See Crawford, *supra* note 66, at 286.

¹⁴⁶ E.g., David A. Hensher, *Hypothetical Bias, Choice Experiments and Willingness to Pay*, 44 TRANSP. RSCH. PART B: METHODOLOGICAL 735 (2010).

¹⁴⁷ *Id.* at 739.

¹⁴⁸ See Johnston et al., *supra* note 73, at 321 (discussing this problem with stated preference surveys).

¹⁴⁹ See Daniel Kahneman & Jack Knetsch, *Valuing Public Goods: The Purchase of*

very similar values for very different interventions.¹⁵⁰ For instance, they might provide “willingness-to-pay to clean up one lake roughly equal to that for cleaning up five lakes.”¹⁵¹ Such illogical responses produce implausible per-unit variable estimates.¹⁵²

However, these criticisms tend to deal narrowly with surveys that ask respondents to value public goods.¹⁵³ 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.¹⁵⁴ 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.¹⁵⁵ 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 time savings rather than less. 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.¹⁵⁶ In the context of tax compliance, a revealed preference study might observe taxpayers in deciding whether to pay for an accountant, or whether to claim

Moral Satisfaction, 22 J. ENVIR. ECON. & MGMT 57, 58-59 (1990).

¹⁵⁰ *Id.*

¹⁵¹ Peter A. Diamond & Jerry A. Hausman, *Contingent Valuation: Is Some Number Better than No Number?*, 8 J. ECON. PERSPECTIVES 45, 46 (1994).

¹⁵² 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.

¹⁵³ 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.

¹⁵⁴ See Kahneman & Knetsch, *supra* note 149, at 58-59 (identifying the embedding effect only in the context of valuation of public goods); Diamond & Hausman, *supra* note 151, at 46 (explaining that “the embedding effect is usually thought to arise from the nonexistence of individual preferences for the public good in question”).

¹⁵⁵ *Id.* at 47.

¹⁵⁶ *E.g.*, Benzarti, *supra* note 43, at 1 (observing the revealed preferences of taxpayers deciding between itemizing deductions or claiming the standard deduction to estimate certain specific hassle costs of tax compliance).

a certain tax credit.¹⁵⁷ 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 find the greater inclusivity of stated preference surveys to be appealing, despite their limitations.

3. Statistical Validity and Survey Fatigue

Consistent with best research practices,¹⁵⁸ we pre-registered our experimental design and statistical models with the Open Science Framework.¹⁵⁹ We also conduct two robustness checks to confirm the validity of our experimental design. Section E of the Appendix describes how our study passes tests for the Independence of Irrelevant Alternatives, an important statistical prerequisite for the validity of our discrete choice model.¹⁶⁰

An additional concern is that respondents may experience survey fatigue, decreasing the quality of their responses as the survey goes on. We test this by re-estimating implicit prices based only on the responses in our sample to the first six choice cards. Table 9 shows the results from this analysis.

Table 9: Main Effects Implicit Prices

<i>Attribute</i>	<i>Implicit Price</i>	<i>95% CI</i>
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

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

¹⁵⁷ *E.g., id.*

¹⁵⁸ See generally John P. A. Ioannidis, *Why Most Published Research Findings Are False*, 2 PLOS MEDICINE 696 (2005) (how many published findings are false due to publication bias and p-hacking, and suggesting pre-registration as one remedy).

¹⁵⁹ *Subjective Costs of Taxation*, OPEN SCIENCE FRAMEWORK REGISTRIES (Oct. 18, 2022), <https://osf.io/s8q7h>.

¹⁶⁰ Atkinson & Mourato, *supra* note 73, at § 5.2.

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 foregone 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.¹⁶¹ However, our discrete choice experiment revealed implicit prices for time savings that were on average below respondents' hourly wages.¹⁶² For their last hour of tax compliance work, taxpayers are willing to pay *far* less than 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, 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.”¹⁶³ For these taxpayers, it may be that the subjective benefits of tax compliance outweigh any subjective costs they experience. Several of the qualitative answers suggest 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.¹⁶⁴ These findings support scholarship on the diverse benefits of the tax filing process.¹⁶⁵ While much of this scholarship relates to society-wide benefits from tax filing, our results suggest that there are individual-level benefits as well.¹⁶⁶

¹⁶¹ *E.g.* Evans, *supra* note 28, at 451 (stating that psychological costs occur “[i]n addition to” the time costs of tax compliance).

