

Blesse, Sebastian; Buhlmann, Florian; Doerrenberg, Philipp

Article — Published Version

I'm a taxpayer ... get me out of here? Evidence on attitudes towards simplifying the tax jungle

Economica

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Blesse, Sebastian; Buhlmann, Florian; Doerrenberg, Philipp (2025) : I'm a taxpayer ... get me out of here? Evidence on attitudes towards simplifying the tax jungle, *Economica*, ISSN 1468-0335, Wiley, Hoboken, NJ, Vol. 93, Iss. 369, pp. 332-360, <https://doi.org/10.1111/ecca.70020>

This Version is available at:

<https://hdl.handle.net/10419/335640>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

I'm a taxpayer ... get me out of here? Evidence on attitudes towards simplifying the tax jungle

Sebastian Blesse¹ | Florian Buhlmann² | Philipp Doerrenberg³

¹Leipzig University, KOMKIS, ZEW Mannheim and Ifo Institute

²ZEW Mannheim and University of Mannheim

³University of Mannheim, CESifo, IZA and ZEW

Correspondence

Sebastian Blesse, Leipzig University, Universitätsstrasse 16, 04109 Leipzig, Germany
Email: blesse@wifa.uni-leipzig.de

Abstract

Why are tax systems complex despite a conventional wisdom that simplification is desirable? We study attitudes towards tax simplification using new survey and experimental data. We first document that most respondents generally desire tax simplification. Using various survey techniques (including survey experiments and open-ended questions), we show that people are not fully aware of the objectives, implications and trade-offs involved in tax simplification. Overall, our findings suggest that people generally support tax simplification, but have preferences to keep certain complexity-adding elements of the tax system, particularly those that compensate for circumstances (rather than choices). Our results imply that tax complexity is generally not desired *per se*, but potentially arises as an unintended consequence of many individually demanded tax expenditures.

KEYWORDS

attitudes towards tax simplification, randomized survey experiment, tax complexity

JEL CLASSIFICATION

H2; D72; C9

1 | INTRODUCTION

Should tax systems be simplified? The conventional wisdom seems to be: *yes*, tax systems should be simpler!¹ Many economists advertise tax reforms that make the system less complex, and the conventional wisdom among politicians and journalists also holds that simplifying taxes is generally desirable.² However, there may be a trade-off, and there also are economic arguments in support of a certain level of tax complexity;³ some legal scholars also argue that complexity may be necessary for equity reasons, and to provide certainty in legal application (Paul 1997). Although such arguments feature less prominently in the debate, and simplification

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Economica* published by John Wiley & Sons Ltd on behalf of London School of Economics and Political Science.

is a commonly stated objective, the data show that many tax systems remain very complex and tend to become even more complex over time (e.g. OECD 2010b; Benzarti and Wallossek 2024).

One potential explanation for this puzzle is that attitudes towards tax simplification among the general population are more nuanced than they seem on first glance. In this paper, we aim to shed light on how the general public perceives the complexity of personal income taxes. Our analysis first documents and characterizes public attitudes towards tax complexity, and subsequently studies the role of awareness regarding the trade-offs involved in tax simplification. Our results also shed light on the questions of which tax system the general public favours and why tax systems remain so complex.

To address our research objectives, we collect new survey and experimental data among a representative sample of the German population. Our large set of survey questions is embedded in the *German Internet Panel*, a probability-based survey of the German population ($N = 2432$; see Section 2 for details). The concept of tax complexity is complex in itself. To generate a setting that is comprehensible to survey respondents, we focus on one particular dimension of tax complexity, namely tax expenditures, that is, deductions and exemptions in personal income tax declarations. While there are clearly more dimensions of tax complexity, tax expenditures are a main source of tax complexity and a major issue in the debate about complexity of personal income taxation. Germany is an ideal laboratory in which to study opinion formation on tax simplification since it is a country with a fairly complex income tax system and no pre-populated tax returns, implying that all expenditures have to be explicitly filed in the tax return.⁴

We begin our analysis with a documentation of tax complexity attitudes (Section 3): in a situation in which respondents reflect on tax simplification without any further context, more than two-thirds of respondents express a strong general preference to simplify the tax system. We further document a large heterogeneity among respondents with respect to the preferred type of complexity-reducing tax reform.

We then move on from the finding that a majority supports tax simplification, and use four distinct strategies to study if the matter becomes more multilayered as we provide more context and highlight the trade-offs behind simplification.

First (Subsection 4.1), we elicit if people are in favour of specific complexity-adding tax expenditures. We have a series of questions in which we present the living situations of two fictitious taxpayers, and then survey if any of the two should pay lower taxes than the other one. The two fictitious persons are similar in all tax-relevant means, except for one aspect of their living situation: (i) the necessity to spend money on the care of an elderly family member; (ii) the amount they donate to charity; or (iii) the distance between their home and workplace. The results show for all three scenarios that a considerable fraction of respondents indicate that the two persons should *not* pay the same amount of taxes, and that the taxpayer with the additional cost burden should pay less. The share of people who indicate in *all three* scenarios that the tax system should *not* account for differences in living situation (and thus be as simple as possible) is 29.8%, which is considerably smaller than the share of respondents who generally support simpler taxes. Interestingly, we further find that 70% of those respondents who, in an earlier survey question, say that they wish to abolish *all* expenditures indicate at least once that the two fictitious taxpayers should *not* pay the same amount of taxes. Treating the two otherwise identical taxpayers differently requires the existence of tax expenditures, suggesting that many people prefer a tax system with a certain degree of complexity. Our findings are indicative of the notion that people are not fully aware of the implications and involved trade-offs behind simplification.

The results further show that the 'elderly care' scenario induces more participants to vote for differential taxation than the other two scenarios. This is interesting in light of the fact that the costs for elderly care are circumstantial (i.e. outside of the control of taxpayers), while the other two are choices. Our respondents thus exhibit preferences that are consistent with the rationale in the optimal taxation and equality of opportunity literature. We designed an additional follow-up survey experiment to shed more light on the role of choices and circumstances

(Subsubsection 4.1.1). This exercise confirms that survey respondents are more willing to add complexity-adding expenditures if these expenditures compensate for circumstances, rather than choices.

Second (Subsection 4.2), we implement two (independently) randomized experiments to study if preferences for tax simplicity are elastic to arguments in favour of and against tax simplification (one experiment towards the beginning of the survey, and one towards the end). In each of the two experiments, we have three randomized groups: (i) a control group that is exposed to a neutral statement about the debate about tax complexity (which is also shown to each treatment group); (ii) one treatment group in which preferences are possibly shifted towards stronger support for tax simplification; and (iii) one treatment group in which preferences are possibly shifted towards less support for tax simplification. The arguments that we use to shift preferences in the four treatment groups are inspired by some of the most debated issues in the context of tax complexity: (i) redistribution and social policy aspects; (ii) tax avoidance and evasion; (iii) efficiency; and (iv) lobbying and special interest groups.

We find the following consistent pattern across both randomized experiments: the support for simplification is elastic to arguments *against* simplification, while arguments *in favour of* simplification do not have an effect, suggesting that people are less aware of arguments against simplification. The null results in response to arguments in favour of simplification testify to more existing awareness and fewer misperceptions in the context of pro-simplification arguments.⁵

Third (Subsection 4.3), we explore if exposure to the topic affects the support for tax simplification. We find that the support for tax simplification declines between the first and second elicitations of simplification preferences (despite the known tendency that survey respondents aim to give consistent replies). Importantly, we find this pattern among respondents who are in the control group in both of our survey experiments. Since these respondents are never exposed to any arguments relating to tax complexity, the falling support among these respondents can be attributed to the mere fact that they replied to questions related to tax simplicity and were thereby exposed to the topic. This finding is consistent with other findings in our paper that awareness reduces preferences for tax simplicity. In addition, it provides evidence that the experimental treatment effects are not simply caused by a mechanical re-weighting towards the respective argument to which people are exposed in the experimental intervention.

Fourth (Subsection 4.4), we use open-end questions (in the absence of any priming; Ferrario and Stantcheva 2022) in a follow-up survey to examine the extent to which people (who were never exposed to any information treatments) are aware of the objectives of existing tax expenditures. While a substantial share of respondents indicate that they are not aware of any objectives ($\approx 19.5\%$), the most frequently mentioned objective (22.1%) is that deductions allow a reduction in the personal tax bill. These findings support the view that many people are not fully aware of the arguments against and in favour of tax complexity.

We identify the following main contributions of our paper. We show that while most people express a general preference for more tax simplification, they realize that simplification comes at a cost when they are faced with its concrete implications.⁶ To the best of our knowledge, we are the first paper to show within a single approach that people are not fully aware of the trade-off between tax simplicity and tax complexity, and that the support for tax simplification depends on awareness about its consequences. This relates to literature emphasizing the potential costs of simplification (Hines 2016; OECD 2010a) and to papers on the political economy of taxes and feasibility of tax reforms (e.g. Brennan and Buchanan 1980; Meltzer and Richard 1981; Bierbrauer *et al.* 2021).

Our findings provide a potential answer to the puzzle that tax complexity remains high in the real world despite the seemingly wide support for tax simplification: not fully aware of the overall implications of (complexity-enhancing) tax expenditures, people support the existence of various single tax expenditures. As a result, there may be an unintended amount of tax expenditures, and thus complexity, in the existing system. In other words, the demand for single expenditures creates

an externality in the form of more tax complexity that people do not internalize. To this end, we relate to Hettich and Winer (1988), who model the existing tax system with several expenditures as the result of a political process and a government that maximizes political support.⁷

Our results imply that people support tax-simplifying tax reforms, but they do not necessarily desire the simplest possible tax system, and have preferences to keep certain complexity-adding elements of the tax system. In particular, relating to optimal taxation theory (e.g. Alesina and Angeletos 2005; Durante *et al.* 2014; Ooghe and Peichl 2015), we contribute the finding that people prefer maintaining expenditures that compensate for circumstances rather than choices.

2 | THE SURVEY

2.1 | The German Internet Panel

Our questions are embedded in the *German Internet Panel* (GIP), a longitudinal survey that is operated and administered at the University of Mannheim in Germany.⁸ GIP data are collected online on a bi-monthly basis. The survey is probability-based and representative for the German population aged 16 to 75 (see Blom *et al.* (2015)). for more details on the representativity of the GIP). Recruitment was conducted offline with face-to-face interviews, during which respondents were invited to the online panel. To ensure the representativeness of the sample, the GIP includes respondents without prior computer or internet access by providing them with the necessary equipment and training (Blom *et al.* 2017).

The survey includes repeated questions (included in every wave) as well as questions included only in single waves. We included a block of questions in wave 36, which went to the field in July 2018 and included 2464 participants (2432 of which replied to our key survey question of interest). Summary statistics are presented and discussed in Online Appendix A.

We also report the results of a follow-up survey fielded in March 2020 (GIP wave 46) at several points in the paper. These subsequent survey questions include an additional randomized survey experiment (to shed light on the role of circumstances) and an open-ended question on the purpose of existing tax expenditures. For reasons of brevity, we here describe the structure and content of our main (initial) 2018 survey. The details of the follow-up survey questions are described as we mention them in the text (in particular in Subsubsection 4.1.1), in the notes to the figures that summarize the follow-up results (Figure 5 and Online Appendix Figure B.5), and in Subsection 4.4. We always refer to the 2020 follow-up survey as the ‘follow-up survey’ when we report its results, and refer to the main survey if not mentioned explicitly otherwise.

2.2 | Main objectives of the survey

We designed a block of ten survey questions, and integrated these questions into the GIP.⁹ The ten questions are designed to speak to the two main objectives that we define below.

2.2.1 | First objective

First, we aim to document preferences for tax simplification among a representative sample of the population, and to understand further related aspects of tax simplification. Our *ex ante* expectation is that the support for tax simplification is fairly high. This expectation builds on the

observation that both the public debate and the professional discussion (in academic literature and the media) are centred around critiques about overly complex tax systems and proposals to simplify taxes, while economic arguments in support of a certain degree of tax complexity play a less prominent role. Although the public and policy debates incline us to expect that the support for simplification will be fairly high, it is not clear *ex ante* how high exactly the support is, especially because of the existence of good arguments for keeping certain complexity-adding components in the tax system.

