I am a Taxpayer ... Get Me Out of Here?!? Evidence on Attitudes Towards Simplifying the Tax Jungle *

Sebastian Blesse
Florian Buhlmann
Philipp Doerrenberg

February 28, 2023

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 then 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 individual and publicly demanded tax expenditures.

JEL Classification: H2, D72, C9
Keywords: Tax Complexity, Attitudes Towards Tax Simplification, Randomized Survey Experiment

*__Blesse__: Ludwig Erhard ifo Center for Social Market Economy and Institutional Economics, ZEW Mannheim and CESifo. Email: blesse@ifo.de. **Buhlmann**: ZEW Mannheim and University of Mannheim. Email: buhlmann@zew.de. **Doerrenberg**: University of Mannheim, CESifo, IZA and ZEW. Email: doerrenberg@uni-mannheim.de. 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 of Leibniz ScienceCampus MannheimTaxation (project 140330). Blesse is thankful for funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project ID 139943784—SFB 884. This paper uses data from the German Internet Panel, wave 36 and wave 46. A study description can be found in Blom et al. (2015). The German Internet Panel (GIP) is funded by the German Research Foundation (DFG) through the Collaborative Research Center 884 "Political Economy of Reforms" (SFB 884).

1 Introduction

Should tax systems – sometimes described as tax jungles (New York Times 1964; Kenyon 1997; Financial Times 2006) – be simplified? The conventional wisdom seems to be: *yes*, tax systems should be simpler! As the literature shows, there are indeed many good reasons for supporting tax simplification.[1] It is therefore maybe not surprising that many economists propose implementing tax reforms that make the system less complex, and the conventional wisdom among politicians and journalists also holds that simplifying tax systems is generally desirable.[2] However, there also are economic arguments in support of a certain degree of tax complexity (e.g., OECD 2010b, Hines 2016, 2019, also see Section 2 below), but these arguments feature less prominently in the debate and the discussion therefore perhaps misses out some of the arguments in support of some complexity. On the other hand, the data show that most tax systems remain very complex and tend to get even more complex over time (e.g., OECD 2010b and Figure 1 for the US). It is thus puzzling why most tax systems remain so complex, although the general wisdom seemingly holds that substantial tax simplification is desirable. Against this background, it may be 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. Aside from documenting and characterizing public attitudes towards tax complexity, we particularly study the role of awarness with respect to the arguments against and in favor of tax complexity. Taken together, our results then also shed light on the questions of which tax system the general public favors and of why tax systems remain so complex (Section 2.3 provides an overview of our objectives and related hypotheses).

To address our research objectives, we collect new survey and experimental data among a representative sample of the German population. Our large set of tailor-made survey questions is embedded in the *German Internet Panel* (GIP), a probability-based survey of the German population ($N = 2432$; see Section 3 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 for the purpose of our study: tax expenditures. While there are clearly more dimensions of tax complexity (see below), tax expenditures are a main source of tax complexity and a major issue in the debate about complexity. We first document patterns of tax complexity attitudes (Section 4): More than 2/3 of respondents have a strong general preference for tax simplification. When tax simplicity is considered without any further context, the prevailing view generally is in support of tax simplification. In addition, a question about the preferred type of complexity-reducing tax reform reveals that there is large

[1] See Section 2 below for references. 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. For example, Gale (2001), Rohaly and Gale (2004) and Gravelle and Hungerford (2012) for the US, James et al. (1997) for Australia, New Zealand and the United Kingdom, Tran-Nam (2000) for Australia, and Fuest et al. (2008) and Wagner (2006) for Germany. Newspaper coverage for the US shows that many politicians and journalists also make a case for a simplified tax system: e.g., Economist (2005), Economist (2013), NYT (2015), NPR (2015), Forbes (2017), as well as Vox (2017).

heterogeneity among respondents about strategies to improve the tax system.

We then move on from the finding that tax simplification is supported by a considerable majority and use four distinct strategies to study if the matter becomes more nuanced as we highlight the implications of simplification. First (Section 5.1), we elicit if people are in favor of specific complexity-adding tax expenditures. To this end, we have a series of questions in which we present participants with the living situations of two fictitious taxpayers, and then survey if any of the two should pay less taxes than the other person. We designed the questions in a way where the two fictitious persons are similar in all tax-relevant means except for one particular aspect of their living situation. In particular, the two fictitious taxpayers are different w.r.t. i) the necessity to spend money on the elderly care of a family member, ii) the amount they donate for charity, and iii) the distance between their home and work place. 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.[3] The share of people who indicate in *all* of the 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 to have a simpler tax system. Interestingly, we further find that 70% of those respondents who, in another earlier survey question, say that they wish to abolish *all* allowances and deductions indicated at least once that the two fictitious taxpayers should *not* pay the same amount of taxes.

Obviously, a differential treatment of two otherwise identical taxpayers can only be achieved through tax expenditures. The results therefore suggest that many people have a preference for a tax system that is more complex than a system without any tax expenditures. Our findings are indicative of the notion that people are not fully aware of the implications and involved trade-offs behind simplification.[4] This corresponds with the results of a follow-up study with open ended questions where we show that people are not very much aware of the intended purposes of tax expenditures (see below for more).

The results in our fictitious-taxpayer questions 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 (Section 5.1.1). This exercise confirms that sur-

[3]We show that the answers to these questions are not solely driven by self interest; the result picture remains as we condition on *not* benefiting personally from the respective tax expenditure. Our findings thus show that preferences for having certain tax expenditures in the tax system are not (only) driven by the desire to keep those tax expenditures from which someone benefits personally.

[4]Note that we do not mean to argue that a preference for tax simplification is in logical conflict with wanting some tax expenditures to exist (see the Conclusion for more discussion of this). Also note that we of course acknowledge that governments also have non-tax tools to account for different living situations (e.g., in-kind transfers). However, taxes are frequently used to account for alternative living situations and taxes are the focus of our survey questions here.

vey respondents are more willing to add complexity-adding expenditures if these expenditures compensate for circumstances, rather than choices. In an open-end question in our follow-up survey study, we find that around 10% of respondents also explicitly mention tax justice and compensation for circumstantial differences as a main objective of existing deductions. These results speak to which type of tax system people actually prefer.

Second (Section 5.2), we implement two randomized experiments to study if preferences for tax simplicity are elastic to arguments in favor and against tax simplification, one 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, ii) one treatment group in which preferences are possibly shifted towards stronger support for tax simplification, iii) one treatment group in which preferences are possibly shifted towards less support for tax simplification. This set up thus allows us to study if preferences for tax simplicity are elastic at all, and if yes, if they are elastic in both directions. 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 (see Section 2 below): i) redistribution and social-policy aspects, ii) tax avoidance and evasion, iii) efficiency and iv) lobbyism 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 favor* of simplification do not have an effect. In particular, the arguments against simplification significantly reduce the support for simplification, whereas we see no effects in response to the arguments in favor of simplification. Overall, the results in both experiments suggest that people are less aware of arguments against simplification. The null-results in response to arguments in favor of simplification testify to more existing awareness and less misperceptions when it comes to arguments that support more simplification.[5] These findings are consistent with our working hypotheses (see section 2.3) and with the observation that arguments against simplification feature less prominently in the public debate about tax simplification.

Third (Section 5.3), 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 elicitation of simplification preferences (despite the known tendency that survey respondents aim to give consistent replies). We find this pattern across all participants and, more importantly, among respondents who are in the control group in both of our survey experiments. Since these respondents are never exposed to any arguments (neither against nor in favor) relating to tax complexity, the falling support among these respondents can be attributed to the mere fact that they replied to questions in the context of tax simplicity and were thereby exposed to the topic. This finding is consistent with the other findings in our paper that awareness reduces

[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 favor of simplification). Statistical inference in both experiments is robust to multiple hypothesis testing and exact significance tests. 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 some room for the pro-simplification argument to increase support.

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 that people are exposed to in the experimental intervention.

Fourth (Section 5.4), we use open-end questions (in the absence of any priming; Ferrario and Stantcheva 2022) in a follow-up survey to examine to which extent people (who were never exposed to any information treatments) are aware of the objectives of existing tax expenditures (we already alluded to some of the results above). 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. Such findings support the view that people are not fully aware of the arguments against and in favor of tax complexity. The follow-up survey also provides the opportunity to investigate the persistency of the survey experiments' treatment effects 20 months later. We find that the arguments in favor of complexity have a persistent effect on the probability of answering the open-ended text question, while the arguments against complexity do not trigger any effects 20 months later. Resembling the main experimental results, this suggests that the arguments in favor of complexity stick for longer and were more surprising to participants.