¹⁶² *See supra* Part II.B.1.c (discussing this result).

¹⁶³ Another 20% reported that they find tax compliance work to be neither pleasant nor unpleasant.

¹⁶⁴ *See supra* Part II.B.3.

¹⁶⁵ *See supra* Part I.B.2.

¹⁶⁶ *E.g.*, ZELENAK, *supra* note 59, at 4; *cf.* WILLIAMSON, *supra* note 56, at 180-82 (describing taxpaying as a source of pride among survey respondents and Americans generally).

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.¹⁶⁷ Analysts use market wages to value people's time for nearly all cost-benefit analysis across government agencies, not just for tax policy analysis.¹⁶⁸ Our findings suggest that these 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.¹⁶⁹ 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.¹⁷⁰ Contrary to some conventional economic theory, our respondents experience *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 costs 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

¹⁶⁷ See *supra* notes 34-36 and accompanying text.

¹⁶⁸ This valuation method has been repeatedly and persuasively challenged elsewhere. See, e.g., Samaha, *supra* note 21, at 328-36 (explaining why hourly wages are not an accurate basis for monetizing time burdens); Feather & Shaw, *supra* note 21; McConnell & Strand, *supra* note 21.

¹⁶⁹ The latter interpretation relies on questionable assumptions, including that a person can easily increase their hours of labor.

¹⁷⁰ E.g. Liu, *supra* note 106, at 188.

all publications that use such estimates. Such context might explain, for instance, how to narrowly interpret the market-based cost estimates.

Our specific findings also counsel policymakers to think big when it comes to reducing tax compliance work, focus on anxiety over aggravation, and prioritize reforms that affect low-income taxpayers. Our final section 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 support to proposals for return-free filing, in which many (most) individual taxpayers do not need to file a tax return.¹⁷¹ 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.¹⁷² 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.¹⁷³ 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 taxpayers, especially audits triggered when taxpayers' reported information conflicts with third-party reported information.¹⁷⁴

This discussion of return-free filing is especially salient now. Congress recently mandated, in the Inflation Reduction Act of 2022, that the IRS study

¹⁷¹ See *supra* notes 100-102 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>.

¹⁷² See Tax Pol'y Ctr., *What Is Return Free Filing and How Would It Work?*, in TAX POL'Y CTR. BRIEFING BOOK (2020), <https://www.taxpolicycenter.org/briefing-book/what-return-free-filing-and-how-would-it-work>.

¹⁷³ *Id.*

¹⁷⁴ 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.

how to implement a free direct e-file program.¹⁷⁵ Although a direct e-file program differs from a return-free system, it could come quite close. For instance, a government-run e-file system could pre-populate tax returns with all third-party information the IRS has received, as well as information from prior years' returns. For many (perhaps most) taxpayers, no additional information would be required. Thus, a direct e-file system could operate the same as a return-free system that uses tax agency reconciliation, as described above. Such a system would nearly eliminate aggravation costs for most taxpayers. And, by including all the information the IRS currently has, government-prepared returns would dramatically reduce the risk of audits that currently occur when taxpayers neglect to include third-party reported information. A direct e-file program would also eliminate the direct spending on fees for tax preparation software or accountants. Our respondents spend, on average, \$236.25 per household on such costs annually.

Our findings counsel toward adopting the most expansive version of a direct e-file program, one that includes the maximum amount of information and requires the least amount of taxpayer input. A partially populated government return would still require a significant investment by the IRS, with limited benefits for taxpayers. Moreover, convincing taxpayers to switch from their current tax preparation method to a public method may require more than merely offering a free service. Taxpayers may be loath to adopt a bad service, no matter how free it is; and as our results show, taxpayers are relatively indifferent as to whether the service is offered by the government or a private company, instead focusing on the quality of the service. To ensure the success and widespread adoption of such a program, the IRS may need to offer a *better* service. For all these reasons, the IRS will get more bang for its buck by providing taxpayers with a fully prepared tax return.

Of course, notwithstanding declining marginal costs of time, policymakers should continue to pursue reforms that simplify the tax-filing process. Many reforms have low costs or could even decrease net administrative costs, such as attempts to modernize tax filing with new technology.¹⁷⁶

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 and simplify eligibility standards for claimants.¹⁷⁷ Although such reforms would likely simplify how taxpayers

¹⁷⁵ Inflation Reduction Act of 2022, P.L. 117-169 § 10301(1)(B) (authorizing \$15 million toward this effort).