2.2.2 | Second objective

Second, we aim to understand if the (presumably large) overall support for tax simplicity is driven by a lack of awareness about the implications and consequences of tax simplification. The respective results can then shed light on the question of why tax complexity is prevalent, although many people do not like it. Considering all our findings together (first and second objectives), we further aim to shed light on the question of which type of tax system the general public favours.

We employ four strategies to study the role of awareness.

In our first strategy (questions Q6–Q8 in the survey, see below), we force people to reflect on concrete tax-relevant situations that are potentially familiar to them, then let them decide how the tax system should address these situations. Because the debate about complexity seems to lean towards simplification, we deliberately confront individuals with scenarios that they do not immediately relate to the debate about complexity, and deliberately do not mention to them that their decisions could have implications for tax complexity. This allows us to investigate individuals' preferences towards complexity-adding components in the tax system in the absence of the public-discussion-induced prejudices towards more tax simplicity.¹⁰

The second strategy builds on randomized survey components that expose participants to arguments against and in support of tax simplification (questions Q2 and Q9). Altogether, we run two information experiments in our main survey. The rationale for this approach is that if the provided arguments shift individuals' preferences, then individuals are not aware of certain aspects of tax simplification. We hypothesize that misperceptions and a lack of awareness are less prevalent in the context of arguments *against* tax simplification, implying that shifts in preferences are presumably larger in response to arguments against simplification than to responses in support of simplification. This hypothesis builds on the observation that both the public debate and the professional discussion focus on tax simplification.

Since we elicit preferences towards tax simplification twice—towards the beginning and towards the end of the survey—a third strategy explores if the support for tax simplification changes over the course of the survey (i.e. compare questions Q2 and Q9). Focusing on subjects who are in the control group in both experiments and who are therefore never exposed to any arguments in favour or against simplification, this (within-variation) strategy allows us to study whether the mere exposure to questions on tax complexity—and a resulting higher reflection and engagement with the topic—affects preferences. A decrease in preferences for tax simplification over the course of the survey would imply that a lack of awareness about the topic indeed explains (at least a fraction of) the initially high support. Considering the public debate about the topic, we hypothesize that the support for simplification is lower at the end of the survey than at the beginning.

The fourth strategy is to use open-ended survey questions in a separate survey module to elicit what people believe to be the objectives of existing tax expenditures. This analysis enables us to study how people reflect on how complexity-adding tax expenditures work and why they exist in the first place.

2.3 | Survey structure and questions

Our survey questions and the respective reply categories are shown in Online Appendix Subsection G.1.¹¹ We provide an overview of the survey flow in Figure 1, and describe its structure (in chronological order) below.

- *Introduction.* Opener stating that the next set of questions will be about the tax system in Germany, and in particular about whether the German income tax system is complicated or easy to understand. The opener also states how we define tax complexity in our context, and includes a general statement that the degree of complexity particularly depends on the number of possible tax expenditures.¹² This latter statement thus explains to participants which dimension of complexity we are particularly interested in. The opener also makes all respondents, independent of treatment status, aware of the topic, and it ensures that the topic is made equally salient to all respondents and that all respondents apply the same definition of tax complexity.
- *Q1: Difficulty of filing a tax return.* We ask participants how difficult they find it to file their tax return. We use this question to derive a proxy for the perceived difficulty of the tax system and to investigate whether other questions and treatment responses depend on the degree of perceived tax complexity.
- *Randomized experiment 1.* Participants are randomly assigned to three groups that are exposed to different arguments in the context of tax simplification. See Subsection 2.5 below for more information.
- *Q2: Preferences for tax simplification.* We ask participants (in all treatment groups) if they think the income tax system in Germany generally needs to be simplified. This question elicits preferences for tax simplification and also is the outcome variable in the analysis of the effects of the randomized argument provision.

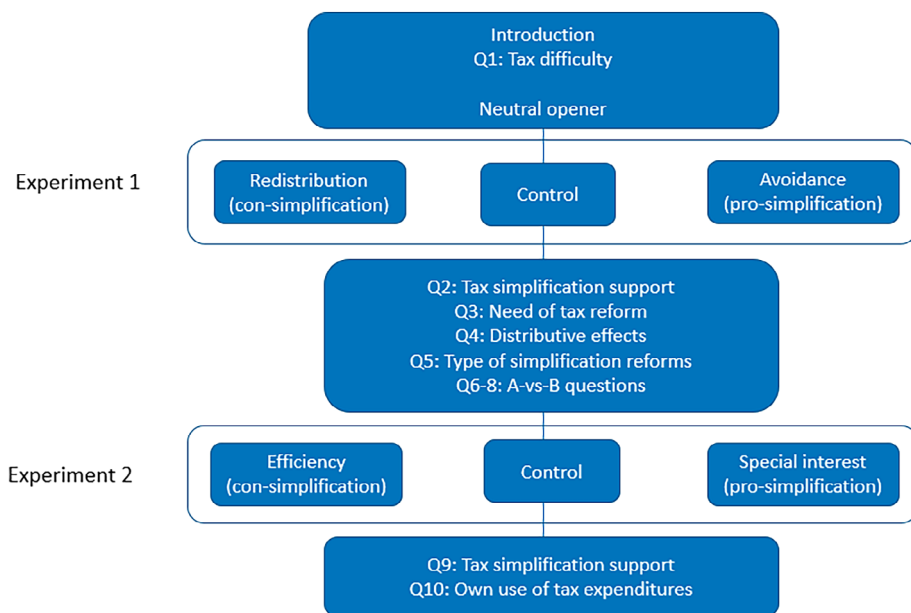


FIGURE 1 Design and flow of main survey (including experiments). *Notes:* This figure illustrates the design and the flow of the main survey, including its randomized components. Experimental groups on the left are con-simplification arguments. Groups on the right represent pro-simplification arguments. Source: Own compilation.

- *Q3: General need for tax reform.* This question surveys if participants in general think that the German tax system is in need of reform.
- *Q4: Distributional implications of tax expenditures.* This question measures participants' beliefs about the distributional effects of tax expenditures. We particularly survey whether someone believes that tax expenditures contribute to a fairer distribution of income, or that they tend to benefit high-income taxpayers.
- *Q5: Which type of tax-simplifying reform?* We offer different types of tax reforms that potentially contribute to simplification of the tax system. Participants are asked to indicate which type of reform they prefer, under the assumption that all the listed reforms are revenue-neutral.
- *Q6–Q8: Should the tax system account for differences in living situations?* We present to respondents the living situations of fictitious taxpayers, and survey who of the two should pay more in taxes. See Subsection 2.4 below for more details.
- *Randomized experiment 2.* Participants are again randomly assigned to three groups that see different arguments in the context of tax simplification (renewed randomization). See Subsection 2.5 for more information.
- *Q9: Preferences for tax simplification.* We again elicit preferences for tax simplification (as in Q2). We explain to participants that we ask the same question again because the topic was the subject of some of the previous questions. The question serves as an outcome variable for the second set of experiments. We also use this question to study how support for simplification evolves over the course of the survey.
- *Q10: Own use of tax expenditures.* We survey which tax expenditures participants usually make use of in their own annual income tax declaration.

2.4 | Different tax burdens for taxpayers in different living situations?

In questions Q6–Q8, we present to survey respondents the living situations of two fictitious taxpayers, *A* and *B*. (For similar survey question techniques, see Weinzierl 2014, 2017; Saez and Stantcheva 2016; Fisman *et al.* 2020). We ask respondents in all three questions to imagine two fictitious taxpayers, *A* and *B*, who are comparable in all tax-relevant aspects, and differ only along one of the following dimensions.

- Person *A* has to spend a considerable amount for the care of her elderly mother, while person *B* does not have to bear such costs (Q6).
- Person *A* spends a considerable amount of income on charitable giving, while person *B* does not donate (Q7).
- Person *A* has to travel a considerable distance to work, while person *B* lives close to work (Q8).

We then ask who of the two fictitious persons should pay more taxes (where the order of reply categories and the order of presented scenarios were randomized).

The choice of these three dimensions is motivated by their economic importance and real-world prevalence. Tax expenditures for commuting, charitable giving and elderly care are substantial and large in size, with commuting expenditures being the largest income-related expense recognized by tax authorities in Germany. In 2017, tax expenditures related to commuting alone amounted to about 5 billion euros (Ramthun 2017). Moreover, these items are used by many taxpayers; about 35% of the taxpayers use the commuting expenditure, about 24% of the taxpayers deduct charitable contributions from their tax base, and approximately 8% of the taxpayers use the elderly care deductions (own calculations based on German administrative tax

records (FAST)).¹³ Overall, the tax expenditures that we chose to rely on in these questions are very important, salient and likely to be familiar to most of our survey respondents. In addition, these three tax expenditures represent three different rationales for tax expenditures—namely (i) circumstances that are outside of the control of the respective taxpayer (elderly care), (ii) positive externalities associated with the (self-chosen) expenditure (charitable giving), and (iii) items representing job-related choices of taxpayers (commuting)—allowing us to study if the share of people who believe that circumstantial living situations should be accounted for in the tax system is different from the share of people who believe that chosen living situations should be accounted for.

2.5 | Randomized survey experiments

We include two randomized components into our survey block on tax simplification (before questions Q2 and Q9). The two experiments are preceded by separate randomization processes. In both experiments, respondents are randomly assigned to either a control group or one of two treatment groups (i.e. between-subjects design with three groups; see the tree in Figure 1). Therefore participants can be either not treated or treated once or twice. Overall, we have nine comparison groups, but refer for the most part to the respective effects of the treatment groups to the never-treated control group.¹⁴

Both experiments are structured in the same way.

- (i) We first have a short opener that serves as a connecting passage to the subsequently provided arguments. Importantly, all participants (control group and treatment groups) see this opener, ensuring that the issues of tax expenditures and complexity are made equally clear to control group participants and treatment group participants. In addition, the opener ensures that respondents in all three experimental groups (including control) apply the same definition of tax complexity. Any treatment effects are therefore not driven by differences in the extent of topic awareness across the groups, or differences in the definition of tax complexity across groups.
- (ii) After the opener, respondents in the two treatment groups are provided with short arguments in the context of tax simplification.
- (iii) Respondents in all three groups move on to the next survey question.

The arguments that we provide in the treatments pick up some of the most frequently debated issues and empirical findings in the context of tax simplification. In each of the two experiments (which are not augmented), one treatment aims to shift preferences towards tax simplification, and the other aims to shift preferences away from tax simplification. Our treatments thus reflect that there are arguments both against and in support of tax simplification, and allow us to investigate whether preferences for tax simplification are more elastic with respect to arguments in support of or against tax simplification. We describe and motivate the two experiments in the following. Screenshots of the experimental treatments and the opener statements provided to all experimental groups (including the control conditions) are provided in Online Appendix H.

2.5.1 | Experiment 1

The two treatments in Experiment 1 are preceded by an opener that is shown to all respondents (i.e. both treatment groups and the control group), and that is everything that control group

respondents see in the context of the first experiment before they move to the next survey question. The opener reads as follows:

In Germany, there is an ongoing debate on whether the income tax system is too complicated because of many possible deductions and allowances.

The *Redistribution* treatment highlights that tax expenditures, which add to tax complexity, potentially have redistributive effects and can be used to reduce the tax liability of taxpayers who are disadvantaged by circumstances. The treatment addresses the point that tax expenditures can serve as a social policy measure, and presents a potential argument in support of a certain degree of complexity. The treatment text follows directly after the opener, and reads as follows:

However, it is sometimes also argued that a tax system with many possible deductions and allowances has an important social and redistributive policy role. For example, tax deductions can be used to reduce the tax burden of taxpayers who are disadvantaged by circumstances.