Contribution and Key Take-Aways. We identify the following main contributions of our paper (see section 2.1 for an overview of the related literature). First, we study if the support for tax simplification depends on the extent of awareness about the consequences and implications of tax simplification. We show that, while most people express a general preference for more tax simplification, they persistently realize that simplification comes at a cost when they are faced with the implications of simplification. 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. Our results speak to the puzzle that real-world tax systems remain complex despite of a consensus that tax systems should generally be simpler. People express a large general support for simplification, but at the same time demand specific expenditures. In addition, there is large heterogeneity among the population with respect to the desired expenditures in the tax system. This suggests that, while there is no explicit policy advocacy in favor of tax complexity, complex tax systems can be viewed as an unintended consequence of a set of many publicly demanded single tax expenditures (without appreciating the respective implications on the overall tax system).[6]

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 (also see Conclusion for more). In particular, and this is our second contribution, we show that people prefer maintaining complexity-adding expenditures that compensate for circumstances rather than choices. This relates to the literature on optimal taxation and equality of opportunity. Third, we implement the first nuanced survey in the context of tax complexity and integrate our questions into several waves of an established representative probability-based survey. The survey design itself therefore stands as a contribution. To this end, we add to a recent literature using tailor-made surveys to study the political

[6]See the discussion in the concluding section 6 for more on this.

economy and feasibility of reforms in different domains of economic policy.

2 Related Literature, Tax Complexity in Germany, and Hypotheses

In this section, we first summarize the related literature (section 2.1) and then present an overview of the (institutional) context in Germany (2.2). Building on the literature and the German context, we then summarize the main objectives of our paper and derive working hypotheses (2.3).

2.1 Related Literature

We relate to the following strands of literature. First, we speak to the literature studying tax complexity. The starting point of our paper is literature showing that tax complexity comes with costs. 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; Paetzold and Winner 2016; Tsankova et al. 2019), reduces the take up of tax refunds by firms (Zwick 2020), 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). Other papers demonstrate that the complexity of tax systems and other policy-measures distorts 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 et al. 2018; Kopczuk and Pop-Eleches 2007; Chetty and Saez 2013; Bhargava and Manoli 2015; Kotakorpi and Laamanen 2016; Fochmann et al. 2018).[7]

A further motivation for our study are those few studies which highlight the potential upsides of tax complexity and express a rationale for the implementation of tax expenditures. It was noted that a fairly complex tax system with a reasonable amount of tax expenditures can contribute to efficiency, for example through tax rates that are tailored to individual situations, 'tagging' components, lower rates on elastic goods, avoidance of tax compounding (e.g., favorable tax treatment of pensions and retirement savings), and the inclusion of Pigouivian elements that correct for market failures or internalize negative externalities, e.g. research tax credits (Hines 2016; Hines 2019). Complex components of the tax system also have redistributive purposes – e.g., deducting the costs of elderly care of family members or allowances for dependent children – and might therefore be viewed as equity enhancing. James Hines discusses several potential justifications to have a comprehensive and simple system with a broad base and low rates, and elaborates that many of these justifications for simplicity do not withstand economic reasoning (Hines 2016; Hines 2019). It is therefore not clear that it is necessarily desirable to

[7] Somewhat related also is the finding by Brown et al. (2017) that complexity complicates the ability of consumers to value life annuities (such as social security benefits).

have the simplest tax system with a broad base and low rates.[8] OECD (2010a) also discuss the rationale for implementing tax expenditures.[9]

Second, we touch upon a literature on the political economy of taxes and tax reforms (e.g., Brennan and Buchanan 1980; Meltzer and Richard 1981; Bierbrauer et al. 2020). We point to an apparent puzzle that tax complexity remains high in the real-world despite the seemingly wide support for tax simplification. Our findings provide a potential answer to this: Not aware of the potential overall implications of (complexity enhancing) tax expenditures, people support the existence of various single tax expenditures. As a result, there is an unintended amount of many tax expenditures, and thus complexity, in the existing system. To this end, we for example 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. A few papers explicitly study tax complexity in a political-economy set-up and investigate how tax complexity arises in the interaction between voters and politicians (Warskett et al. 1998; Galli and Profeta 2009). Our paper also speaks to these papers as it provides novel evidence that the general public has limited awareness of the possible implications of tax expenditures. Our results also suggest that arguments against tax simplicity could play a more prominent role in the voting process if voters were more aware of the trade-offs behind tax complexity and simplicity. Showing that individuals have a biased and uninformed understanding about taxes, we also relate to literature showing that such information frictions may indeed induce the government to implement inefficient tax policy (Boccanfuso and Ferey 2019).

Third, we speak to papers in the context of optimal taxation and equality of opportunity, which show that circumstantial differences should be accounted for to a larger extent by the tax system than deliberate and self-chosen differences (e.g., Alesina and Angeletos 2005; Durante et al. 2014; Ooghe and Peichl 2015).

Fourth, we join a set of papers that set up tailor-made surveys with randomized components to study a particular research topic (see Haaland et al. 2021 and Stantcheva 2022 for extensive overviews). The vast majority of existing survey papers rely on commercial providers who conduct the surveys online and establish representativity through a reweighting of the initially non-representative sample. We implement our questions within the GIP, an established survey with a representative sample of the German population that was explicitly build up for research purposes. Using a representative, probability-based survey (over several waves, thereby allowing to study persistency) therefore is a contribution of our study relative to the majority of

[8]The role of economic theory in this discussion is addressed by Hines (2016) who concludes (in the Abstract) that: "Economic theory does not say that an efficient and equitable income tax system has a broad base and a low rate, and in fact the theory has never said that." Gordon and Kopczuk (2014) study the selection of the income tax base and show that it is advantageous (in the sense of approximating a tax on ability as good as possible) to allow for particular tax expenditures (such as the dependents' deduction). Thus, there is an implicit rationale for not having the simplest possible tax system with a broad base and without any tax expenditures.

[9]They particularly point out arguments of i) tax administration costs (costs of broadening the base might exceed the corresponding efficiency gains), ii) equity and social-policy considerations (tax provisions might have the same purposes as social benefits), iii) correcting of market failures (internalize positive external effects), and iv) a political-economy argument, that they borrow from Hettich and Winer (1999), according to which the elimination of tax expenditures possibly reduces tax revenues (abolishing tax expenditures implies that government will be less able to discriminate among heterogeneous taxpayers and voters, which will lead to an increased overall opposition to taxation).

other papers in the context of randomized survey experiments. In contrast to many papers in the recent survey literature, yet similar e.g., to Stantcheva (2021), we use qualitative arguments in the survey experiments and exploit open-ended survey questions using machine-learning based techniques of text analysis (see the 'guidelines' in Ferrario and Stantcheva 2022).

2.2 Tax Complexity in Germany

Our survey is conducted in a country with a fairly complex income-tax system. For example, Germany's tax schedule presumably includes more than 500 deduction possibilities, according to Kirchhof (2011). Germany does not use pre-populated tax returns. As a result, all expenditures have to be explicitly filed by the taxpayer when completing the tax return. The German income-tax system is probably quite representative for the income-tax systems in many industrialized countries. A study by Blaufus et al. (2014) finds that the large number of tax expenditures, along with other particularities of the German tax system, translate into considerable income-tax compliance costs of filing taxes. Using survey data, the study estimates aggregate compliance costs for Germany of 6-9 billion EUR, corresponding to 3.1-4.7% of total 2007 tax revenues. Studying the topic of tax complexity in the context of Germany thus appears a sensible choice given its complex income tax system. In light of the large number of tax expenditures, studying complexity through its dimension of the number of tax expenditures is also reasonable. Indeed, as we show in a follow-up survey in the GIP (see Section 3), 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 Figure B.5). In contrast, only about a quarter of the respondents think that complexity is caused by the tax rate.