¹⁷⁶ See I.R.S., IRS INTEGRATED MODERNIZATION BUSINESS PLAN (2019) (describing proposals to improve tax administration through new technology).

¹⁷⁷ See Jacob Goldin & Ariel Jurow Kleiman, *Whose Child Is This? Improving Child-*

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.¹⁷⁸ Tax filing software like TurboTax largely automates the calculations required to file taxes, which taxpayers previously would have had to compute by hand. Legislators have responded by massively increasing the amount of computation required to file a tax return.¹⁷⁹ While scholars have pointed out both costs and benefits to this increase in computational complexity,¹⁸⁰ these scholars generally assume that computational complexity is essentially costless to taxpayers.¹⁸¹ Some provisions that increase computational complexity—like 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 one¹⁸²—may not require taxpayers to laboriously commit pen to paper, but they surely increase the likelihood that taxpayers will make a mistake. And, just as importantly, they

Claiming Rules in Safety-Net Programs, 131 YALE L.J. 1719, 1764-82 (2022) (suggesting and evaluating such reforms to child-claiming rules).

¹⁷⁸ See generally Lawrence Zelenak, *Complex Tax Legislation in the TurboTax Era*, 1 COLUM. J. OF TAX L. 91 (2010) (coining the phrase “computational complexity” and describing this phenomenon).

¹⁷⁹ *Id.*

¹⁸⁰ 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); David I. Walker, *Tax Complexity and Technology*, 97 IND. L.J. 1095, 1105-17 (2022) (arguing that computational complexity can discourage inefficient tax planning).

¹⁸¹ E.g., *id.* at 1097 (describing computational complexity as “trivial and acceptable” given tools like TurboTax); Zelenak, *supra* note 179, 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 I.R.C. § 55 (2022); Zelenak, *supra* note 179, at 99-104 (discussing the computational complexity of the Alternative Minimum Tax).

may also increase taxpayers' *perception* that they are likely to make a mistake.

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.¹⁸³ 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.¹⁸⁴ 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.¹⁸⁵ While anxiety may motivate compliance, it also imposes psychological costs on individuals in the form of discomfort and stress.¹⁸⁶ 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 taxpayers.¹⁸⁷ These considerations suggest that fomenting anxiety¹⁸⁸ 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 the full balance right away.¹⁸⁹ The IRS could publicize these

¹⁸³ See, e.g., GREENSTEIN, WANCHECK & MARR, *supra* note 22, at 3-4 (explaining how the complexity of the EITC rules accounts for most erroneous claims); Goldin & Jurow Kleiman, *supra* note 177, at 1759 (explaining how current child-claiming rules in tax credits often misalign with taxpayers' intuitions about caregiving).

¹⁸⁴ E.g. York & Muresianu, *supra* note 125, at 3-5 (describing how the TCJA simplified tax filing by increasing the standard deduction).

¹⁸⁵ 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).

¹⁸⁶ 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> ("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.").

¹⁸⁷ Leandra Lederman & Joseph C. Dugan, *Information Matters in Tax Enforcement*, 2020 B.Y.U. L. REV. 145, 160-65 (2020) (emphasizing the importance of third-party reporting to tax compliance and enforcement).

¹⁸⁸ 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 185, at 17 tbl.1.

¹⁸⁹ *Additional Information on Payment Plans*, I.R.S.,

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 engage in good faith.¹⁹⁰ For most taxpayers, a mistake on a tax return simply means paying any additional taxes plus interest.¹⁹¹ 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.¹⁹² 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.¹⁹³ Publicizing the difference between good faith tax filing and abusive tax behavior, as well as the remedies available to taxpayers who make an honest mistake, might help to reduce taxpayer anxiety.

3. Focus on Simple, Low-Income Returns

We find that taxpayers' willingness to pay to reduce aggravation and anxiety is relatively constant across income levels.¹⁹⁴ Given declining marginal utility of income, this finding suggests that the actual disutility of aggravation and anxiety may be higher for low-income individuals. Moreover, while the monetary benefits are relatively constant, the *administrative costs* of simplification are likely much higher for high-income taxpayers and those with complex returns. For example, while the IRS could almost entirely pre-populate a tax return for someone who receives only wage income, it would be very difficult for the IRS to assemble enough data to pre-populate a tax return for someone with a small business. Therefore, policymakers and the IRS should focus relatively more resources on reducing subjective costs of tax compliance for low-income taxpayers with simple tax filings.