The *Avoidance* treatment highlights the frequently debated point that the existence of many complexity-adding tax expenditures potentially facilitates tax avoidance and evasion.¹⁵ Assuming that most people disapprove of tax avoidance and evasion, the treatment is an argument in favour of tax simplification. The treatment text follows directly after the opener, and reads as follows:

In this context, one argument is that a tax system with many possible deductions and allowances offers greater opportunity for tax avoidance and tax adjustment. For example, when individuals have a better knowledge of the tax system or make unjustified declarations, they can reduce their tax burden by taking advantage of certain allowances or deductions.

2.5.2 | Experiment 2

The two treatments in the second experiment are, again, preceded by an opener that is shown to all respondents (i.e. both treatment groups and the control group). The opener reads as follows:

We would like to once again address the ongoing debate concerning whether the income tax system is too complicated due to the many possible deductions and allowances.

The *Efficiency* treatment highlights the argument that efficiency is potentially higher in a complex system with many tax expenditures, because such a system provides the opportunity to tailor taxes to individual situations, and thus to tax individual capacity and ability. The treatment therefore increases awareness for a potential argument against tax simplification. The treatment is presented immediately after the opener, and reads as follows:

One argument that is often used against tax simplification and that has not been addressed so far is that a tax system with many deductions and allowances provides better opportunities to tax individuals in accordance with their ability to pay, and is therefore economically more efficient.

The *Special-interest* treatment highlights that a complex system with many tax expenditures is more vulnerable to the lobbying activities of special interest groups. The argument is that

special interest groups try to bargain favourable tax treatment, and the existence of many complexity-adding exemptions facilitates the groups' efforts. Provided that most people agree that special interests should not be accounted for in the tax system, this second treatment provides an argument in support of tax simplification. The treatment text is formulated as follows:

One argument that is often used in favour of tax simplification and that has not been addressed so far is that a tax system with many deductions and allowances offers special interest groups greater opportunity for obtaining exemptions for their clientele.

2.5.3 | **Balancedness of experimental groups**

Online Appendix Table A.2 presents the results of balancing checks for our first experiment. Following the strategy in Alesina *et al.* (2018), we test balance across groups as follows. For each covariate, we run three ordinary least squares (OLS) regressions of the form $y_i = \beta \text{Covariate}_i + \varepsilon_i$, where Covariate is the respective covariate that we test. The three dependent variables for which we run the regressions are dummies indicating the treatment groups—redistribution, avoidance and control group. From this procedure, we have the results of 30 OLS regressions (one regression for each combination of 10 covariates and 3 outcome dummies). Reassuringly, we find strong evidence that randomization worked well and our covariates do not systematically predict treatment status. Out of 60 estimated coefficients, only five are significant at the 10% level and only one is significant at the 5% and 1% levels, respectively. This is well in line with these coefficients being significant by chance within their margin of error.¹⁶ The equivalent strategy was used for testing balancedness of the second experiment; see Online Appendix Table A.3, which is structured just as the corresponding table for the first experiment (Table A.2). We here restrict the sample to respondents who were in the control group of the first experiment (because we are interested in the effect of the second experiment for this 'unencumbered' group; see Subsection 4.2 for a more detailed explanation). The results are again reassuring. Out of 60 coefficients, seven are significant at conventional levels of significance (10% or lower). Overall, randomization apparently worked out well, which is not surprising given that the GIP computer system assigned respondents randomly to treatment groups, and selection into groups was not possible. Later in our regressions, we present specifications that condition on all observable covariates to mitigate all remaining concerns regarding balancedness of experimental groups.

2.5.4 | **Discussion of prior knowledge, experimenter demand effects and statistical power**

While we are not able to measure to what extent the provided treatment information constitutes news to the respondents, we can rely on a follow-up survey that we fielded 20 months after the main survey to gain an idea how people generally reason about tax expenditures (see Subsection 4.4 for details). Focusing on those respondents who are in the control groups of both our experiments, and who are thus not affected by the treatment information, we find that a substantial fraction of respondents are not aware of any objectives of expenditures, and many believe that their prime purpose is to reduce the tax burden (i.e. have pocketbook interests). This suggests that people generally have very limited knowledge about the exact purposes of tax expenditures.

One frequently raised concern with survey experiments is that experimenter demand effects drive the survey responses and results. Our main survey question, on preferences for tax simplicity, asks respondents for their view on a specific aspect of policy. There is no correct or incorrect

answer to this question, nor is it in any way ethically critical, nor is the topic ideologically loaded. Participants are therefore not under the impression that they must provide a particular answer, and social desirability bias should thus not be critical here. In addition, the question is very similar to the questions that GIP participants are used to. The arguments that treatment participants receive prior to replying to the tax-simplicity preferences are provided in a neutral and objective way, thus do not induce subjects to provide a certain answer. Overall, the intention behind our survey question and experimental interventions was certainly considerably more subtle than in studies that test for demand effects (and do not find demand effects; see Mummolo and Peterson 2019; de Quidt *et al.* 2018). Furthermore, our empirical findings are very consistent across different survey techniques. It is unlikely that all survey techniques are subject to the same type and degree of experimenter demand effects. In particular, our results using within-subject variation in Subsection 4.3 should not be affected by experimenter demand effects.

With respect to statistical power, the review paper by Haaland *et al.* (2023), suggests that randomized survey experiments should have about 700 observations per treatment arm. The numbers of observations in our experimental groups are in accordance with this suggestion ($N = 2432$, with an even distribution of observations across three conditions).

3 | RESULTS SET 1: DOCUMENTING AND CHARACTERIZING ATTITUDES TOWARDS TAX SIMPLIFICATION

This section documents and characterizes preferences for tax simplification. Note that the figures in the following analyses are based on our entire sample. It is, of course, possible that the reported statistics for the entire sample are partly affected by the arguments and information that we randomly provide to some respondents in the context of our first survey experiment (except for the question of perceived difficulty of filing a tax return, which was asked before treatment exposure). We therefore always report along (in the body of the text) the analogous statistics for control group respondents who did not receive any particular information; that is, control group statistics are not affected by the survey experiment. We decided to report the results for the entire sample in the figures in order to present a complete picture for all respondents in our survey.

3.1 | Preferences for tax simplification

Using a 6-point scale, from 1 ‘Absolutely not’ to 6 ‘Absolutely’, we ask if people believe that the income tax system in Germany needs to be simplified (Q2). As shown in Figure 2, a large majority believe that the tax system needs to be simplified: among all question respondents (i.e. those who gave a non-missing and non-*I-don't-know* reply), about 76% (control group 79%) of the respondents have strong (category 5 or 6) preferences for tax simplification.¹⁷ Only about 8% (control group 7.5%) of the respondents chose categories 1, 2 or 3, which indicate rather weak preferences for tax simplification. The mean response across all respondents is 5.16 (control group 5.21). Overall, the results provide clear evidence of strong preferences for tax simplicity.

3.2 | Which simplifying tax reform?

To understand which tax system people desire, we include a question to survey *how* policy should reform the tax system to make it simpler (Q5). We provide respondents with a list of potential

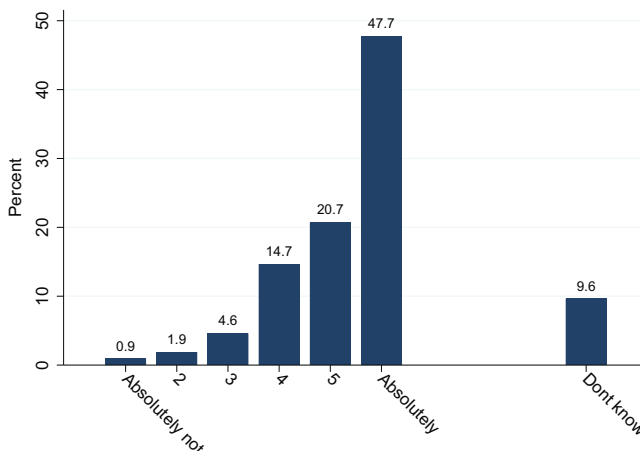


FIGURE 2 Preferences for tax simplification. *Notes:* This figure depicts the percentage shares of respondents in the respective categories for the question ‘Do you generally think that the income tax system in Germany needs to be simplified?’ (Q2). Respondents could pick one of the following categories: 1 Absolutely not; ... ; 6 Absolutely; I don’t know. The figure is based on 2423 non-missing observations. The mean answer is 5.16. Source: Own calculations based on the GIP.

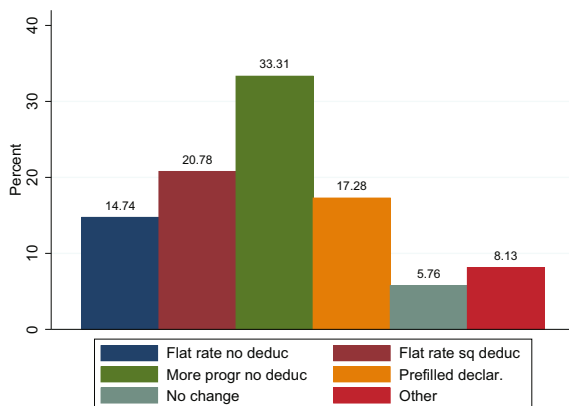


FIGURE 3 Which revenue-neutral reform? *Notes:* This figure depicts the percentage shares of respondents in the respective categories for the question: ‘Which of the following measures to simplify the income tax system would you like the most? Assume the proposed measures will lead to unchanged tax revenues in each case.’ (Q5) Respondents could pick one of the following categories: Same rate for all but no deductions and allowances; Same rate for all and same deductions and allowances as under current system; More progressive tax rates and no deductions and allowances; Automatic determination of amounts in income tax declaration; No change; Other measure [insert text]; I don’t know. The figure is based on 1771 non-missing observations. Source: Own calculations based on the GIP.

tax-simplifying reforms and they can choose which of the offered alternatives they prefer. While this list is, of course, not exhaustive—for instance, we did not ask participants whether they would prefer a reform where only some but not all tax expenditures could get abolished or be kept, which is the focus of Q6–Q8 as shown in Subsection 4.1—the list features some of the most debated type of reforms. Respondents are explicitly informed that they should consider each of the reforms under the assumption of tax revenue neutrality in order to abstract from revenue considerations.

Figure 3 shows that the most frequently chosen type of reform at 33% (control group 33%) increases the progressivity of the tax system and abolishes all types of tax expenditures. About

21% (control group 22%) of the respondents would prefer a flat-rate system that features the same amount of tax expenditures as in the status quo. Some 15% (control group 15%) of respondents also want a flat-rate system, but without any possibilities for deductions or allowances. About 17% (control group 17%) of the respondents prefer a different type of tax simplification. Instead of reforming rates or the amount of tax expenditures, they prefer to change the tax-filing process through pre-filled tax returns that require less effort to file. Just 6% (control group 6%) of our respondents have a preference for keeping the status quo and not implementing any tax reform.

Overall, the policy support results suggest that there is no consensus with respect to the tax-simplifying reform to be implemented, and that there is heterogeneity in the preferred approach for moving towards a simplified tax system. These observations add to an explanation for the puzzle that real-world tax systems are complex although the conventional wisdom holds that simplicity is strongly desirable. We also investigate if the policy reform preferences are affected by our first experimental intervention (recall that the second experiment was implemented after the reform survey question), and do not find any evidence that this is the case.

3.3 | Further results and survey questions

Investigating the *anatomy* of simplification preferences (see Online Appendix Table B.1), we find that age and gender are important demographic correlates of simplification preferences. In addition, respondents who find it easy to file a tax return have lower simplification preferences than those who find it difficult. Interestingly, respondents who do not file their tax return themselves, or employ a tax preparer, are more supportive of simplification than those who file themselves and find it easy. These non-filers, however, have lower support for simplification than self-filers who find it difficult to prepare the tax return.