A further reason for why Germany is an interesting case to study complexity is that there are frequently returning debates about tax simplification in the public, media and among politicians. One prominent example of this debate is the proposal by prominent politicians (particularly in the conservative center-right party) to simplify the tax system in a way that makes it possible to file the income-tax return on a sheet of paper that is not larger than a usual German beer coaster (such proposals were originally made in 2003 and kept coming back ever since; see FAZ 2004 or Goettinger Tageblatt 2018). Another salient example is the proposal of a prominent academic tax lawyer (Paul Kirchhof) during election campaigns to introduce an income-tax system with a flat rate of 25% and considerably less tax expenditures (see e.g., FAZ 2005).

2.3 Main Objectives and Hypotheses

In this section, we summarize the two main objectives of our paper and, based on the related literature and the public discussion about tax complexity (as described above), derive hypotheses regarding the objectives.

First Objective. Our first objective is to document preferences for tax simplification among a representative sample of the population and to understand 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 media) are centered 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 considerably less prominent role. Asking people about their desire to simplify taxes without providing any context then presumably triggers an obvious reply, namely that the tax system needs simplification.

However, 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 will eventually be among a representative sample of the population. There are good arguments for keeping certain complexity adding components in the tax system and we empirically observe that real-world tax systems remain complex (see Figure 1). These might be indications that the support is not as enormous as one would believe on first glance. We therefore view the documentation of simplification support as one contribution of our paper.

In the context of our first objective, we additionally aim to understand and characterize attitudes towards simplification in even more depth. For example, we study the anatomy of simplification preferences in order to understand the heterogeneity in tax-simplification preferences and we survey the perceived complexity of the system and the perceived distributional implications of tax simplification. We also elicit which type of simplifying tax reforms our survey respondents prefer.

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

We employ four strategies to study the role of awareness. In the context of our first strategy, we force people to reflect on concrete tax-relevant situations that are potentially familiar to them and then let them decide how the tax system, in their view, should address these situations. Because the debate about complexity seems to lean towards simplification, we deliberately confront individuals with scenarios which 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. Our expectation is that the share of respondents who indicate that the tax system should *not* account for certain differences in living situation (and thus be as simple as possible) is considerably smaller than the share of respondents who generally support to have a simpler tax system.

It is of course possible to believe that the system should account for the described differ-

ences in living situations and at the same time think that the overall tax system should generally be simpler. However, if the general 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 behavior would also point towards awareness deficits regarding the trade-offs involved in tax simplification. Please 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 objective solely is to study if people consider the involved trade-offs when they express a strong support for simplification.

The three scenarios that we present to respondents differ with respect to their degree of being circumstantial (exogenous) to taxpayers or the result of a choice. Building on the literature on optimal taxation and equality of opportunity (e.g., Alesina and Angeletos 2005; Durante et al. 2014; Ooghe and Peichl 2015), we further hypothesize that the share of people who believe that circumstantial living situations should be accounted for in the tax system is higher than the share of people who believe that chosen living situations should be accounted for. Such findings then provide insights on whether people would prefer the simplest possible tax system without any expenditures or if they would like a tax system that includes tax expenditures which compensate for circumstantial and/or choice-based situations. We conduct an additional survey experiment in a follow-up study to shed more light on this aspect.

The second strategy builds on randomized survey components that expose participants to arguments against and in support of tax simplification. The rationale for this approach is: if the provided arguments shift individuals' preferences, then individuals are not aware of certain aspects of tax simplification in the relevant public debate. We hypothesize that misperceptions and a lack of awareness are less prevalent when it comes to arguments *against* tax simplification. As a result, shifts in preferences are presumably larger in response to arguments against simplification than to responses in support of simplification. This hypothesis (again) builds on the observation that both the public debate and the professional discussion are centered around 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. This (within-variation) strategy allows us to study if 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. In light of 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. We particularly focus on subjects who are in the control group in both experiments, because these respondents are never exposed to any arguments in favor or against simplification, and the development of simplification preferences over the course of the survey is solely driven

by awareness due to exposure to the topic.

The fourth and last strategy in the context of our second objective 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 and to investigate the persistence of our experimental effects. This analysis enables us to study how people reflect on how complexity-adding tax expenditures work and why they exist in the first place.

Overall, our paper and the hypotheses relate to studies that emphasize the importance of the public opinion and the role of taxpayers as voters for the design of the tax system (e.g., Hettich and Winer 1988). We expect that the public opinion is dependent on context, awareness and the understanding of arguments, and that attitudes towards tax simplicity of the general public may be more nuanced when people are exposed to implications of tax simplification.

3 The Survey

3.1 German Internet Panel

Our questions are embedded in the *German Internet Panel* (henceforth: GIP). The GIP is a longitudinal survey that is operated and administered at the University of Mannheim in Germany.[10] The main purpose of the panel survey is to collect 'data on individual attitudes and preferences relevant in political and economic decision making processes'. 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 its representativity). 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 only included 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 among which replied to our key survey question of interest). Summary statistics are presented and discussed in Appendix A.

We also report the results of a later follow-up survey questions fielded in March 2020 (GIP wave 46) at several spots 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 (and initial) 2018 survey. The details of the follow-up survey questions are described as we mention them in the text (in particular in section 5.1.1),

[10]To be more precise, the survey is based at the "Collaborative Research Center 884 on Political Economy of Reforms", which is funded by the German Science Foundation (*Deutsche Forschungsgemeinschaft*, SFB 884). See `http://reforms.uni-mannheim.de/` for background information on the research center. Also see the general survey description in Blom et al. (2015) and on `https://www.uni-mannheim.de/en/gip/`. Examples of GIP-based papers include Kerschbamer and Müller (2020), Müller and Renes (2021), Dolls and Wehrhoefer (2021), Engelmann et al. (2020), Doerrenberg and Peichl (2022), Gsottbauer et al. (2022) and Blesse and Heinemann (2020).

in the notes to the Figures that summarize the follow-up results (Figures 5 and B.5), as well as in Section 5.4). We always refer to the 2020 follow-up survey as "follow-up survey" when we report its results. That is, we refer to the main survey if not mentioned explicitly otherwise.

3.2 Survey Structure and Questions

We designed a block of 10 survey questions and integrated these questions into the GIP. For our analysis, we can complement the results from our questions with the results from other questions in the same wave and other waves of the GIP (for example, 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.

Our survey questions and the respective reply categories are shown in full in Appendix section G.1.[11] The GIP has a professional and experienced team of survey experts who supported us in developing and formulating our survey questions. Our questions therefore meet up-to-date standards of survey methodology. The survey and its structure (in chronological order) are summarized in the following.

- **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. This latter statement thus explains to participants which dimension of complexity we are particularly interested in.[12] 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 below for more

[11] These are the translated survey questions. The original German questions are available on the GIP website (`https://www.uni-mannheim.de/en/gip/for-data-users/questionnaires-and-documentation/`).

[12] We focus on tax expenditures because they probably constitute the largest source of tax complexity; 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, readibility, perception and salience of taxes) also apply to tax expenditures and are particularly relevant in the context of tax expenditures. Furthermore, a follow-up survey shows that tax expenditures (and their direct consequences) are perceived to be the prime source of tax complexity. A straight forward definition of tax expenditures is provided by the Tax Policy Center (2019): 'Tax expenditures are special provisions of the tax code such as exclusions, deductions, deferrals, credits, and tax rates that benefit specific activities or groups of taxpayers.' Our definition of tax complexity is consistent with Slemrod and Kopczuk (2002) and Kopczuk (2005) who characterize an income tax system as complex when it features many deductions (also see the discussion about tax-complexity measurement in Abeler and Jaeger 2014).

info.

- **Q2: Preferences for tax simplification:** We ask participants if they think whether the income-tax system in Germany generally needs to be simplified. This question elicits preferences for tax simplification and also is a potential outcome variable in the analysis of the effects of the randomized argument provision. We use this question to document the degree of support for tax simplification and to analyze the anatomy of simplification preferences.

- **Q3: General need for tax reform:** The question surveys if participants in general think that the German tax system is in need of reform.

- **Q4: Distributional implications of tax expenditures:** The question measures participants' beliefs about the distributional effects of tax expenditures. We particularly survey whether someone beliefs that tax expenditures contribute to a fairer distribution of income or if 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?:** In each of these three questions, we present to the survey respondents the living situations of two fictitious taxpayers A and B (see Weinzierl 2014, Saez and Stantcheva 2016, Weinzierl 2017 and Fisman et al. 2020 for similar survey question techniques). Respondents are told that A and B earn the same gross income and are very similar in all other (tax relevant) means, but only differ in one particular dimension. We have three different scenarios of varying living situations, and for each scenario we ask participants if A and B should pay the same amount of taxes or if any of the two should pay more. In the three presented scenarios, A and B differ with respect to i) the amount that has to be paid for elderly care of a poor mother, ii) the amount that is given to charity, and iii) the distance they commute to work. The three scenarios are presented in random order to avoid any order effects.