<https://www.irs.gov/payments/payment-plans-installment-agreements#plandef> (last visited Dec. 14, 2022).

¹⁹⁰ 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).

¹⁹¹ I.R.C. § 6601.

¹⁹² See I.R.C. § 6604(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).

¹⁹³ *Supra* Section II.B.3.

¹⁹⁴ See *supra* Section II.B.2.

Declining marginal utility of income means that a lower-income person derives more “utility” (happiness, satisfaction, wellbeing, etc.) from an additional dollar of income compared to a higher-income person.¹⁹⁵ In other words, \$50 is worth more to a low-income person than to a high-income person. Thus, if a low-income person is willing to pay the same amount as a high-income person to reduce the aggravation or anxiety associated with tax compliance, the low-income person is giving up something more valuable to them. Their willingness to pay, then, implies a higher amount of discomfort.¹⁹⁶

Even if we ignore individuals’ wellbeing and focus only on monetary estimates, a rough cost-benefit analysis supports focusing on automating simple returns filed by low-income taxpayers. One major obstacle to return-free filing is that the government lacks the information required to determine whether taxpayers are eligible for the constellation of deductions available under the modern tax code. Collecting this information is administratively costly but eliminating the deductions could potentially make the tax system less targeted and therefore less fair. Our findings, as presented in Section II.B.2, suggest that we could capture 44.31% of the total benefits from return-free filing by focusing on low-income taxpayers with no small business income and no itemized deductions. These are the taxpayers with the very simplest returns, which could be pre-populated with relative ease. It turns out that the lowest-hanging fruit is just as sweet.

These considerations counsel toward focusing greater resources on tax simplification targeting low-income taxpayers, rather than simplification efforts that primarily benefit high-income taxpayers. For example, we should prioritize return-free filing for simple returns over streamlining the process to file estate taxes, which only apply to estates above \$12.92 million.¹⁹⁷ This is the *opposite* of the usual logic of cost-benefit analysis. Because rich people earn higher wages and generally have higher willingness to pay for all services, cost-benefit analysis notoriously privileges the rich at the expense of the poor.¹⁹⁸ For tax compliance, though, the cost-benefit calculus flips.

¹⁹⁵ See Bankman & Griffith, *supra* note 25, at 1946-48 (discussing declining marginal utility).

¹⁹⁶ Neoclassical economists sometimes argue that we should simply maximize aggregate willingness to pay and then redistribute using taxes and transfers. See, e.g., Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. LEGAL STUD. 669 (1994) (giving the classic articulation of this view). This argument has come under recent criticism, *infra* note 198, and in any case our findings suggest it would be better to simplify compliance for low-income taxpayers even on willingness-to-pay-maximization grounds alone.

¹⁹⁷ The exemption is per individual, so married couples may exclude twice as much. I.R.C. § 2010 (2022).

¹⁹⁸ See, e.g., Zachary Liscow, *Is Efficiency Biased?*, 85 U. CHI. L. REV. 1656 (2018)

Reforms that simplify the EITC (and, to a lesser extent, the Child Tax Credit) might be especially worthwhile.¹⁹⁹ As one example, although it would be a significant departure from current law, providing a universal child benefit to all households with children would significantly reduce both the aggravation and anxiety associated with tax filing for low-income families.²⁰⁰ The current complex rules underlying the EITC and Child Tax Credit largely arise due to the work requirement and means-test built into these child-based credits.²⁰¹ With a universal benefit, the IRS could automatically send payments to all households with children. Although a universal credit would certainly be more expensive than the current limited approach, this additional cost would be offset by the elimination of subjective costs of tax compliance among recipient households.

4. 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.²⁰² In their view, it doesn't matter whether tax automation is 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 a government-provided service.

On the first point, recent experience with the so-called "Free File" program offers compelling evidence in support of direct government provision.²⁰³ IRS Free File is a public-private partnership between the IRS and the Free File Alliance, a group of tax software companies that agreed to

(arguing that cost-benefit analysis is generally rich-biased); POSNER AND ADLER, *NEW FOUNDATIONS OF COST-BENEFIT ANALYSIS* 16-19 (2008) (same); Eric A. Posner & Matthew D. Adler, *Rethinking Cost-Benefit Analysis*, 109 *YALE L.J.* 184 (1999) (same).