We further observe that about 50% of respondents who file a tax return find it *difficult to file their return* (see Online Appendix Figure B.1). Germans further believe that the tax system in their country is in strong *need of reform* (see Online Appendix Figure B.2). We also survey beliefs about the *distributional implications of tax expenditures*, and observe that the large majority believe that tax expenditures mostly benefit richer taxpayers (see Online Appendix Figure B.3). Finally, the *most frequently used tax expenditures* in our sample are: commuting-to-work allowance, deduction of other type of work expenses, charitable donations, deductible expenses for pension and retirement savings, child allowances, and ‘standard deductions’ (see Online Appendix Figure B.4).¹⁸

4 | RESULTS SET 2: ROLE OF AWARENESS

This section presents the results in the context of our second objective, where we analyse to what extent attitudes towards tax simplification are affected by awareness with respect to the pros and cons of tax simplification using four different strategies (see Subsection 2.2).

4.1 | Different tax burden for taxpayers in different living situations?

Figures 4(a) to 4(c) present the results of the ‘taxpayer *A* versus taxpayer *B* questions’ (Q6–Q8) for all subjects. We again report along the results for respondents who were in the control group of the first experiment.

Figure 4(a) shows that a majority of almost 60% (control group almost 60%) of the survey respondents believe that the costs for elderly care should reduce the tax burden. Almost 40%

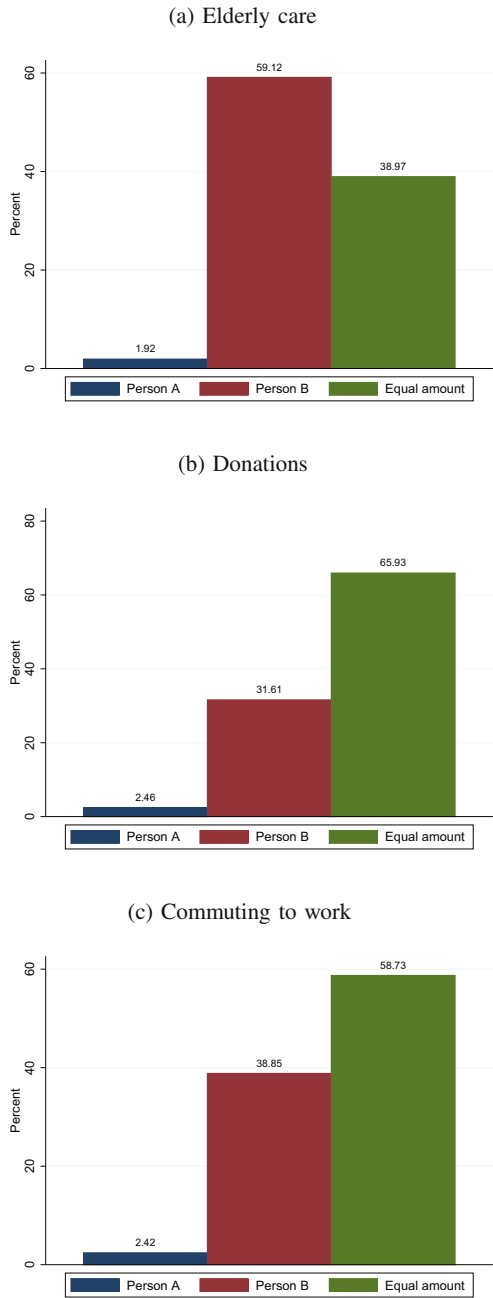


FIGURE 4 Who should pay more taxes?. *Notes:* This figure depicts the percentage shares of respondents in the respective categories for the question. (a) ‘In contrast to person *B*, person *A* has a poor mother in need of elderly care and has to spend a considerable amount of her income for the care of her mother. Person *A* and person *B* have the same gross income and are very similar in all other respects.’ (Q6) (b) ‘Person *A* spends a considerable amount of her income on charitable giving. Person *B* does no such thing. Both person *A* and person *B* have the same gross income and are very similar in all other respects.’ (Q7) (c) ‘Person *A* has to travel a considerable distance to work. Person *B* lives very close to work. Both person *A* and person *B* have the same gross income and are very similar in all other respects.’ (Q8) Respondents could pick one of the following categories (where the order of answer categories was randomized): person *A* should pay higher taxes; person *B* should pay higher taxes; person *A* and person *B* should pay equal tax amounts. The numbers are based on: (a) 2397 non-missing observations, (b) 2398 non-missing observations, (c) 2394 non-missing observations. Source: Own calculations based on the GIP.

(control group 40%) indicate that taxpayers with and without costs for elderly care should pay the same amount of taxes, and almost nobody thinks that person *A* should pay more in taxes. Online Appendix Figure E.1(a) shows that these effects are not driven by self-interest. We split the sample into those who make use of deductions for care costs themselves, and those who do not. The survey responses among these two groups look very similar.¹⁹ Overall, these survey responses provide evidence that a majority favour a system in which the costs for elderly care are deductible.

For the survey questions regarding charitable donations (Figure 4(b)) and expenses for commuting (Figure 4(c)), we see that a majority of about 66% (control group 67%) and 59% (control group 59%) of the respondents believe that both persons *A* and *B* should pay the same amounts of taxes, respectively. However, fractions of 32% (control group 30%) and 39% (control group 39%) of all respondents think that higher donations and commuting costs should imply lower tax burdens. That is, roughly one-third of the respondents believe that differential expenses in these areas should result in a reduced tax burden. This is a considerably smaller share than in the case of elderly care, but it still constitutes a substantial fraction that is in favour of accounting for these living situations in the tax system.

The difference between the ‘elderly care’ situation and the other two situations is interesting. Costs for elderly care are circumstantial and outside the control of the respective taxpayer, while donations and commuting distance are choices of the taxpayer.²⁰ Consistent with the literature on optimal taxation and equality of opportunity (see the Introduction for the related literature), our survey respondents have the intuition that circumstantial differences should be accounted for to a larger extent by the tax system than by deliberate and self-chosen differences. We further explore the role of circumstances and choices in evaluating complexity-adding expenditures in a follow-up survey experiment—see Subsubsection 4.1.1 below.

Consistent with the findings regarding circumstances and choices, we further find that the responses for donations and commuting expenses are more affected by self-interest, as compared to the responses for elderly care. Online Appendix Figures E.1(b) and E.1(c) present the results separately for those who use the respective tax expenditure themselves, and those who do not. In the case of donations, a quarter (control group 23.5%) of those respondents who do not use the donation expenditure themselves think that donations should reduce the tax burden, while this share is 45% (control group 41.2%) among those who do use the donation expenditure. The pattern is similar for the case of commuting expenditures: among those who do not use the commuting expenditure, 34% (control group 34.6%) believe that it should result in a lower tax burden. Among those who do use the commuting expenditure, the share of people who believe that commuting should reduce tax payments stands at 47% (control group 46.7%). While the differences between those who use the respective tax expenditure and those who do not are thus larger in the case of donations and commuting than for the case of elderly care, even for donations and commuting we see that a large share of those who do not use the expenditure support the notion that the tax system should account for the respective living situation. This suggests that the result for none of the three different tax expenditures is entirely driven by self-interest.

We also explore which share of respondents indicate in all three scenarios that the tax system should *not* account for the respective difference in living situation (and thus be as simple as possible). This share is 29.8%, which is considerably smaller than the 90% share of respondents who generally support having a simpler tax system. Online Appendix B reports and discusses the associations for the *A*-vs-*B* exercise and other survey questions.

4.1.1 | Follow-up survey experiment on the role of circumstances and choices

To shed more light on the role of circumstances and choices in evaluating complexity-adding tax expenditures, we conduct an additional randomized survey experiment in a follow-up survey

(see Section 2). In the experiment, we again present two fictitious taxpayers, *A* and *B*, and ask respondents who of the two taxpayers should pay more taxes. The experiment is in the context of commuting subsidies, and the two fictitious taxpayers differ with respect to their commuting distance to work; person *B* always lives close to work, and person *A* has to commute a significant distance. Across three experimental groups, we varied the reasons for why person *A* has to commute. In a control group, we did not specify the reason for person *A*'s long commute. In treatment *Circumstance*, it is specified that 'Person *A* was relocated by his employer and has had to travel a considerable distance to work ever since'. In treatment *Choice*, we specify: 'The possible professional activities in the vicinity of person *A*'s place of residence do not correspond to his preferences and qualifications. Person *A* therefore chooses a job with a very long journey to work.' More details about the survey and experiment are in the Notes to Figure 5.

Figure 5 presents the results, and plots the share of respondents in each experimental group who believe that person *B* (who lives close to work) should pay more taxes than person *A* (who commutes to work). The results provide clear evidence that the reason for the work commute matters significantly to respondents. The share of respondents who think that person *B* should pay more is considerably higher in group *Circumstance* than in the control group, whereas this share is considerably lower in group *Choice* than the control group. Online Appendix Table E.1 shows the regression-based results, and indicates that the respective group differences are statistically

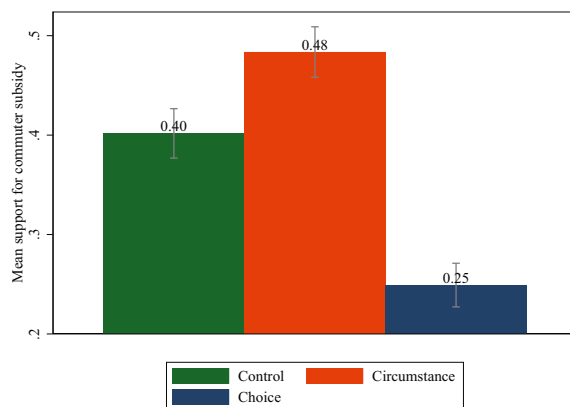


FIGURE 5 Follow-up survey: support for commuter subsidies—the role of circumstances and choices. *Notes:* The figure depicts the average support for commuter subsidies across experimental groups that are based on a follow-up experiment in wave 46 of the GIP in March 2020. The experiment adopted a version of the commuter subsidy question (Q8) discussed in Subsection 4.1 on preferences about different tax burdens for taxpayers in alternative living situations. Specifically, we augment Q8 from wave 36 of the GIP in which our main survey is implemented by two treatment conditions that provide alternative scenarios of living situations with respect to commuting expenses. The question reads: 'Imagine two persons, *A* and *B*. Which person do you think should pay more taxes in the following situation?' The control group receives the following scenario: 'Person *A* has to travel a considerable distance to work, while person *B* lives very close to work. Both person *A* and person *B* have the same gross income and are very similar in all other respects.' The *Circumstance* treatment group receives the following scenario: 'Person *A* was relocated by his employer and has had to travel a considerable distance to work ever since, while person *B* lives very close to work. Both person *A* and person *B* have the same gross income and are very similar in all other respects.' The *Choice* treatment group receives the following scenario: 'The possible professional activities in the vicinity of person *A*'s place of residence do not correspond to his preferences and qualifications. Person *A* therefore decides on a job with a very long way to work, while person *B* lives very close to work. Both person *A* and person *B* have the same gross income and are very similar in all other respects.' Response categories are: person *A* should pay higher taxes; person *B* should pay higher taxes; person *A* and person *B* should pay equal tax amounts. The figure plots the frequency of whether respondents chose the answer that person *B* should pay higher taxes, implying a subsidy for the commuting person *A*. Average support by experimental group is displayed with 95% confidence bars. Total number of observations is 4480, with 1496, 1491 and 1493 observations for the control, *Circumstance* and *Choice* groups, respectively. Source: Own calculations based on the GIP.

significant at the 1% level. Overall, these results thus provide evidence that the evaluation of complexity-adding expenditures depends on the reasons, circumstance or choice, because of which a taxpayer has to bear additional costs.