  Motivation: The choice of these three type of tax expenditures 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 Euro (WiWo 2017). Moreover, these items are used by many taxpayers; about 35% of the taxpayers use the commuting expenditure, about 24 percent of the taxpayers deduct charitable contributions from their tax base, and approximately 8 percent of the taxpayers use the elderly care deductions (own calculations based on German administrative tax records

(FAST); see FAST 2010).[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).

- **Randomized Experiment 2:** Participants are again randomly assigned to three groups that see different arguments in the context of tax simplification (renewed randomization). See below for more info.

- **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 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 the 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.

3.3 Randomized Survey Experiments

We include two randomized components into our survey block on tax simplification (see the survey structure above). 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. Tests for balance across experimental groups are presented and discussed in Appendix section A.2). 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 set-up 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 set-up with two experiments (each three groups) is advantageous w.r.t. statistical power.[14]

[13]The Factually Anonymous Income Tax Statistic (FAST) is a 10% stratified random sample of the German Income Tax Statistics, comprising information about taxable income, family situation, income sources, granted deductions and exemptions, revenues and sources of revenues, income tax burden, etc. The data are available as cross-section scientific use files. For the tax-expenditure calculations here we use the most recent available year of 2010. See Boenke and Schroeder (2017) for more information.

[14]Survey experiments are not part of the seminal classification of experiments (from lab to natual field experiment) provided by Harrison and List (2004). Generally, our survey experiment combines a high level of control (because we design the treatments and randomly assign people) and generalizability (because our experiment is embedded in a survey that is designed to be representative of the German population). The obvious difference to the experiments classified in Harrison and List (2004) is that we measure stated attitudes and preferences rather than actual behavior/revealed preferences.

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. The opener again defines how we interpret tax complexity in the context of the survey and explains that tax expenditures potentially contribute to the complexity of the tax system. Importantly, all participants (control group and treatment groups) see this opener. The opener therefore ensures that the issues of tax expenditures and complexity are made equally aware to control-group participants and treatment-group participants. In addition, the opener ensures that respondents in all three experimental groups (incl. 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 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; all the content of the treatments is reflected in the cited literature and presented arguments in section 1 and 2.1. In each of the two experiments, one treatment aims to shift preferences towards tax simplification and the other one aims to shift preferences away from tax simplification. Our treatments thus reflect that there are arguments both against and in support of tax simplification. In addition, we are able to investigate if preferences for tax simplification are more elastic with respect to arguments in support or against tax simplification.[15] We describe and motivate the two experiments in the following. Screenshots of the experimental treatments and of the opener statements provided to all experimental groups (including the control conditions) are provided in Appendix H.

Experiment 1. The first experiment includes two treatments which we label *Redistribution* treatment and *Avoidance* treatment. The two treatments are preceded by an *opener that is shown to all respondents (i.e., both treatment groups and control group).* The opener 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 redistributional 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. Those respondents who are initially not aware of the link between tax expenditures and social-policy aspects might reconsider their tax-simplicity

[15] The treatment structure is *not* augmented, meaning that respondents in the second treatment group do not see both the arguments in the first treatment group and the arguments from the second treatment group, but only see the respective statements from the second treatment group.

preferences in response to the treatment and become less supportive of tax simplification. 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-policy role, particularly in relation to income redistribution. 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. Assuming that most people disapprove tax avoidance and evasion, respondents who were initially not fully aware of the potential link between tax complexity and avoidance/evasion might shift their preferences towards more simplification in response to being exposed to this treatment. 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.*

Experiment 2. The second experiment includes two treatments, labeled *Efficiency* treatment and *Special interest* treatment. The two treatments are preceded by an *opener that is shown to all respondents (i.e., both treatment groups and control group).* As before in the first experiment, this opener is everything that control-group respondents see in the context of this second experiment. 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, and potentially shifts preferences away from tax simplification – at least among those respondents who did not consider this argument initially. 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 favorable tax treatment and the existence of many complexity-adding exemptions facilitates the groups' efforts; a system with a narrow tax base and without tax expenditures would make it more to difficult to implement special interests in the tax system. 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, that follows right after the opener, is formulated as follows:

> *One argument that is often used in favor 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.*

Discussion of Experimenter Demand Effects, Statistical Power and Prior Knowledge. One frequently raised concern with survey experiments (and surveys and experiments in general) is that experimenter demand effects drive the survey responses and results. Our main survey question, preferences for tax simplicity, asks respondents for their view on a specific aspect of policy. There is neither a correct or false answer to this question, nor is it in anyhow 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, and 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 below). 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. We also provide evidence that our pro-complexity arguments indeed increase awareness for the purpose of tax expenditures in the very long run, as shown in Section 5.4, thus also alleviating concerns of experimenter demand bias in the main intervention (see the discussion of follow-up surveys and experimenter demand in information experiments by Haaland et al. 2021).[16]

With respect to statistical power, the review paper by Haaland et al. (2021) suggests that randomized survey experiments should have about 700 observations per treatment arm. The number of observations in our experimental treatment arms are in accordance with this suggestion ($N = 2432$ with an even distribution of observations across the three experimental groups).[17]

[16] Recent studies such as Mummolo and Peterson (2019) and de Quidt et al. (2018) explicitly study experimenter demand effects and do not find strong evidence for their existence.

[17] Note that performing ex-ante power analyses during the design stage of our survey experiment was very difficult. The main survey question, preferences for tax simplicity, has neither been included in the GIP before

While we are not able to measure to which extent the provided treatment information constitute news to the respondents, we can rely on a 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 5.4 for details). Focusing on those respondents who are in the control group of both our experiments and who are thus not affected by the treatment information, we find that a substantial fraction of respondents is not aware of any objectives of expenditures and many believe their prime purpose is to reduce the tax burden.

4 Results Set 1: Documenting and Characterizing Attitudes Towards Tax Simplification

This section presents the survey results in the context of the paper's first objective, where we aim to document and characterize preferences for tax simplification. We first document the preferences for tax simplification (Q2) and then report the answers to a question about the desired tax system and summarize further results that help us to understand simplification preferences in more depth. Appendix B reports more findings that help to improve our (descriptive) understanding of tax complexity attitudes (e.g., we study the anatomy of tax complexity preferences there).

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; i.e., 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. Note in this context that regressions to study the anatomy of simplification preferences control for the experimental group of a given participant.

Preferences for Tax Simplification. We elicit preferences for tax simplicity using a question which surveys whether people believe that the income-tax system in Germany needs to be simplified (Q2 in the survey structure above). The reply categories were on a 6-point scale from 1 'Absolutely not' to 6 'Absolutely'.

Figure 2 presents the share of respondents in each reply category across all survey participants. A large majority believes that the tax system needs to be simplified: Among all

nor are we aware of any other survey that includes a similar question. It was therefore not possible to rely on any reliable predictions regarding the standard deviation (and mean) for our main survey question at the point of time when we designed the survey experiment. In light of a lack of comparable studies, we could neither form any good expectations regarding the effect sizes that would occur from our treatment interventions. However, these parameters are of course crucial for a meaningful power analysis. In addition, we faced a given number of participants in the GIP and it would have been difficult to adjust the sample size in response to the results of an ex-ante power analysis. We therefore do not present the results of any ex-ante power analyses.

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 (either category 5 or 6) preferences for tax simplification.[18] 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. We thus confirm that the prevailing view indeed is in favor of tax simplification.

Which Simplifying Tax Reform? In light of the conventional wisdom that tax simplification is desirable and to understand which tax system people desire, we included a question to survey *how* policy should reform the tax system in order to make it simpler (Q5). For the purpose of this question, we provided respondents with a list of potential tax-simplifying reforms and they could chose which of the offered alternatives they prefer. This list is of course not exhaustive, it yet 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.[19]

The results for this question are summarized in Figure 3. The most frequently chosen type of reform (33%, control group: 32.7%) 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 which features the same amount of tax expenditures as in the status quo. 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 a return. 6% (control group: 6%) of our respondents have a preference for keeping the status-quo and do not implement any tax reform.