¹⁹⁹ For discussion of how to simplify the child-claiming rules for the EITC and Child Tax Credit, see generally Goldin & Jurow Kleiman, *supra* note 177 (assessing how child-claiming rules contribute to higher compliance costs, among other negative outcomes, and offering reform proposals).

²⁰⁰ In 2021, Congress enacted the American Rescue Plan Act, which temporarily created a near-universal child benefit that was available to all families with income below a certain level, regardless of employment status. American Rescue Plan Act, Pub. L. No. 117-2, § 9621, 135 Stat. 4, 144 (2021) (removing the Child Tax Credit (CTC) refundability limits for 2021).

²⁰¹ See Goldin & Jurow Kleiman, *supra* note 177, at 1765-1770 (explaining how child-claiming rules could be dramatically simplified under a universal benefit structure).

²⁰² See *infra* Appendix Part D for the regression table with survey results.

²⁰³ See *Free File: About the Free File Alliance*, I.R.S., <https://www.irs.gov/e-file-providers/about-the-free-file-alliance> (last visited Jan. 10, 2023).

provide free online tax filing for low- and moderate-income taxpayers.²⁰⁴ In exchange, the IRS agreed not to create its own free tax filing software.²⁰⁵ In theory our findings support this arrangement, since taxpayers should be indifferent as to whether a private company or the IRS provides tax filing services.

In practice, however, the Free File partnership proved fraught. In 2019, ProPublica revealed that Intuit, arguably the most powerful member of the Free File Alliance, charged low-income taxpayers for tax preparation services even though they were eligible for free services.²⁰⁶ Reporting also revealed that Intuit made efforts to hide its free products from search engines.²⁰⁷ Intuit has since left the Free File Alliance.²⁰⁸ The future of the program is currently in doubt.²⁰⁹

Our finding about taxpayers' indifference cuts both ways. Taxpayers wouldn't mind using a private service, but they also wouldn't mind the government providing tax services directly. In other words, mistrust of government does not seem to be an overriding concern in this context.

These considerations might counsel toward the IRS providing tax automation services directly. In 2019, Free File members removed the limitation on the IRS providing tax software, paving the way for a public filing option.²¹⁰ A public filing system could more easily pre-populate simple tax returns using IRS data, eliminating tax preparation work for many while also reducing the risk of making a mistake. And, unlike members of the Free File Alliance, the IRS is not subject to profit maximizing motives and is subject to public oversight.

²⁰⁴ *Id.*

²⁰⁵ See 1st Memorandum of Understanding on Service Standards and Disputes Between the Internal Revenue Service and Free File Alliance, LLC 4 (2005), <https://www.irs.gov/pub/irs-utl/2006-first-ff-mou.pdf>.

²⁰⁶ 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>.

²⁰⁷ *Id.*

²⁰⁸ Justin Elliott & Paul Kiel, *TurboTax-Maker Intuit Will Leave Free Tax Filing Partnership With IRS*, PROPUBLICA (Jul. 16, 2021), <https://www.propublica.org/article/turbotax-maker-intuit-will-leave-free-tax-filing-partnership-with-irs>.

²⁰⁹ See Aravind Boddupalli, *The Uncertain Future of the IRS Free File Program*, TAX VOX (Mar. 9, 2022), <https://www.taxpolicycenter.org/taxvox/uncertain-future-irs-free-file-program>.

²¹⁰ See Addendum to the Eighth Memorandum of Understanding on Service Standards and Disputes Between the Internal Revenue Service and Free File Alliance, LLC 1 (2019), <https://www.irs.gov/pub/irs-utl/FFI%20Signed%20MOU%20Addendum%2012-26-19.pdf>.

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 find 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 find 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 find that the benefits of eliminating the subjective costs of taxation are not much higher for high-income taxpayers with complicated returns than for low- and moderate-income taxpayers with simple returns.

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. Especially as Congress considers the possibility of further 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.

APPENDIX

A. Survey Contents

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 post-tax 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;²¹¹ (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?"

"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 provides 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.

²¹¹ 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). See Che Cheng, Keng-Ling Lay, Yung-Fong Hsu & Yi-Miau Tsai, *Can Likert Scales Predict Choices? Testing the Congruence Between Using Likert Scale and Comparative Judgment on Measuring Attribution*, 5 METH. PSYCH. 100081, 100081 (2021).