4.2 | Randomized survey experiments

4.2.1 | Empirical strategy

To analyse the first of the two experiments in the main survey, we use OLS regressions (with robust standard errors) in which we regress the respective outcome variable Y for subject i on dummy variables indicating the two randomly provided arguments. The resulting coefficients then present the effect of the respective treatment relative to the omitted control group. In our preferred specification, we include control variables to improve precision and to alleviate any concerns about potential imbalances across our experimental groups. Specifically, we include the control variables gender, age, marital status, household size, employment status, retirement status, education, perceived difficulty to declare taxes, net household income and political ideology:

$$Y_i = \alpha + \beta_1 \text{Redistribution}_i + \beta_2 \text{Avoidance}_i + \gamma' X_i + \varepsilon_i. \quad (1)$$

The treatments of the first experiment could impact the treatment effects of the second experiment. For example, consider a respondent who was assigned to the con-simplification treatment in the first experiment and to the pro-simplification treatment in the second experiment. A positive effect of the pro-argument in the second experiment might then cancel out the negative effect of the con-argument of the first experiment, and as a result, we see no effect in the second experiment, although there actually is a positive effect. We circumvent this concern in our analysis of the second experiment by estimating the following model:

$$\begin{aligned} Y_i = & \alpha + \beta_1 \text{Redistribution}_i + \beta_2 \text{Avoidance}_i + \beta_3 \text{Efficiency}_i + \beta_4 \text{Special-interest}_i \\ & + \beta_5 \text{Efficiency}_i \times \text{Redistribution}_i + \beta_6 \text{Efficiency}_i \times \text{Avoidance}_i \\ & + \beta_7 \text{Special-interest}_i \times \text{Redistribution}_i + \beta_8 \text{Special-interest}_i \times \text{Avoidance}_i \\ & + \gamma' X_i + \varepsilon_i. \end{aligned} \quad (2)$$

The model fully interacts dummies indicating treatment status of the second experiment with dummies indicating treatment status of the first experiment (the control group always being the reference category), then uses OLS (with robust standard errors) to regress the outcome variable of interest on the full set of interactions. This approach represents the estimation of a long model, including a fully saturated interaction model of cross-cutting experiments, according to Muralidharan *et al.* (2025). We report only the coefficients of the treatment dummies of the second experiment (and not the interactions); that is, the effects of interest in the analysis of the second experiment are β_3 and β_4 , which present the effects of the second experiment for those respondents who were in the control group of the first experiment. These respondents have not received any prior treatment in the context of simplification, and therefore are ‘unencumbered’ when they enter the second experiment.²¹ As with the first experiment, our preferred specifications include control variables that improve the precision of the treatment coefficients of interest.

The main outcome variables are the responses to the question of whether the tax system should be simplified; that is, Q2 in the case of the first experiment, and Q9 in the case of the second experiment. These are the variables that follow immediately after the respective randomized intervention. The variable that we use in the regressions is coded as the original survey question, on a 6-point scale, to not throw away any information. In the context of the first experiment, we further study the treatment effects on the survey question regarding the

perceived distributional effects of expenditures (Q4, coded as the original survey variable, that is, on a 6-point scale). We use OLS for reasons of eased interpretation. Ordered probit models, which account for the discrete and ordered nature of the outcome variables, are presented in robustness checks.

4.2.2 | Experiment 1—main effects

Column (1) of Table 1 shows the effects of the treatment dummies in a regression specification without conditioning on any additional covariates, whereas columns (2)–(5) add several control variables. The regression results in all specifications show negative effects of the *Redistribution* treatment on preferences for tax simplification (all estimates statistically significant at the 5% level). The coefficients are remarkably stable across the five different specifications. In column (5), our preferred specification where we include all covariates, the support for tax simplification is reduced by about 2.6% relative to the control group average ($= -0.133/5.22$). This effect corresponds to a reduction of 13.1% of a standard deviation (i.e. standardized effect). Overall, the regressions thus provide evidence that preferences for tax simplicity are elastic to arguments *against* tax simplification.

The effect size is not very large, but it has to be considered in light of the fact that the overall support for tax simplification is substantial and, given the debate in public and press outlets,

TABLE 1 Experiment 1: effect on preferences for tax simplification.

	(1)	(2)	(3)	(4)	(5)
Experimental group reference category: control					
Redistribution	-0.115** (0.058)	-0.123** (0.058)	-0.133** (0.055)	-0.134** (0.055)	-0.133** (0.055)
Avoidance	-0.042 (0.057)	-0.039 (0.056)	-0.032 (0.053)	-0.032 (0.053)	-0.029 (0.053)
Constant	5.215*** (0.040)	4.610*** (0.149)	3.453*** (0.242)	3.469*** (0.248)	3.456*** (0.249)
<i>N</i>	2190	2132	2109	2109	2109
Demographics	No	Yes	Yes	Yes	Yes
Tax difficulty	No	No	Yes	Yes	Yes
Household income	No	No	No	Yes	Yes
Political preference	No	No	No	No	Yes

Notes: The table presents the effects of the randomized treatment interventions on preferences for tax simplification (Q2). This is estimated by OLS regressions of preferences for tax simplification on treatment dummies. Tax simplification is measured on a 6-point scale based on the question: 'Do you generally think that the income tax system in Germany needs to be simplified?' The experimental groups are control group, redistribution group and avoidance group. Control is omitted, implying that the effects are relative to the control group. All participants receive the following opening statement: 'In Germany, there is an ongoing debate on whether the income tax system is too complicated because of many deduction possibilities and allowances.' Participants in the redistribution group receive the following argument: 'However, it is sometimes also argued that a tax system with many possible deductions and allowances has an important social and redistributive policy role. For example, tax deductions can be used to reduce the tax burden of taxpayers who are disadvantaged by circumstances.' Participants in the avoidance group receive the following argument: 'In this context, one argument is that a tax system with many possible deductions and allowances offers greater opportunity for tax avoidance and tax adjustment. For example, when individuals have a better knowledge of the tax system or make unjustified declarations, they can reduce their tax burden by taking advantage of certain allowances or deductions.' Columns (1)–(5) differ in the included sets of covariates: (1) no covariates; (2) gender, age, marital status, household size, employment status, retirement status and education; (3) column (2) plus perceived difficulty to declare taxes; (4) column (3) plus net household income; (5) column (4) plus political preferences. The scale for the outcome variable is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses.

***, **, * indicate $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.

presumably strongly anchored among respondents. Our treatment thus affects preferences for tax simplification *although* the conventional wisdom on the topic is very clear and strong.

The *Avoidance* treatment does not have a significant effect on simplification preferences. The coefficients are small and not statistically significant throughout the five specifications. The standard errors in all five specifications are considerably greater than the respective coefficient. Statistical precision is thus much weaker than in the case of the *Redistribution* treatment. The coefficients are also considerably different: across all specifications, the coefficients of the *Redistribution* treatment are at least 2.7 times larger than the coefficients of the *Avoidance* treatment, and the difference between the two is statistically significant throughout specifications (3)–(5) (with p -values in the range 0.064–0.057).²² We also tried different specifications of the outcome variable (e.g. a dummy variable indicating very high support for simplification), but never find a significant effect of the *Avoidance* treatment.

Thus we do not find evidence that preferences for simplification are elastic to the argument *in support of* tax simplification. This null result might be explained with the very prominent role of arguments in favour of simplification in the public debate, because of which participants presumably have fewer misperceptions regarding arguments that support simplification. Note that a substantial fraction of control group respondents did not select the highest possible support for simplification (see above), suggesting that there would have been room for the pro-simplification argument to increase the support for simplification.²³

4.2.3 | Experiment 1—robustness and heterogeneous treatment effects

We investigate if the (robust) OLS standard errors that we reported above are robust to other ways of computing standard errors. In particular, we adjust standard errors using (i) randomization tests in the spirit of Fisher (1935), and (ii) tests for multiple comparisons that follow the procedure proposed by Westfall and Young (1993). We present the details and results in Online Appendix D. These exercises show that statistical inference is robust to alternative types of computing standard errors throughout. Results for the experimental effects on both tax simplification attitudes and distributional views are also robust to using ordered probit regressions that account for the discrete nature of the outcome variables; see Online Appendix Tables C.3 and C.4.

Investigating if particular groups of respondents responded differently to the treatments of the first experiment than other groups, we find the effects of the treatments to be homogeneous. We mostly do not see any significant interactions between treatments and demographic variables or variables such as political preferences, difficulty of filing a tax return, trust in government, the perceived quality of tax use for public spending, taste for redistribution (from wave 34 of the GIP), beliefs in luck or effort, and social mobility perceptions (from wave 33 of the GIP).²⁴

4.2.4 | Experiment 2—main effects

The main results for our second randomized intervention are presented in Table 2, which is organized as the corresponding table for the first experiment. Consistent with the results from the first experiment, we observe that preferences for tax simplicity are elastic towards an argument against tax simplification, and not elastic to an argument in favour of simplification. The estimated coefficient for the *Efficiency* treatment is negative and statistically significant throughout all five specifications of the regression table. Column (5), our preferred specification with all covariates, shows that the efficiency argument reduced support for tax simplicity by about 5% ($= 0.240/5.084$), relative to the control group average. This effect corresponds to a reduction of 22.4% of a standard deviation (i.e. standardized effect); that is, the effect is larger than the (standardized) effect that we estimated for the *Redistribution* treatment in the first experiment, and it

TABLE 2 Experiment 2: effect on preferences for tax simplification.

	(1)	(2)	(3)	(4)	(5)
Experimental group reference category: control					
Economic efficiency	-0.197*	-0.216**	-0.229**	-0.237**	-0.240**
	(0.109)	(0.109)	(0.105)	(0.105)	(0.105)
Special-interest	-0.064	-0.062	-0.055	-0.068	-0.067
	(0.097)	(0.097)	(0.094)	(0.095)	(0.095)
Constant	5.084***	4.588***	3.928***	3.993***	3.960***
	(0.066)	(0.160)	(0.222)	(0.232)	(0.232)
<i>N</i>	2187	2134	2114	2114	2114
Demographics	No	Yes	Yes	Yes	Yes
Tax difficulty	No	No	Yes	Yes	Yes
Household income	No	No	No	Yes	Yes
Political preference	No	No	No	No	Yes

Notes: The table presents the effects of the randomized treatment interventions of the second experiment on preferences for tax simplification (Q9). This is estimated by OLS regressions of preferences for tax simplification on treatment dummies and a full set of interactions of the treatment groups of the first and second experiments. Tax simplification is measured on a 6-point scale based on the question: 'Now that we have dealt extensively with various aspects of the German tax system in this survey, we would like to ask again: do you generally believe that the income tax system should be simplified in Germany?' The experimental groups are control group, economic efficiency group and special-interest group. Control is omitted, implying that the effects are relative to the control group. Specifically, we estimate β_3 and β_4 from equation (2) as treatment effects of the economic efficiency and special-interest groups, respectively (unencumbered by potential interactions with interventions in the first experiment). All participants receive the following opening statement: 'We would like to once again address the ongoing debate concerning whether the income tax system is too complicated due to the many possible deductions and allowances.' Participants in the economic efficiency group receive the following argument: 'One argument that is often used against tax simplification and that has not been addressed so far is that a tax system with many deductions and allowances provides better opportunities to tax individuals in accordance with their ability to pay and is therefore economically more efficient.' Participants in the special-interest group receive the following argument: 'One argument that is often used in favour of tax simplification and that has not been addressed so far is that a tax system with many deductions and allowances offers special-interest groups greater opportunity for obtaining exemptions for their clientele.' Columns (1)–(5) all include a full set of interactions of the treatment groups of the first and second experiments. They differ in the additionally included sets of covariates: (1) no additional covariates; (2) gender, age, marital status, household size, employment status, retirement status and education; (3) column (2) plus perceived difficulty to declare taxes; (4) column (3) plus net household income; (5) column (4) plus political preferences. The scale for the outcome variable is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses. ***, **, * indicate $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.

appears non-negligible to us. In addition, the effect size should again be considered in light of the fact that the general wisdom clearly holds that tax simplification is desirable.

The estimates for the effect of the *Special-interest* group argument in favour of tax simplicity are very close to zero and non-significant in all of the regression specifications. Notably, the coefficient of the *Efficiency* treatment is at least three times larger than the coefficient of the *Special-interest* group treatment across the five specifications. However, these differences between the two treatment estimates are not statistically significant, presumably due to power reasons since we compare reactions only for participants who have been in the control group in the first experiment. The null result of the *Special-interest* group treatment is, again, likely to be driven by the more prominent role of arguments in favour of simplification in the public debate, which reduce misperceptions regarding pro-simplification arguments.