Importantly, note that these results do *not* mean that only half of the respondents desire tax simplification by means of reducing deduction possibilities and allowances (on first glance, one might be tempted to think so because reply categories 'flat-rate system without any deductions and allowances' and 'increased progessivity along with abolishment of all deductions and allowances' were chosen by 15% and 33% of the respondents, respectively). In contrast, a strong preference for tax simplification can, as we show below, also imply that some, but not all, deduction and allowance possibilities are eliminated. Hence, the finding that half of the sample is in favor of eliminating *all* deductions and allowances is quite staggering. In addition, supporting a system with pre-filled tax returns (as done by 17% of respondents) potentially also indicates a preference to eliminate or reduce deductions and allowances because filing these deductions and allowances is obsolete in a system with pre-filled returns. Overall, the results

[18] 53% (= 47.7/(100 − 9.6)) checked reply category six, meaning that the system 'absolutely' needs to be simplified. In the control group, 54% of the respondents checked reply category six. Another 23% of the respondents chose the second highest reply category 5 (control group: 25%). 16% (control group: 13.4%) are in category 4, which also implies a certain preference for tax simplification.

[19] See Q5 in Appendix section G.1 for the detailed question design.

suggest that the share of respondents who favor to eliminate some deductions and allowances is considerably larger than 50%. This is also reflected in the previously mentioned observation from our follow-up survey that tax expenditures and the direct consequences of tax expenditures are believed to be the prime source of tax complexity in the German tax system (see Figure B.5).

Overall, the policy-support results suggest that there is no consensus w.r.t. 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 investigated if the policy-reform preferences are affected by our first experimental intervention (recall that the second experiment was implemented after the reform-survey question). We do not find any evidence that this is the case.

Further Survey Questions. Investigating the *anatomy* of simplification preferences (Appendix Table B.1), we find that age and gender are important demographic correlates of simplification preferences. Another important correlate is the perceived difficulty of filing a tax return. Respondents who find it easy to file a tax return have lower simplification preferences than respondents 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* (Appendix Figure B.1). Overall, there is a tendency that tax returns are perceived to be fairly difficult, but the picture is not as strong as in the question on preferences for tax simplification. This corresponds with the 'anatomy' result above: the positive correlation between simplification preferences and perceived difficulty to file a return is not perfectly linear. This suggests that the strong preference for tax simplification is not entirely motivated by own experiences with too-difficult tax returns.

Germans further believe that the tax system in their country is in strong *need of reform* (Appendix Figure B.2). We further survey beliefs about the *distributional implications of tax expenditures* and observe that the large majority believes that tax expenditures mostly benefit richer taxpayers (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' (Appendix Figure B.4).[20]

[20]This 'standard deduction' represents the lump sum deduction amount for taxpayers who do not exceed the thresholds in other deduction categories.

5 Results Set 2: Role of Awareness

This section presents the results in the context of our second objective, where we analyze to which extent attitudes towards tax simplification are affected by awareness with respect to the pros and cons of tax simplification. We use four strategies towards this objective. First, we study in section 5.1 if participants believe that the tax system should account for differences in living situations of taxpayers. Second, we present the results of two randomized survey experiments in section 5.2. Third, section 5.3 reports how the support for tax simplification evolves over the course of the survey. Fourth, section 5.4 provides a text analysis of an open-ended question on the perceived objectives of tax expenditures and studies persistence of treatment effects.

5.1 Different Tax Burden for Taxpayers in Different Living Situations?

We confront respondents with two fictitious taxpayers who differ in one aspect of their living situation, and then ask if these two taxpayers should pay the same amount of taxes (Q6-Q8). While a differential tax treatment of the two fictitious taxpayers would add complexity to the tax system, we do not mention this complexity aspect of the presented scenario explicitly to respondents. These questions allow us to evaluate if respondents prefer to account for different living situations through the tax system at the cost of adding complexity to the tax system. In other words, if people indicated that specific differences in living situations should matter for the tax burden, this would imply that they do not desire the simplest possible tax system and have a preference to keep certain complexity-adding features of the tax system.

We ask respondents in all three questions to imagine two fictitious taxpayers, A and B, who are comparable in all tax relevant aspects, and only differ along one of the following dimensions: i) Person A has to spend a considerable amount for the elderly care of her mother, while Person B does not have to bear such costs (Q6). ii) Person A spends a considerable amount of income on charitable giving, while Person B does not donate (Q7). iii) 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 persons, A or B, should pay more taxes (where the order of reply categories and the order of presented scenarios was randomized). The results are presented in Figures 4a to 4c. Note that the responses here were not affected by the randomized experimental interventions. We yet report along the results for respondents who were in the control group of the first experiment.

Figure 4a 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% (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 A should pay more in taxes. Figure E.1a 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. 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 bit more than 60% (control group: 60%) think that Person B should pay more.[21] Overall, these survey responses provide clear evidence that people favor a system in which the costs for elderly care are deductible from the tax base.

For the survey questions regarding charitable donations (Figure 4b) and expenses for commuting (Figure 4c) 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, a fraction of 32% (control group: 30%) and 39% (control group: 39%) of all respondents yet think that higher donations and commuting costs should imply lower tax burdens. That is, roughly 1/3 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 favor 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.[22] Consistent with the literature on optimal taxation and equality of opportunity (see review of related literature), our survey respondents have the intuition that circumstantial differences should be accounted for to a larger extent by the tax system than 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 Section 5.1.1 further 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. Figures E.1b and E.1c 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 reduce in a lower tax burden. Among those who do use the commuting expenditure, the share of people who believe commuting should reduce tax payments stands at 47% (control group: 46.7%). Overall, 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. However, even for donations and commuting we still 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

[21] 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.

[22] Commuting might be perceived as a circumstance in certain specific situations; for example when a worker is relocated to another branch of her firm. However, the decision where to live (close to work or not) will usually be a choice.

self-interest.

In addition, we also explore which share of respondents indicate in all of the 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 to have a simpler tax system.

Relation to other Survey Questions. How are the results of this exercise associated with other survey questions (which we reported on in our first set of results in Section 4)? In particular, we relate the results of our 'A-vs-B' questions to the question on preferences for tax simplicity (Q2) and to the question surveying *how* policy should reform the tax system in order to make it simpler (Q5).

If preferences for tax simplification were not driven by awareness deficits, we would expect that there is a substantial correlation between reporting strong preferences for tax simplification (in Q2) and not wanting to account for different living situations in the tax burden. The reason is that the preference for tax simplicity should persist as we ask respondents about concrete and understandable situations if people were fully aware about the practical implications for tax simplification. To shed light on the relationships, we compute correlations between a dummy variable indicating that the tax system should *not* differentiate between person A and person B and preferences for tax simplification. We find that the correlation is fairly weak for all of the three living situations considered. The correlation coefficient (based on control-group respondents) is 0.054 for the commuting-to-work question, 0.086 for the question on donations, and 0.055 for the elderly-care scenario.[23] Expressed differently, among those with the highest support for tax simplicity, only 40% think that both taxpayers in the elderly-care situation should pay the same amount of taxes. For the scenarios relating to commuting-to-work and donations, 61% and 69%, respectively, of respondents with the strongest support for tax simplification believe that both taxpayers should pay the same amount in taxes.

We also examine how responses regarding the preferred tax reform (Q5) align with our results for the 'A-vs-B' questions. 70% of those respondents who prefer to abolish *all* allowances and deductions (see Figure 3) indicated at least once in the three 'A-vs-B' questions that the two taxpayers should not pay the same amount of taxes. In other words, among those who first indicate that they wish to eliminate all allowances and deductions, 70% indicated at least once to maintain a deduction/allowance when asked in the context of a concrete application. This is further evidence supporting the notion that people are not fully aware of the implications of simplification; the concrete application makes them aware of these implications and they then respond differently than in a setting without concrete application. Supporting this interpretation, the open-ended questions in our follow-up survey are also indicative for a lack of awareness of the general public regarding the concrete purpose of tax expenditures (see Section 5.4).

[23]The non-parametric spearman's rank correlation coefficients give similar results (0.055 for donations, 0.078 for commuting-to-work and 0.07 for elderly care).