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.²¹²

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.²¹³ Each choice card contained four attributes, with the following values and descriptions:

Table 10: Choice Attributes

<i>Time Savings</i>	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)
<i>Risk Elimination</i>	Full elimination of risk (taxes are 100% accurate, no risk of future audit)
	No reduction (no change in accuracy or audit risk)
<i>Provider</i>	Private company
	The government
<i>Fee</i>	\$5
	\$15
	\$25
	\$40
	\$50
	\$100

²¹² See Ian J. Bateman, Matthew Cole, Philip Cooper, Stavros Georgiou, David Hadley & Gregory L. Poe, *On Visible Choice Sets and Scope Sensitivity*, 47 J. ENV. ECON. & MGMT. 71, 90 (2004).

²¹³ *Supra* Part II.A.2.

4. Fractional Factorial Design

Each choice card set includes 2 hypothetical services with varying attributes as well as the option to choose the status quo. Because there are 4 attributes with 4, 2, 2, and 6 levels, respectively, there are $(4 \cdot 2 \cdot 2 \cdot 6)^2 = 9216$ 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” design²¹⁴ that presented a subset of the possible choice card sets. Doing so reduces the number of respondents needed and maximizes statistical power.

To select the subset of choice card sets, we employed Expert Choice, a package developed for R statistical software.²¹⁵ 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%.²¹⁶ 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. 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:

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

²¹⁴ 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. OF PHARMACOECON. & OUTCOMES RSCH. 319 (2002).

²¹⁵ JED STEPHENS, THEORETICAL INTRODUCTION TO EXPERTCHOICE (2010), https://cran.r-project.org/web/packages/ExpertChoice/vignettes/include_theory.pdf; Jed Stephens, *Expert Choice*, GITHUB (Apr. 4, 2020), <https://github.com/JedStephens/ExpertChoice>.

²¹⁶ We utilized the Expert Choice setting that applies the widely used Street/Burgess fractional factorial design. See F. Reed Johnson et al., *Constructing Experimental Designs for Discrete-Choice Experiments*, 16 VALUE IN 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 PHARMACOECON. 661, 669 (2008).

That is, the Johnson-Orme sample sizing method suggests 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 1708 discrete choices, we substantially exceed 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 receive one version of the following question:

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].

Section II.B.3 discusses the qualitative responses in more detail. Word clouds summarizing the most common phrases in the qualitative responses are available in the Online Appendix. 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.

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

B. Attrition and Rejected Respondents

Like any survey, our discrete choice experiment relies 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:

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 is 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.

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.

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:

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)	No Change (continue to spend 10 hours, no change in accuracy or audit risk)
Risk Elimination	Choice Not Available (please select the "Neither Service" option)	Choice Not Available (please select the "Neither Service" option)	
Provider	Private Company	Government	
Fee	\$15	\$25	

Your choice: Service 1 Service 2 Neither

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

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).

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].”

“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 quite 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 11 shows, our final sample had 244 responses out of 475 completed responses.

²²¹ See IAN J. BATEMAN ET AL., ECONOMIC VALUATION WITH STATED PREFERENCE TECHNIQUES: A MANUAL (2002) (discussing strategies for identifying protest responses through open-ended questions).

Table 11: Attrition Statistics

Statistic	All	Pre-Screened, Started Survey	Completed Survey	Passed Comprehension Check	Passed Attention Check	Non-Protester
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
Hispanic	0.135	0.125	0.118	0.091	0.084	0.086
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

C. Discrete Choice Model

To model respondents' discrete choices regarding whether to purchase the tax services offered to them, we specify 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 generate 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 bias²²²), $Hours$ is the number of hours that individual i spends on tax compliance, Fee is the cost in dollars of the package of services the individual buys to reduce tax compliance burdens, and $Risk$, $Time_1$, $Time_{half}$, $Time_{all}$, and $Government$ are indicator variables corresponding with the various attributes:

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

We presented each respondent with three discrete choices at once: two

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

alternative packages of tax services, or the status quo (no service). We use the conditional logit model described in McFadden (1973) to predict the probability that respondent i selects any alternative g (including the status quo) as²²³:

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

Given this model, we can calculate 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 formula²²⁴:

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

As an additional alternative, we can model the interaction between individual respondent characteristics and choice attributes by using interaction terms.²²⁵ As discussed in Section II.B.1, we analyzed a wide variety of demographic and respondent-specific variables to see whether specific groups had different implicit prices for the services offered. Formally, given any respondent-specific variable $Demographic_k$, we can model the impact of that variable as:

²²³ See ZSOLT SÁNDOR, COMPUTATION, EFFICIENCY AND ENDOGENEITY IN DISCRETE CHOICE MODELS 8, equation 2.5 (2001), <https://pure.rug.nl/ws/portalfiles/portal/3083715/c2.pdf>; Atkinson & Mourato, *supra* note 73, at § 5.2 (describing an identical formula, assuming that the scale parameter μ is set to 1).