4.2.5 | Experiment 2—robustness and heterogeneous treatment effects

Statistical inference is robust to adjusted standard errors using randomization tests and tests for multiple comparisons—see Online Appendix D for details. As shown in Table C.5, the above results are robust to using ordered probit models.

As with the first experiment, we find that effects of the treatments are very homogeneous across different demographic groups and respondents with different characters and preferences. We also study if the effects of the second experimental intervention depend on experimental group status of the first experiment. The results are shown in Online Appendix Table C.6, which displays the effect of the treatments of the second experiment by treatment group status of the first experiment (all results relative to the second experiment's control group). As seen in our main specification, we do find a negative effect of the *Efficiency* treatment for those respondents who were in the control group of the first experiment (and no *Special-interest* treatment effect in this group). As we discussed above, we believe that studying the effect of the second experiment among those who were in the control group of the first experiment is the cleanest approach because they did not receive any treatment before they were exposed to the second experiment. As a result, the effects of the second experiment do not mix and interact with treatments of the first experiment. This approach is supported by the observation in Table C.6 that the second experiment did not trigger statistically significant effects among those who were in one of the treatment groups of the first experiment (with the exception of a positive effect of the *Special-interest* treatment among those who were in the *Avoidance* group of the first experiment in one single specification).²⁵ Despite not being statistically significant at conventional levels, the directions of the respective estimates are plausible in the sense that those who received a con-simplification argument in both experiments are less in favour of tax simplification than those who received it only once. Also, respondents who in both experiments received arguments that are in favour of tax simplification (i.e. *Avoidance* and *Special-interest* group arguments in the first and second experiments, respectively) support tax simplification more than those who received a pro-simplification argument only in the first experiment. Receiving conflicting arguments—that is, an argument in favour of simplification in the first experiment, and a con-argument in the second one (or vice versa)—does not alter the support for tax simplification relative to being exposed to only one argument in favour of tax simplification (or con-tax-simplification) in the first experiment.

4.3 | How does support for tax simplification evolve over the course of the survey?

We now analyse how the support for tax simplification differs between the beginning of the survey (Q2) and the end of the survey (Q9). Table 3 presents the differences in means between Q9 and Q2 for each combination of treatment groups in our two survey experiments. Negative numbers indicate that the support for simplification has declined over the course of the survey; that is, support is lower in Q9 than in Q2. The statistical difference in means between Q2 and Q9 is tested using a *t*-test, and the inference results are presented in parentheses in the table.

Our prime focus is on those survey respondents who were in the neutral control group in both of the randomized survey experiments, because they were not exposed to the provision of any arguments (neither against nor in favour) relating to the desirability of tax complexity. Between reporting their simplicity preferences in Q2 and Q9, they only replied to questions Q3–Q8, and thus were forced to reflect on the topic of tax complexity. We see that the support for simplification is statistically significantly lower towards the end of the survey in Q9 than towards the beginning in Q2. The difference in mean is -0.193 (with standard error 0.055). This result clearly suggests that reflection on the topic reduces the high initial support, and indicates that awareness and reflection matter.²⁶ The difference may not appear enormously large, but should be seen in the context of the following two points. First, in such a within-design where respondents reply to two similar questions twice, it is well known that many respondents try to be consistent across their replies and are averse to deviating much between their two answers. Experimenter demand effects do not seem to play a role in our within-design either (or they are outweighed by the priming effect): we see throughout all three groups of the second experiment that support for tax simplicity

TABLE 3 Support for tax simplification: difference between Q9 and Q2, by treatment combination.

<i>Experiment 2:</i>	Control	Redistribution	Avoidance
<i>Experiment 1:</i>			
Control	-0.193*** (0.055)	-0.114** (0.051)	-0.136** (0.059)
<i>N</i>	218	246	242
Economic efficiency	-0.306*** (0.063)	-0.248*** (0.051)	-0.264*** (0.058)
<i>N</i>	219	214	250
Special-interest	-0.188*** (0.054)	-0.081* (0.049)	-0.224*** (0.060)
<i>N</i>	255	246	219

Notes: The table presents the differences in means in preferences for tax simplification between the beginning of the survey (Q2) and the end of the survey (Q9), by combination of treatment assignment across the two randomized survey experiments. Negative values indicate that preference for tax simplification has decreased on average between Q2 and Q9. Standard errors are estimated via a *t*-test testing for the equality of means of Q2 and Q9. Across all respondents (i.e. without considering the combination of treatment assignments), the difference in means is -0.193 (standard error 0.019, *N* = 2108). We include only respondents who have given a preference in Q2 and Q9. Preference for tax simplification is measured on a 6-point scale based on Q2 'Do you generally think that the income tax system in Germany needs to be simplified?' and Q9 'Now that we have dealt extensively with various aspects of the German tax system in this survey, we would like to ask again: do you generally believe that the income tax system should be simplified in Germany?' The scale for the outcome variables is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses.

***, **, * indicate $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.

decreases more strongly for those who were in the *Avoidance* group in the first experiment relative to those who were assigned to the *Redistribution* group of the first experiment. Experimenter demand bias would work in the opposite direction. Second, as we show in the first part of our paper, the support for simplification is very large, and the entire debate tends to focus on tax simplification. The support for simplification therefore likely roots deeply in people's mindsets, thus even small shifts in preferences testify that mere exposure to the topic indeed matters.

Across all respondents (thus not analysing by treatment groups), we also see a significant decline in simplification preferences between Q2 and Q9. A Kolmogorov–Smirnov test for differences in distributions clearly rejects the null that the distributions of answers to Q2 and Q9 are the same (p -value 0.000; distributions plotted in Online Appendix Figure E.2).

Other combinations of treatment group status are also interesting, and are displayed in Table 3. For example, it makes intuitive sense that the difference between Q2 and Q9 is largest among those respondents who saw no treatment (control group) or the *Avoidance* treatment in favour of simplification in the first experiment, and were then presented with an argument against tax simplicity in the *Efficiency* group of the second experiment (differences are -0.306 and -0.264). Similarly, it is also sensible that the difference between Q2 and Q9 is smallest (-0.081) for those respondents who were provided an argument against tax simplification before replying to Q2 (i.e. those in the *Redistribution* group of the first experiment), and were then presented with an argument in favour of tax simplicity before replying to Q9 (i.e. those in the *Special-interest* group of the second experiment).

4.4 | Reasoning about tax expenditures

To improve our understanding of how people reason about tax expenditures, we field an open-ended question in our follow-up survey (see Ferrario and Stantcheva 2022).²⁷ Considering that our survey may have persistent effects, the unprimed picture particularly arises among those

subjects who were in the control groups of both experiments. The following results, displayed in Online Appendix Table F.1, therefore report the findings of this ‘pure’ control group. The survey question reads: ‘There are tax deductions and allowances in the system of the German income tax. In your opinion, what is the primary goal of the existing deduction possibilities and allowances?’ Overall, 4656 people participated in the survey, and 4130 ultimately answered the question at hand.

We use modern text analysis techniques to analyse the open-ended question. In a first step, we prepare the data, for example by removing punctuation, removing superfluous spaces, and unifying the text to lower-case letters. In a second step, we apply a spell correction, remove stop words, and use lemmatization and stemming to unify the text further. Our data cleaning procedure removes 279 observations, because they did not include useable information; for example, some of these answers were just single letters such as X, or a full point. In a third step, we manually create a training dataset consisting of 3046 observations by categorizing our dataset into several topics, namely *incentive*, *complexity*, *pro-rich*, *redistribution*, *special interest*, *tax rebate*, *tax equality*, *tax justice*, *simplification*, *special expenditures*, *don’t know* and *miscellaneous*. Online Appendix F provides further details on these categories. We use this training dataset to predict the remaining 805 observations using an SGD Classifier. Online Appendix F outlines and discusses the specifics of the empirical approach (see examples in Online Appendix Table F.2).

Our results are summarized in Online Appendix Table F.1, which displays the frequency with which different types of objectives were stated by the respondents. We focus here on the members of the control group of our main experiment, who were unencumbered by information on the effects of tax expenditures. The most frequently stated objective of tax expenditures among these subjects is the reduction of one’s own tax burden (chosen by 22%). This is followed by a considerable share of respondents who explicitly say that they do not have an idea about the objectives of tax expenditures (19.5%). At a minimum, this means that a significant minority share of the population does not understand the role of tax expenditures. These findings suggest that the understanding and reflection of complexity-adding tax expenditures and their economic purposes are not particularly sophisticated among the general population, and are generally heterogeneous. Instead, a plurality of people tend to view tax expenditures as vehicles to reduce tax payments, or they openly state that they are not aware of their purpose. Survey respondents named other categories of objectives less frequently. They seem to be somewhat aware that tax expenditures can serve to redistribute income ($\approx 10\%$) and to improve tax justice ($\approx 12.6\%$). Other objectives were stated by only a small share of people; such other classes of objectives include (frequency given in parentheses) incentivizing economic behaviour (6.8%), serving the rich or special interests (both 4.2%), preserving tax equality (5.3%), and accounting for special circumstances (9.5%). Here, 1.6% and 3.7% believe that tax expenditures explicitly exist to make the tax system complex or easy, respectively. Note that there is also a fraction of about 15.3% of overall responses that our algorithm could not categorize as one of the aforementioned categories, but still resemble valid answers.

5 | CONCLUSION

Using new experimental and survey data for a representative sample of the German population, we find that most people are in favour of a simpler tax system, but they are not well aware of the purposes of complexity-adding tax expenditures, and have heterogeneous beliefs on what they should achieve. Preferences for simplicity are mitigated once we make people aware of the trade-offs behind tax complexity and simplification. Apparently, individuals frequently express desires for tax simplification without knowing and appreciating the exact implications (externalities). A significant share of the general public also either do not know the role of tax

expenditures or view them merely as a vehicle to reduce one's own tax payments. In this vein, it is conceivable that complex tax systems can be viewed as an unintended consequence of a set of many well-intentioned and publicly demanded policies. Tax complexity may prevail since political debates about single tax expenditures do not internalize that the introduction of single (useful) tax expenditures potentially creates an externality on the overall complexity of the tax system. Our results indicate that awareness about the exact implications of simplifying reforms can change attitudes towards reforms. More awareness can moderate the strong views that people apparently have in the context of simplification, implying that the (policy, academic and public) debate about tax simplification potentially benefits from a more nuanced discussion of the pros and cons of tax reforms (and simplification in particular).

What are the implications for the political feasibility of tax reforms? Our findings regarding the role of choices and circumstances (in our main survey and the survey experiment in the follow-up survey) for example suggest that tax-simplifying reforms could focus on the abolition of self-chosen deduction possibilities (e.g. expenses for craftsmen services or professional literature in Germany), while preserving tax expenditures that compensate for exogenous living conditions that are not self-inflicted (e.g. costs for elderly care).

ACKNOWLEDGMENTS

We thank Friedrich Heinemann, James Hines, Andreas Peichl, Nadine Riedel, Sebastian Siegloch, Stefanie Stantcheva, Johannes Voget, Daniel Waldenstroem, Alfons Weichenrieder, and various seminar/conference participants for helpful comments and suggestions. We are grateful for project funding from Leibniz ScienceCampus Mannheim Taxation (project 140330). Blesse is thankful for funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project ID 139943784—SFB 884. Blesse is also grateful to KOMKIS (Competence Center for Local Infrastructure Saxony) for funding.

This paper uses data from the German Internet Panel, waves 36 and 46. A study description can be found in Blom *et al.* (2015). The German Internet Panel is funded by the DFG through the Collaborative Research Center 'Political economy of reforms' (SFB 884). Open Access funding enabled and organized by Projekt DEAL.