5.1.1 Follow-up Survey Experiment on the Role of Circumstances and Choices

In order to shed more light on the role of circumstances and choices in evaluating complexity-adding tax expenditures, we conducted an additional randomized survey experiment in a follow-up survey (see section 3). 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 w.r.t. 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 two commute. In a control group, we did not specify the reason for A's long commute. In treatment *Circumstance*, it is specified that "Person A was relocated by his employer and has to travel a considerable distance to work ever since." In treatment *Choice*, we specify that "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 for a job with a very long way 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 matter significantly to respondents. The share of respondents who think that 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. 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.

5.2 Randomized Survey Experiments

Empirical Strategy. We now present the results of the two randomized survey experiments that were implemented in our main survey. To analyze the first experiment, we use OLS regressions (with robust standard errors) in which we regress the respective outcome variable on dummy variables indicating the two randomly received 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 of the treatment effects.

We expect that the treatments of the first experiment 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 with 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 as follows in our analysis of the second experiment: We first fully interact 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),

and then use OLS (with robust standard errors) to regress the outcome variable of interest on the full set of interactions. We only report the coefficients of the treatment dummies of the second experiment (and not the interactions). These reported coefficients then 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.[24] As with the first experiment, our preferred specifications include control variables which improve precision of the treatment coefficient of interest.

The main outcome variables are the responses to the question of whether the tax system should be simplified; i.e., Q2 in the case of the first experiment and Q9 in 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 just as the original survey question, on a six-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). This variable is also coded as the original survey variable (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.

5.2.1 Experiment 1

Main Effects. Table 1 presents the main results for the first experiment, in which we provide arguments about the social-policy role of tax expenditures (*Redistribution* group) and about expenditure-induced tax avoidance opportunities (*Avoidance* group) in complex tax systems.

Column (1) of the table 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% ($-0.133/5.22$ = coefficient/control-group average), relative to the control-group average. 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 the public and press outlets, presumably is strongly anchored among respondents. Our treatment thus affects preferences for tax simplification *although* the conventional wisdom on the topic is very clear and

[24]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.

strong. For these reasons, we argue that the effect size should be interpreted as non-negligible.

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) to (5) (with p-values in the range of 0.064 to 0.057).[25] 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.

The results thus show that preferences for simplification are not elastic to the argument *in support of* tax simplification. This null-result might be explained with the very prominent role of arguments in favor of simplification in the public debate. As a result of these salient arguments, participants presumably have less 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); this suggests that there would have been some room for the pro-simplification argument to increase the support for simplification.[26]

We also investigated the effect of the first experimental intervention on respondents' views about the distributional implications of tax complexity (Q4) – see Appendix Table C.2. The *Redistribution*-treatment does not affect these beliefs (relative to the control group). However, the argument about possible complexity-induced avoidance possibilities in the *Avoidance*-group somewhat affects the distributional beliefs. The treatment coefficient is statistically significant in specifications (3) to (5), and indicates that the treatment increases beliefs that tax expenditures add to income inequality (the coefficients in specifications (1) and (2) are imprecisely measured). Considering the specification in column (5), which includes all covariates, the treatment increased the distributional-beliefs variable by about 4% (0.169/4.285), relative to the control-group average. Comparing the coefficients of the *Redistribution*-treatment and the *Avoidance*-treatment, we find statistical significant differences for specifications (3) to (5) with p-values ranging from 0.09 to 0.064.

Studying the effects on further survey questions (Q3, Q5, Q6, Q7, Q8), we do not detect any treatment effects and therefore do not discuss the results any further (Appendix Table C.3).

Robustness. 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

[25] Negative coefficients of the *Avoidance*-treatment are consistent with the treatment having no effect. If two independent samples are drawn from the same population, it is very likely that one sample is smaller than the other one.

[26] 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. This suggests that our control variables can explain a considerable part of the high baseline support for simplification.

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 Appendix Section 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 Appendix Tables C.4 and C.5.

Heterogenous Treatment Effects. In a next step of the analysis, we investigate if particular groups of respondents responded differently to the treatments of the first experiment than other groups. For this purpose, we interact the treatment-group dummies with the following observable characteristics: age, gender, marital status, household size, income, education, 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). Overall, the effects of the treatments seem to be homogenous. We mostly do not see any significant interactions. 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 homogenous 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 Appendix Section D).

5.2.2 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. This second experiment includes a control group, a group that is presented an *Efficiency* argument against tax simplification, and a group that is presented a *Special interest group* argument in favor of tax simplicity. The dependent variable is the question surveying tax-simplicity preferences (note that Q9 is the dependent variable here, not Q2 which we use 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 favor 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 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 favor 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 only compare reactions 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 favor of simplification in the public debate, which reduce misperceptions regarding pro-simplification arguments.

Robustness. Statistical inference is robust to adjusted standard errors using randomization tests and tests for multiple comparisons – see Appendix Section D for details. As shown in Table C.6, the above results are robust to using ordered probit models.

Heterogenous Treatment Effects. As in the case of the first experiment, we investigate if particular groups of respondents respond differently to the treatments of the second experiment than other groups. We run the same interaction models as in the case of the first experiment (with the same interacted observable variables) and again find that effects of the treatments are very homogenous across different demographic groups; we mostly do not detect any significant interactions. For reasons of brevity and given these results, we again do not report the regression results. We acknowledge, again, that it is possible that the interaction models for detecting heterogeneity lack statistical power, rather than providing evidence of homogenous effects. However, the lack of heterogeneity is consistent with the finding that exact p-values following Young (2018) are very similar to the classical p-values (see Appendix Section D).

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 Table C.1, 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.1 that the second effect 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). Despite not being statistically significant at conventional levels, the direction of the respective estimates are plausible in the sense that those who received a contra-

simplification argument in both experiments are less in favor of tax simplification than those who received it only once. Also, respondents who, in both experiments, received arguments which are in favor of tax simplification (i.e., Avoidance and Special-Interest-Group arguments in the first and second experiment, respectively) support tax simplification more than those who only received a pro simplification argument in the first experiment. Receiving conflicting arguments, i.e., an argument in favor 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 favor of tax simplification (or contra tax simplification) in the first experiment.

5.3 How does the Support for Tax Simplification Evolve over the Course of the Survey?

We now analyze how the support for tax simplification differs between the beginning of the survey (Q2) and the end of the survey (Q9). Differences in simplification preferences between Q2 and Q9 potentially suggest that the mere exposure to the survey questions and the reflection on the topic in the course of the survey affect preferences. This sheds further light on the question of whether a lack of awareness can explain the initial high support for simplification. This approach is similar in spirit to the approach in Alesina et al. (2023), where the order of question blocks is varied to investigate if the exposure to questions about a certain topic affects beliefs.

The main results of this exercise are presented in Table 3. In this table, we present 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. These respondents were not exposed to the provision of any arguments (neither against nor in favor) 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 a 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. The difference may not appear enormously high, 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 outweighted by the priming effect): We see throughout all three groups of the second experiment that the support for tax simplicity 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 and therefore even small shifts in preferences testify that the mere exposure to the topic indeed matters.

Across all respondents (thus not analyzing by treatment groups), we also see a significant decline in simplification preferences between Q2 and Q9. A Kolmogorov-Smirnoff 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). These distributions are plotted in Figure E.2.

Other combinations of treatment-group status are also interesting and all 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 favor simplification in the first experiment and were then presented 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 then were presented an argument in favor of tax simplicity before replying to Q9 (i.e., those in the *Special interest* group of the second experiment).

5.4 Reasoning about Tax Expenditures and Persistence of Treatment Effects

To improve our understanding of how people reason about tax expenditures and to study the persistence of our experimental interventions, we fielded an open-ended question in our follow-up survey.[27] The motive for using an open-ended question was to obtain an unprimed picture regarding the question of which objectives and goals people associate with tax expenditures (Ferrario and Stantcheva 2022). Considering that our survey effects may have persistent effects, the unprimed picture particularly arises among those subjects who were in the control group of both experiments. Table F.1 therefore reports the findings of this 'pure' control 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 participated in the survey and 4130 ultimately answered the question at hand.

We use modern text-analysis techniques to analyze 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 stopwords 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

[27] This was implemented in GIP wave 46, March 2020, about 20 months after the initial main experiment. This is the same survey wave that includes the additional survey experiment on the role of choice vs. circumstances (see section 5.1.1).

example, some of these answers were just single letters like "X" or ".". In a third step, we manually create a training data set consisting of 3046 observations by categorizing our data set 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*. Appendix Section F provides further details on these categories. We use this training data set to predict the remaining 805 observations using a SGD Classifier. Appendix Section F outlines and discusses the specifics of the employed empirical approach.