²²⁴ *Id.* We calculate standard errors and confidence intervals for implicit prices using the delta method. See *How Can I Estimate the Standard Error of Transformed Regression Parameters in R Using the Delta Method?*, UCLA, <https://stats.oarc.ucla.edu/r/faq/how-can-i-estimate-the-standard-error-of-transformed-regression-parameters-in-r-using-the-delta-method> (last visited Dec. 31, 2022).

²²⁵ For theory behind interaction terms in discrete choice experiments, see MANDY RYAN ET AL., USING DISCRETE CHOICE EXPERIMENTS TO VALUE HEALTH AND HEALTH CARE 47-50 (2008). For examples in practice, see Jianhua Wang et al., *Urban Chinese Consumers' Willingness to Pay for Pork with Certified Labels: A Discrete Choice Experiment*, 10 SUSTAINABILITY 603 (2018); Margaret E. Kruk et al., *Rural Practice Preferences Among Medical Students in Ghana: A Discrete Choice Experiment*, 88 BULL. OF WORLD HEALTH ORG. 333 (2010).

$$\begin{aligned}
V_{ij} = & \beta_0 \cdot ASC + \beta_1 \cdot Risk_j + \beta_2 \cdot Time_{1j} + \beta_3 \cdot Time_{halfj} + \beta_4 \\
& \cdot Time_{allj} + \beta_5 \cdot Government_j + \beta_f \cdot Fee_j \\
& + \sum_k (\beta_{1k} \cdot Risk_j \cdot Demographic_{ik} + \beta_{2k} \cdot Time_{1j} \\
& \cdot Demographic_{ik} + \beta_{3k} \cdot Time_{halfj} \cdot Demographic_{ik} \\
& + \beta_{4k} \cdot Time_{allj} \cdot Demographic_{ik} + \beta_{5k} \cdot Government_j \\
& \cdot Demographic_{ik} + \beta_f \cdot Fee_{jk} \cdot Demographic_{ik})
\end{aligned}$$

D. Regression Table

Table 12 below summarizes the results of the main effects regression, excluding respondent-specific interaction terms. Additional information, including the full regression table used to generate monetary benefit

estimates, is available in the Online Appendix.

Table 12: Regression Table - Main Effects

Attribute	Coefficient
<i>Risk</i>	2.154*** (0.110)
<i>Time₁</i>	-0.263* (0.149)
<i>Time_{half}</i>	0.302** (0.146)
<i>Time_{all}</i>	1.592*** (0.141)
<i>Government</i>	0.055 (0.072)
<i>Fee</i>	-0.030*** (0.002)
<i>ASC_{status quo}</i>	0.263 (0.195)
<i>ASC_{option 2}</i>	0.081 (0.075)
Observations	1694
Log Likelihood	-1308.060
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

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 includes 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 is correct, the subset regression will 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.²²⁶

We conduct 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 13: Hausman-McFadden Test – Subset Excluding Risk

χ^2	-554.89
degrees of freedom	7
p -value	1

Table 14: Hausman-McFadden Test – Subset Excluding Time

χ^2	-61.613
degrees of freedom	5
p -value	1

Table 15: Hausman-McFadden Test – Subset Excluding Government

χ^2	-0.57904
degrees of freedom	7
p -value	1

Each of the tests produces a p -value of 1. Based on this p -value, we fail to reject the null hypothesis, suggesting that the IIA assumption holds. Each of the tests also produces a negative χ^2 value, which Hausman and McFadden argue is evidence that the IIA assumption has not been violated.²²⁷

²²⁶ See generally Jerry Hausman & Daniel McFadden, *Specification Tests for the Multinomial Logit Model*, 52 *ECONOMETRICA* 1219 (1984) (detailing the Hausman-McFadden test).

²²⁷ *Id.* at 1226.