ENDNOTES

- ¹ Tax systems in many countries are sometimes described as *tax jungles* (*New York Times* 1964; Kenyon 1997; *Financial Times* 2006). The title of our paper is a play on the famous TV series 'I'm a celebrity ... get me out of here!', in which celebrities live together in a jungle and undertake challenges.
- ² The simplification of the tax system is a key objective of many income tax reform proposals by economists in various countries, e.g. Gale (2001), Rohaly and Gale (2004), Gravelle and Hungerford (2012), Tran-Nam (2000), and Fuest *et al.* (2008). Newspaper coverage for the USA shows that many politicians and journalists also make a case for a simplified tax system, e.g. *Economist* (2005, 2013), *New York Times* (2015), Kurtzleben (2015), Thorndike (2017), and Matthews (2017).
- ³ For example, complexity-adding tax expenditures can contribute to efficiency, such as through tax rates that are tailored to individual situations, 'tagging' components, lower rates on elastic goods, avoidance of tax compounding (e.g. favourable tax treatment of pensions and retirement savings), and the inclusion of Pigouvian elements (e.g. research tax credits) (Hines 2016; OECD 2010a). Complex components of the tax system can also have redistributive purposes.
- ⁴ Germany's tax schedule presumably includes more than 500 tax exceptions, according to Kirchhof (2011). A study by Blaufus *et al.* (2014), finds that the large number of tax expenditures, along with other particularities of the German tax system, translates into considerable income tax compliance costs of filing taxes. Using survey data, the study estimates aggregate compliance costs for Germany of 6–9 billion euros, corresponding to 3.1–4.7% of total 2007 tax revenues. There are also frequently returning debates about tax simplification in the public and media, and among politicians.
- ⁵ Finding consistent patterns across *both* experiments suggests that the results are not driven by how convincing a particular argument or particular aspect of the debate is to respondents (since it is unlikely that the arguments against simplification are more convincing in both of the experiments than the arguments in favour of simplification). Note that a substantial fraction of control group respondents did *not* select the highest possible support for simplification, suggesting that there generally would have been room for the pro-simplification argument to increase support.

- ⁶ A large literature in economics studies the costs of complex taxes. For example, tax complexity induces the self-employed to leave money on the table (Aghion *et al.* 2017; Benzarti 2020), facilitates tax evasion (Kleven *et al.* 2011; Paezold and Winner 2016; De Neve *et al.* 2021), reduces the take-up of tax refunds by firms (Zwick 2021), lends scope to lobby groups to achieve beneficial tax treatments (Brusco *et al.* 2014), affects tax revenues (Kopczuk 2012), and possibly comes with resource costs (Pitt and Slemrod 1989). Complexity of tax systems and other policy measures also distort the responses to (tax) government interventions, reduces their take-up, and induces people to misperceive tax incentives (e.g. Abeler and Jaeger 2014; de Bartolome 1995; Liebman and Zeckhauser 2004; Blaufus and Ortlieb 2009; Blumkin *et al.* 2012; Blaufus *et al.* 2013; Ito 2014; Feldman *et al.* 2016; Rees-Jones and Taubinsky 2016; Gideon 2017; Ballard and Gupta 2018; Kopczuk and Pop-Eleches 2007; Chetty and Saez 2013; Bhargava and Manoli 2015; Kotakorpi and Laamanen 2016; Fochmann *et al.* 2018).
- ⁷ A few papers explicitly study tax complexity in a political economy setup, and investigate how tax complexity arises in the interaction between voters and politicians (Warskett *et al.* 1998; Galli and Profeta 2009).
- ⁸ A general survey description is in Blom *et al.* (2015), and online at <https://www.uni-mannheim.de/en/gip> (accessed 2 October 2025).
- ⁹ For our analysis, we can complement the results from our questions with the answers from other questions in the same wave and other waves of the GIP (e.g. background characteristics of the participants are available, although they are not surveyed in our question block). Our questions were embedded in a regular wave of the GIP, and were surrounded by other GIP questions.
- ¹⁰ It is, of course, possible to believe that the system should account for the described differences in living situations and at the same time think that the overall tax system should generally be simpler. However, if the general, context-free support for simplification is considerably higher than the share of respondents who think that the system should not account for differences across people, then this could point in the direction that some people are not aware of certain aspects of tax complexity. In addition, our survey allows us to study if people express a preference to abolish *all* tax expenditures in a general context, while indicating in the context of our concrete applications that the specific expenditure should exist. Such survey behaviour would also point towards awareness deficits regarding the trade-offs involved in tax simplification. Note that we do not wish to argue that it is logically inconsistent to express a preference for a simpler tax system while wanting some tax expenditures to exist. Our sole objective is to study if people consider the involved trade-offs when they express a strong support for simplification.
- ¹¹ These are the translated survey questions. The original German questions are available on the GIP web-repository at https://search.gesis.org/research_data/ZA6956 (accessed 22 October 2025).
- ¹² Our definition of tax complexity is consistent with that of Slemrod and Kopczuk (2002), and Kopczuk (2005), who characterize an income tax system as complex when it features many deductions; see also the discussion about tax-complexity measurement in Abeler and Jaeger (2014). Moving to a system without any tax expenditures would clearly make any existing tax system simpler, easier and more comprehensible. In addition, many of the other dimensions of tax complexity (such as documentation requirements, administration and filing costs, readability, perception and salience of taxes) also apply to tax expenditures and are particularly relevant in the context of tax expenditures. Furthermore, our follow-up survey (see Section 2) reveals that Germans believe that tax expenditures and the direct consequences of tax expenditures are the prime source of tax complexity in the German tax system (see Online Appendix Figure B.5). In contrast, only about a quarter of the respondents think that complexity is caused by the tax rate.
- ¹³ The factually anonymous income tax statistic (FAST) is a 10% stratified random sample of the German income tax statistics, including information about granted deductions and exemptions. The data are available as cross-section scientific use files, and for the tax-expenditure calculations here, we use the most recent available year of 2010. See Bönke *et al.* (2017), for more information.
- ¹⁴ An alternative to having two separate experiments would have been one single experiment with more treatment groups. We chose to implement two experiments for two reasons. (i) The setup with two experiments and two elicitations of simplification preferences (Q2 and Q9) allows us to study how the support for simplification evolves over the course of the survey. (ii) In light of the number of participants and the rather subtle experimental interventions, a setup with two experiments (each with three groups) is advantageous with respect to statistical power.
- ¹⁵ Note that we do not distinguish between legal avoidance and illegal evasion activities here since both may pose a possible argument (albeit for possibly different reasons) in favour of fewer deduction possibilities and exemptions, and thus more tax simplification in the context of our survey. We aimed to convey that many deduction possibilities and exemptions increase the scope for reducing taxable income, be it through legal or illegal means, both of which are possible theoretically, and we intended to be agnostic about the source of why or how people take advantage of many deductions and exemptions.
- ¹⁶ With 60 estimated coefficients, one would expect six coefficients with a significance level of 10% even in the absence of any systematic differences between groups.
- ¹⁷ Some 53% (= $47.7/(100 - 9.6)$) checked reply category 6, meaning that the system ‘absolutely’ needs to be simplified. In the control group, 54% of the respondents checked reply category 6. Another 23% of the respondents chose the second highest reply category 5 (control group 25%). Some 16% (control group 13.4%) are in category 4, which also implies a certain preference for tax simplification.

- ¹⁸ This ‘standard deduction’ represents the lump sum deduction amount for taxpayers who do not exceed the thresholds in other deduction categories.
- ¹⁹ Even among those who do not use care deductions themselves, a majority of almost 60% (control group almost 60%) believe that person *B*, who does not have care costs, should pay more taxes. Among those who use the deduction themselves, a little more than 60% (control group 60%) think that person *B* should pay more. We acknowledge that it might be possible that a few taxpayers who do not currently use this tax expenditure expect to use it in the future. A support of this tax expenditure might then be driven by self-interest, even if they do not currently make use of the tax expenditure.
- ²⁰ Commuting might be perceived as a circumstance in certain specific situations, e.g. when a worker is relocated to another branch of her firm. However, the decision about where to live (close to work or not) will usually be a choice.
- ²¹ The coefficients that we report for the second experiment are identical to coefficients that are estimated in regressions in which the sample is restricted to respondents who were in the control group in the first experiment. We use the full interaction model, and not the sample-split variant, because this approach improves precision, and the resulting coefficients are based on the same sample that is used for the regressions for the first experiment.
- ²² Negative coefficients of the *Avoidance* treatment are consistent with the treatment having no effect. If two independent samples are drawn from the same population, then it is very likely that one sample is smaller than the other one.
- ²³ In addition, the control group support for tax simplification was not clear *ex ante* when we designed the experiment with the goal of studying the symmetry of pro- and con-arguments on simplification support. Further note that the constant decreases as we subsequently add control variables across the regression specifications, suggesting that our control variables can explain a considerable part of the high baseline support for simplification. We further see that while the *Redistribution* treatment does not have any effects, the argument in the *Avoidance* group somewhat affects respondents’ views about the distributional implications of tax complexity (Q4)—see Online Appendix Table E.1. Studying the effects on further survey questions (Q3, Q5, Q6, Q7, Q8), we do not detect any treatment effects (see Online Appendix Table C.2).
- ²⁴ For reasons of brevity, and given these results, we do not report the regression results. We acknowledge that it is possible that the interaction models for detecting heterogeneity lack statistical power, rather than providing evidence of homogeneous treatment effects. The finding that treatment effects seem to be rather constant across observable characteristics is consistent with the finding that the classical standard errors and the adjusted standard errors using the Young (2018). procedure are very similar (see above, and in particular Online Appendix D).
- ²⁵ On *ex ante* grounds, we were agnostic about interaction effects between our intervention groups of our two independent experiments. Our focus was on the individual effects of each experiment, but due to space constraints in the GIP, we placed two subsequent but independent experiments into one survey wave. Note that there may be meaningful interaction effects, but we are, for the most part, not able to identify them. Muralidharan *et al.* (2025). suggest that this may be, as is often the case for factorial designs with multiple treatments, due to low power. Additionally, we cannot exclude that not only control members but to some degree also treatment group members from the first experiment self-reflected on the tax simplification question by going through the survey. Since our within-variation shows that merely reading through the survey reduced tax simplification preferences for control-group-only members across both experiments, this kind of self-reflection may at least have reduced the magnitude of positive treatment effects, or reduced negative effects from the first experiment even further.
- ²⁶ Another possible interpretation of this finding is that respondents may have interpreted the repeated question as a signal that their initial answer was ‘incorrect’, which then results in an adjustment of their reply in the second question. While we cannot rule out this interpretation with certainty, we believe that our preferred interpretation—making subjects think about the topic reduces simplification preferences—remains more likely, for the following reasons. First, as stated before, our survey does not present any objectively correct or incorrect survey questions that could induce people to think that their initial answer was ‘incorrect’, nor do we believe that the design induces experimenter demand effects that might compel respondents to modify their answers towards what they perceive to be socially desirable responses. Second, we framed the second simplification question as a reconsideration of the entire topic to avoid giving the impression that we were repeating the question because the initial answer was ‘wrong’. Specifically, before posing the actual simplification question, we introduced it with the justification ‘Now that we have dealt extensively with various aspects of the German tax system in this survey, we would like to ask again: ...’. Third, it is not immediately clear why respondents would consistently reduce their preference for simplification if the second question led them to believe that their initial answer was incorrect. If respondents were indeed adjusting their initial response because the second question induced them to think their initial answer was incorrect, then we would expect such corrections to occur in both directions.
- ²⁷ This was implemented in GIP wave 46, March 2020 (about 20 months after the initial main experiment), which is the same survey wave that includes the additional survey experiment on the role of choice versus circumstances (see Subsubsection 4.1.1).

REFERENCES

Abeler, J. and Jaeger, S. (2014). Complex tax incentives. *American Economic Journal: Economic Policy*, 7(3), 1–28.