Our first results are summarized in Table F.1 which displays the frequency with which different types of objectives were stated by the respondents. The most frequently stated objective of tax expenditures is the reduction of one's own tax burden (chosen by 22% of respondents). 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%). 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 only by a small share of people; such other classes of objectives include (frequency as depicted in parentheses): incentivizing economic behavior (6.8%), serving the rich or special interests (4.2%, respectively), preserving tax equality (5.3%), and accounting for special circumstances (9.5%). 1.6% and 3.7% believe that tax expenditures explicitly exist to make the tax system complex or easy, respectively. Please note that there is also a fraction of about 15.3% of overall responses which our algorithm could not categorize as one of the aforementioned categories, but still resemble valid answers.

We also use the open-ended questions to study the persistency of the treatment effects of our main survey experiment. While we established that arguments in favor of complex taxes lead to lower acceptance of tax simplification and that people have very low understanding of what the primary goals of tax expenditures are, it is yet unclear whether information provision has lasting effects on the awareness for these issues. We test this by studying whether response rates for the open-ended text questions in the follow-up survey are different across different experimental groups from our main survey (i.e., 20 months prior to the follow-up survey). Table 4 shows that treatment groups exposed to arguments in favor of more complexity (redistribution and economic efficiency conditions) have significantly lower rates of not replying in the follow-up survey, while neither argument against complexity shows lasting effects. The effect sizes of being exposed to arguments in favor of complexity on non-response are also sizeable, with about 16.9% (0.034/0.201) and 19.9% (0.04/0.201) relative to the control mean.

Therefore, providing information about tax expenditures in favor of more complexity did not only lead to lower support for tax simplification in the main survey, but also to a higher willingness to answer questions and reflect on the topic even 20 months later. This suggests that these treatment arguments increased the saliency of the topic and motivated people to follow up

on the topic in the long-run. The asymmetry between arguments in favor and against simplicity again indicates that the arguments in favor of complexity were more surprising to respondents. Given that effects persist and are observable in another survey wave with a significant time lag, this also alleviates concerns of experimenter demand in the main survey (see Haaland et al. 2021 for a related discussion). We do not find systematic evidence for persistent treatment effects on the choice of the classified text categories (as described above). This may indicate that while there may be higher salience for the topic per se through certain information, there may not be information updating in the long-run. On the latter, however, we cannot draw definite conclusions since these results are unfortunately confounded by a systematic change in the choice to answer the open-ended question through the experiment.

6 Conclusion

The prevailing view in the literature and public debate seemingly is that most tax systems are overly complex and should therefore be simplified. However, there also are economic arguments in support of a certain degree of tax complexity and it is puzzling why tax systems remain highly complex despite the conventional view in favor of more simplification. Using new experimental and survey data for a representative sample of the German population, we find that most people are indeed in favor 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 of tax simplification. Providing arguments in favor of tax expenditures has also long-lasting effects and increases the saliency of the latter in a sizable manner even 20 months after the main intervention. In addition, there is heterogeneity among the population with respect to the way towards more simplicity.

Our results indicate that awareness about the exact implications of simplifying reforms can change attitudes towards reforms. More awareness can thus moderate the strong views that people apparently have in the context of simplification. This then implies 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). These discussions may even have long-lasting effects on the relative salience of the topic. Recent work shows that information deficits among individuals in the context of taxation can induce governments to implement inefficient tax policy (Boccanfuso and Ferey 2019). More nuanced discussions and better information about the implications of taxation could also mitigate this source of inefficiency. This may also be true for public perceptions of other domains of economic policy since people often fail to take equilibrium effects into account (Dal Bó et al. 2018).

We also speak to the puzzle of why tax systems remain complex although there seems to exist a general consensus to simplify tax systems. For example, our *A-vs-B* questions show that people find certain expenditures useful and therefore demand keeping them. This finding,

together with the considerable general support for simplification and the heterogeneity about the preferred system, then potentially suggests that, while there is no explicit policy advocacy in favor of tax complexity, complex tax systems can be viewed as an unintended consequence of a set of many well-intentioned and publicly demanded policies. Put differently, 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.

What are the implications for the political feasibility of tax reforms? Our results imply that voters do support tax simplifying reforms. However, they do not demand the simplest possible tax system and have preferences to keep certain complexity-adding elements of the tax system. 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 which are not self-inflicted (e.g., costs for elderly care).

References

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

Aghion, P., U. Akcigit, M. Lequien, and S. Stantcheva (2017). Tax simplicity and heterogeneous learning. NBER working paper no 24049.

Alesina, A. and G.-M. Angeletos (2005). Fairness and redistribution. *The American Economic Review 95*(4), 960 – 980.

Alesina, A., A. Miano, and S. Stantcheva (2023). Immigration and redistribution. *The Review of Economic Studies 90*(1), 1–39.

Alesina, A., S. Stantcheva, and E. Teso (2018). Intergenerational mobility and support for redistribution. *American Economic Review 108*(2), 521 – 554.

Ballard, C. L., S. Gupta, et al. (2018). Perceptions and realities of average tax rates in the federal income tax: Evidence from michigan. *National Tax Journal 71*(2), 263–294.

Benzarti, Y. (2020). How taxing is tax filing? using revealed preferences to estimate compliance costs. *American Economic Journal: Economic Policy*.

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review 105*(11), 3489–3529.

Bierbrauer, F., P. Boyer, and A. Peichl (2020). Politically feasible reforms of non-linear tax systems. *American Economic Review*.

Blattman, C., J. C. Jamison, and M. Sheridan (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *American Economic*

Review 107(4), 1165–1206.

Blaufus, K., J. Bob, J. Hundsdoerfer, D. Kiesewetter, and J. Weimann (2013). Decision heuristics and tax perception–an analysis of a tax-cut-cum-base-broadening policy. *Journal of Economic Psychology 35*, 1–16.

Blaufus, K., S. Eichfelder, and J. Hundsdoerfer (2014). Income tax compliance costs of working individuals: Empirical evidence from germany. *Public Finance Review 42*(6), 800–829.

Blaufus, K. and R. Ortlieb (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.

Blesse, S. and F. Heinemann (2020). Citizens' trade-offs in state merger decisions: Evidence from a randomized survey experiment. *Journal of Economic Behavior and Organization 180*, 438–471.

Blom, A. G., C. Gathmann, and U. Krieger (2015). Setting up an online panel representative of the general population: The German Internet Panel. *Field Methods 27*(4), 391–408.

Blom, A. G., J. M. Herzing, C. Cornesse, J. W. Sakshaug, U. Krieger, and D. Bossert (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., B. J. Ruffle, and Y. Ganun (2012). Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods. *European Economic Review 56*(6), 1200–1219.

Boccanfuso, J. and A. Ferey (2019). Inattention and the taxation bias. mimeo, online at `https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxmZXJleWFudG9pbmV8Z3g6NTRlZThiM2UwOWJmYTRhNQ`.

Boenke, Tim, J. B. and C. Schroeder (2017). Fiscal federalism and tax administration: Evidence from germany. *German Economic Review 18*(3), 337–409.

Brennan, G. and J. M. Buchanan (1980). *The Power to Tax*. Cambridge University Press.

Brown, J. R., A. Kapteyn, E. F. Luttmer, and O. S. Mitchell (2017). Cognitive Constraints on Valuing Annuities. *Journal of the European Economic Association 15*(2), 429–462.

Brusco, S., L. Colombo, and U. Galmarini (2014). Tax differentiation, lobbying, and welfare. *Social Choice and Welfare 42*(4), 977–1006.

Chetty, R. and E. Saez (2013). Teaching the tax code: Earnings responses to an experiment with EITC recipients. *American Economic Journal: Applied Economics 5*(1), 1–31.

Dal Bó, E., P. Dal Bó, and E. Eyster (2018). The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies 85*(2), 964–998.

de Bartolome, C. (1995). Which tax rate do people use: Average or marginal? *Journal of Public Economics 56*(1), 79–96.

de Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and bounding experimenter demand. *American Economic Review 108*(11), 3266–3302.