- Aghion, P., Akcigit, U., Lequien, M. and Stantcheva, S. (2017). Tax simplicity and heterogeneous learning. NBER Working Paper no. 24049.
- Alesina, A. and Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, **95**(4), 960–80.
- , Stantcheva, S. and Teso, E. (2018). Intergenerational mobility and support for redistribution. *American Economic Review*, **108**(2), 521–54.
- Ballard, C. L. and Gupta, S. (2018). Perceptions and realities of average tax rates in the federal income tax: evidence from Michigan. *National Tax Journal*, **71**(2), 263–94.
- Benzarti, Y. (2020). How taxing is tax filing? Using revealed preferences to estimate compliance costs. *American Economic Journal: Economic Policy*, **12**(4), 38–57.
- and Wallossek, L. (2024). Rising income tax complexity. *National Tax Journal*, **77**(1), 143–73.
- Bhargava, S. and Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: evidence from an IRS field experiment. *American Economic Review*, **105**(11), 3489–529.
- Bierbrauer, F., Boyer, P. and Peichl, A. (2021). Politically feasible reforms of non-linear tax systems. *American Economic Review*, **111**(1), 153–91.
- Blaufus, K., Bob, J., Hundsdoerfer, J., Kiesewetter, D. and Weimann, J. (2013). Decision heuristics and tax perception—an analysis of a tax-cut-cum-base-broadening policy. *Journal of Economic Psychology*, **35**, 1–16.
- , Eichfelder, S. and Hundsdoerfer, J. (2014). Income tax compliance costs of working individuals: empirical evidence from Germany. *Public Finance Review*, **42**(6), 800–29.
- and Ortlieb, R. (2009). Is simple better? A conjoint analysis of the effects of tax complexity on employee preferences concerning company pension plans. *Schmalenbach Business Review*, **61**(1), 60–83.
- Blom, A. G., Gathmann, C. and Krieger, U. (2015). Setting up an online panel representative of the general population: the German Internet Panel. *Field Methods*, **27**(4), 391–408.
- , Herzog, J. M., Cornesse, C., Sakshaug, J. W., Krieger, U. and Bossert, D. (2017). Does the recruitment of offline households increase the sample representativeness of probability-based online panels? evidence from the German internet panel. *Social Science Computer Review*, **35**(4), 498–520.
- Blumkin, T., Ruffle, B. J. and Ganun, Y. (2012). Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods. *European Economic Review*, **56**(6), 1200–19.
- Bönke, T., Schröder, C. and Jochimsen, B. (2017). Fiscal equalization and tax enforcement. *German Economic Review*, **18**(3), 337–409.
- Brennan, G. and Buchanan, J. M. (1980). *The Power to Tax*. Cambridge: Cambridge University Press.
- Brusco, S., Colombo, L. and Galmarini, U. (2014). Tax differentiation, lobbying, and welfare. *Social Choice and Welfare*, **42**(4), 977–1006.
- Chetty, R. and Saez, E. (2013). Teaching the tax code: earnings responses to an experiment with EITC recipients. *American Economic Journal: Applied Economics*, **5**(1), 1–31.
- de Bartolome, C. (1995). Which tax rate do people use: average or marginal? *Journal of Public Economics*, **56**(1), 79–96.
- De Neve, J.-E., Imbert, C., Spinnewijn, J., Tsankova, T. and Luts, M. (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. *Journal of Political Economy*, **129**(5), 1425–63.
- de Quidt, J., Haushofer, J. and Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, **108**(11), 3266–302.
- Durante, R., Putterman, L. and van der Weele, J. (2014). Preferences for redistribution and perception of fairness: an experimental study. *Journal of the European Economic Association*, **12**(4), 1059–86.
- Economist* (2005). Simplifying tax systems: the case for flat taxes. *The Economist*, 14 April; available online at <https://www.economist.com/special-report/2005/04/14/the-case-for-flat-taxes> (accessed 3 October 2025).
- (2013). Tax reform in America: simpler, fairer, possible. *The Economist*, 13 July; available online at <https://www.economist.com/leaders/2013/07/13/simpler-fairer-possible> (accessed 3 October 2025).
- Feldman, N. E., Katuscák, P. and Kawano, L. (2016). Taxpayer confusion: evidence from the child tax credit. *American Economic Review*, **106**(3), 807–35.
- Ferrario, B. and Stantcheva, S. (2022). Eliciting people’s first-order concerns: text analysis of open-ended survey questions. *American Economic Association Papers and Proceedings*, **112**, 163–9.
- Financial Times* (2006). We must ‘keep it simple’ to clear the tax jungle. Opinion piece by Geoffrey Howe, 21 March; available online at <https://www.ft.com/content/b460b7e8-b90f-11da-b57d-0000779e2340> (accessed 4 October 2025).
- Fisher, R. A. (1935). *The Design of Experiments*. Edinburgh: Oliver and Boyd.
- Fisman, R., Gladstone, K., Kuziemko, I. and Naidu, S. (2020). Do Americans want to tax wealth? Evidence from online surveys. *Journal of Public Economics*, **188**, 104207.
- Fochmann, M., Müller, N. and Overesch, M. (2018). Less cheating? The effects of prefilled forms on compliance behavior. *Journal of Economic Psychology*, **83**, 102365.
- Fuest, C., Peichl, A. and Schaefer, T. (2008). Does tax simplification yield more equity and efficiency? An empirical analysis for Germany. *CESifo Economic Studies*, **54**(1), 73–97.

- Gale, W. (2001). Tax simplification: issues and option. Mimeo.
- Galli, E. and Profeta, P. (2009). Tax complexity with heterogeneous voters. *Public Finance and Management*, **9**(2), 1–35.
- Gideon, M. (2017). Do individuals perceive income tax rates correctly? *Public Finance Review*, **45**(1), 97–117.
- Gravelle, J. G. and Hungerford, T. L. (2012). *The challenge of individual income tax reform: an economic analysis of tax base broadening*. Congressional Research Service, Washington, DC.
- Haaland, I., Roth, C. and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, **61**(1), 3–40.
- Hettich, W. and Winer, S. L. (1988). Economic and political foundations of tax structure. *American Economic Review*, **78**(4), 701–12.
- Hines, J. R. (2016). High tax heresy. Lecture at Center on Finance, Law & Policy at University of Michigan; available online at <https://fordschool.umich.edu/event/2021/high-tax-heresy> (accessed 22 October 2025).
- Ito, K. (2014). Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *American Economic Review*, **104**(2), 537–63.
- Kenyon, D. A. (1997). *Tax Policy in an Intergovernmental Setting: Is it Time for the US to Change?* Dordrecht: Springer, pp. 61–97.
- Kirchhof, P. (2011). *Bundessteuergesetzbuch. Ein Reformentwurf zur Erneuerung des Steuerrechts*. Heidelberg: C. F. Müller Wissenschaft.
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S. and Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, **79**(3), 651–92.
- Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, **89**(11–12), 2093–119.
- (2012). The Polish business ‘flat’ tax and its effect on reported incomes: a Pareto improving tax reform? Mimeo; available online at <http://www.columbia.edu/~wk2110/bin/PolishFlatTax.pdf> (accessed 4 October 2025).
- and Pop-Eleches, C. (2007). Electronic filing, tax preparers and participation in the earned income tax credit. *Journal of Public Economics*, **91**(7–8), 1351–67.
- Kotakorpi, K. and Laamanen, J.-P. (2016). Prefilled income tax returns and tax compliance: evidence from a natural experiment. University of Tampere, School of Management Working Paper no. 104.
- Kurtzleben, D. (2015). Lots of candidates want to simplify tax code; here’s what they get wrong. National Public Radio, 28 September; available online at <https://www.npr.org/sections/itsallpolitics/2015/09/28/443203910/lots-of-candidates-want-to-simplify-the-tax-code-heres-what-they-get-wrong?t=1551777844261> (accessed 3 October 2025).
- Liebman, J. B. and Zeckhauser, R. J. (2004). Schmeduling. Mimeo, Harvard Kennedy School.
- Matthews, D. (2017). Why Democrats should support radically simpler taxes. Vox, 8 May; available online at <https://www.vox.com/policy-and-politics/2017/5/8/15442172/democrats-tax-plan-return-free-filing-trump-ambitious> (accessed 4 October 2025).
- Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, **89**(5), 914–27.
- Mummolo, J. and Peterson, E. (2019). Demand effects in survey experiments: an empirical assessment. *American Political Science Review*, **113**(2), 517–29.
- Muralidharan, K., Romero, M. and Wüthrich, K. (2025). Factorial designs, model selection, and (incorrect) inference in randomized experiments. *Review of Economics and Statistics*, **107**(3), 589–604.
- New York Times* (1964). The tax jungle. *New York Times*, 21 December; available online at <https://www.nytimes.com/1964/12/21/archives/the-tax-jungle.html> (accessed 3 October 2025).
- (2015). The tax code can be simpler. But not three pages. *New York Times*; available online at <https://www.nytimes.com/2015/11/15/upshot/a-three-page-tax-code-not-exactly-simple.html> (accessed 17 May 2025).
- OECD (2010a). Tax design considerations. In *Tax Policy Reform and Economic Growth*. Paris: OECD Publishing; available online at <https://doi.org/10.1787/9789264091085-7-en> (accessed 4 October 2025).
- (2010b). *Tax Expenditures in OECD Countries*. Paris: OECD Publishing; available online at <https://doi.org/10.1787/9789264076907-en> (accessed 4 October 2025).
- Ooghe, E. and Peichl, A. (2015). Fair and efficient taxation under partial control. *Economic Journal*, **125**(589), 2024–51.
- Paetzold, J. and Winner, H. (2016). Taking the high road? Compliance with commuter tax allowances and the role of evasion spillovers. *Journal of Public Economics*, **143**, 1–14.
- Paul, D. L. (1997). The sources of tax complexity: how much simplicity can fundamental tax reform achieve. *North Carolina Law Review*, **76**(1), 151.
- Pitt, M. and Slemrod, J. (1989). The compliance cost of itemizing deductions: evidence from individual tax returns. *American Economic Review*, **79**(5), 1224–32.

- Ramthun, C. (2017). Pendler kosten den Fiskus fünf Milliarden Euro. *WirtschaftsWoche*, 3 December; available online at <https://www.wiwo.de/politik/deutschland/berufspendler-pendler-kosten-den-fiskus-fuenf-milliarden-euro/20654060.html> (accessed 4 October 2025).
- Rees-Jones, A. and Taubinsky, D. (2016). Measuring ‘schmeduling’. NBER Working Paper no. 22884.
- Rohaly, J. and Gale, W. G. (2004). Effects of tax simplification options on equity, efficiency, and simplicity: a quantitative analysis. In H. J. Aaron and J. Slemrod (eds), *The Crisis in Tax Administration*. Washington, DC: Brookings Institution Press.
- Saez, E. and Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, **106**(1), 24–45.
- Slemrod, J. and Kopczuk, W. (2002). The optimal elasticity of taxable income. *Journal of Public Economics*, **84**(1), 91–112.
- Thorndike, J. (2017). Trump wants a simple tax system—history says he won’t get it. *Forbes*, 11 September; available online at <https://www.forbes.com/sites/taxanalysts/2017/09/11/trump-wants-a-simple-tax-system-history-says-he-wont-get-it/#4c3b63e27c1c> (accessed 3 October 2025).
- Tran-Nam, B. (2000). Tax reform and tax simplicity: a new and ‘simpler’ tax system? *University of New South Wales Law Journal*, **23**(2), 241–51.
- Warskett, G., Winer, S. L. and Hettich, W. (1998). The complexity of tax structure in competitive political systems. *International Tax and Public Finance*, **5**, 123–51.
- Weinzierl, M. (2014). The promise of positive optimal taxation: normative diversity and a role for equal sacrifice. *Journal of Public Economics*, **118**, 128–42.
- (2017). Popular acceptance of inequality due to innate brute luck and support for classical benefit-based taxation. *Journal of Public Economics*, **155**, 54–63.
- Westfall, P. H. and Young, S. S. (1993). *Resampling-based Multiple Testing: Examples and Methods for p-value Adjustment*. New York: John Wiley & Sons.
- Young, A. (2018). Channeling Fisher: randomization tests and the statistical insignificance of seemingly significant experimental results. *Quarterly Journal of Economics*, **134**(2), 557–98.
- Zwick, E. (2021). The costs of corporate tax complexity. *American Economic Journal: Economic Policy*, **13**(2), 467–500.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Blesse, S., Buhlmann, F. and Doerrenberg, P. (2026). I’m a taxpayer ... get me out of here? Evidence on attitudes towards simplifying the tax jungle. *Economica*, **93**(369), 332–360. <https://doi.org/10.1111/ecca.70020>