Ding, P. (2017). A paradox from randomization-based causal inference. *Statistical science 32*(3), 331–345.

Doerrenberg, P. and A. Peichl (2022). Tax morale and the role of social norms and reciprocity. evidence from a randomized survey experiment. *FinanzArchiv/Public Finance Analysis 78*(1-2), 44–86.

Dolls, M. and N. Wehrhoefer (2021). Attitudes towards euro area reforms: Evidence from a randomized survey experiment. *European Journal of Political Economy*.

Durante, R., L. Putterman, and J. van der Weele (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association 12*(4), 1059–1086.

Economist (2005). Simplifying tax systems. the case for flat taxes. The Economist article, online at `https://www.economist.com/special-report/2005/04/14/the-case-for-flat-taxes`.

Economist (2013). Tax reform in america. simpler, fairer, possible. The Economist article, online at `https://www.economist.com/leaders/2013/07/13/simpler-fairer-possible`.

Engelmann, D., E. Janeba, L. Mechtenberg, and N. Wehrhoefer (2020). Preferences over taxation of high income individuals: Evidence from survey and laboratory experiments. CESifo working paper no. 8595.

FAST (2010). *Faktisch anonymisierte Lohn- und Einkommensteuerstatistik by the Federal Statistical Office*. Wiesbaden, Germany.

FAZ (2004). Der steuermann. Frankfurter Allgemeine Zeitung (FAZ) article, online at `https://www.faz.net/aktuell/wirtschaft/reformer-merz-der-steuermann-1192866.html`.

FAZ (2005). Kirchhof-modell: Die reichen profitieren am meisten. Frankfurter Allgemeine Zeitung (FAZ) article, online at `https://www.faz.net/aktuell/wirtschaft/wirtschaftspolitik/steuern-kirchhof-modell-die-reichen-profitieren-am-meisten-1260403.html`.

Feldman, N. E., P. Katuščák, and L. Kawano (2016). Taxpayer confusion: Evidence from the child tax credit. *American Economic Review 106*(3), 807–35.

Ferrario, B. and S. Stantcheva (2022). Eliciting people's first-order concerns: Text analysis of open-ended survey questions. *American Economic Association Papers and Proceedings*.

Financial Times (2006). We must 'keep it simple' to clear the tax jungle. Opinion piece by geoffrey howe, march 21, 2006; online at `https://www.ft.com/content/b460b7e8-b90f-11da-b57d-0000779e2340`.

Fisher, R. A. (1935). *The Design of Experiments*. Oliver and Boyd, Ltd.

Fisman, R., K. Gladstone, I. Kuziemko, and S. Naidu (2020). Do americans want to tax wealth? evidence from online surveys. *Journal of Public Economics 188*, 104207.

Fochmann, M., N. Müller, and M. Overesch (2018). Less cheating? the effects of prefilled forms on compliance behavior.

Forbes (2017). Trump wants a simple tax system – history says he won't get it. Forbes article, online at `https://www.forbes.com/sites/taxanalysts/2017/09/11/trump-wants-a-simple-tax-system-history-says-he-wont-get-it/#4c3b63e27c1c`.

Fuest, C., A. Peichl, and T. Schaefer (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 P. Profeta (2009). Tax complexity with heterogeneous voters. *Public Finance and Managemente 9*(2), 1–35.

Gideon, M. (2017). Do individuals perceive income tax rates correctly? *Public Finance Review 45*(1), 97–117.

Goettinger Tageblatt (2018). Friedrich Merz' erster tweet – ein bierdeckel, online at: `http://www.goettinger-tageblatt.de/Nachrichten/Politik/Deutschland-Welt/Friedrich-Merz-erster-Tweet-ein-Bierdeckel`. Technical report.

Gordon, R. H. and W. Kopczuk (2014). The choice of the personal income tax base. *Journal of Public Economics 118*, 97 – 110.

Gravelle, J. G. and T. L. Hungerford (2012). The challenge of individual income tax reform: An economic analysis of tax base broadening. Congressional research service. washington, dc.

Gsottbauer, E., D. MÃ¼ller, S. MÃ¼ller, S. T. Trautmann, and G. Zudenkova (2022). Social Class and (Un)Ethical Behaviour: Causal and Correlational Evidence. Technical Report 647.

Haaland, I., C. Roth, and J. Wohlfart (2021). Designing information provision experiments. *Journal of Economic Literature*. forthcoming.

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic Literature 42*(4), 1009 – 1055.

Hettich, W. and S. Winer (1999). *Democratic Choice and Taxation*. Cambridge University Press.

Hettich, W. and S. L. Winer (1988). Economic and political foundations of tax structure. *American Economic Review 78*(4), 701–712.

Hines, J. R. (2016). High tax heresy. Working paper, available online: `https://pdfs.semanticscholar.org/2fc7/1dfd009ab42b5899815fe75cda368bdc178a.pdf`.

Hines, J. R. (2019). *Sensible taxes and practical policitcs*. Oxford University Press. Forthcoming 2020.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review 104*(2), 537–63.

James, S., A. Sawyer, and I. Wallschutzky (1997). Tax simplifications - a tale of three countries. *Bulletin for International Fiscal Documentation 51*(11), 493–503.

Kenyon, D. A. (1997). *Tax Policy in an Intergovernmental Setting: Is it Time for the U.S. to Change?*, pp. 61–97. Dordrecht: Springer Netherlands.

Kerschbamer, R. and D. Müller (2020). Social preferences and political attitudes: An online experiment on a large heterogeneous sample. *Journal of Public Economics 182*.

Kirchhof, P. (2011). *Bundessteuergesetzbuch. Ein Reformentwurf zur Erneuerung des Steuerrechts*. Heidelberg: C.F. Müller Wissenschaft.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica 79*(3), 651 – 692.

Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics 89*(11-12), 2093–2119.

Kopczuk, W. (2012). The polish business "flat" tax and its effect on reported incomes: a pareto improving tax reform? mimeo, online at `http://www.columbia.edu/~wk2110/bin/PolishFlatTax.pdf`.

Kopczuk, W. and C. Pop-Eleches (2007). Electronic filing, tax preparers and participation in the earned income tax credit. *Journal of Public Economics 91*(7-8), 1351–1367.

Kotakorpi, K. and J.-P. Laamanen (2016). Prefilled income tax returns and tax compliance: Evidence from a natural experiment. University of tampere, school of management, economics.

Liebman, J. B. and R. J. Zeckhauser (2004). Schmeduling. mimeo, Harvard Kennedy School.

Meltzer, A. H. and S. F. Richard (1981). A rational theory of the size of government. *Journal of political Economy 89*(5), 914–927.

Müller, D. and S. Renes (2021). Fairness views and political preferences: Evidence from a large online experiment. *Social Choice and Welfare*.

Mummolo, J. and E. Peterson (2019). Demand effects in survey experiments: An empirical assessment. *American Political Science Review*. forthcoming.

New York Times (1964). The tax jungle. December 21, 1964, page 28; online at `https://www.nytimes.com/1964/12/21/archives/the-tax-jungle.html`.

NPR (2015). Lots of candidates want to simplify tax code; here's what they get wrong. National Public Radio (NPR) online article. `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`.

NYT (2015). The tax code can be simpler. but not three pages. New York Times (NYT) article, online at https://www.nytimes.com/2015/11/15/upshot/a-three-page-tax-code-not-exactly-simple.html.

OECD (2010a). Tax design considerations. in *Tax Policy Reform and Economic Growth*, oecd publishing, paris. doi: https://doi.org/10.1787/9789264091085-7-en.

OECD (2010b). Tax expenditures in OECD countries. Oecd publishing, paris. doi: https://doi.org/10.1787/9789264076907-en.

Ooghe, E. and A. Peichl (2015). Fair and efficient taxation under partial control. *The Economic Journal 125*(589), 2024–2051.

Paetzold, J. and H. Winner (2016). Taking the high road? compliance with commuter tax allowances and the role of evasion spillovers. *Journal of Public Economics 143*, 1 – 14.

Pitt, M. and J. Slemrod (1989). The compliance cost of itemizing deductions: Evidence from individual tax returns. *American Economic Review 79*(5), 1224–32.

Rees-Jones, A. and D. Taubinsky (2016). Measuring "schmeduling". NBER working pape no. 22884r.

Rohaly, J. and W. G. Gale (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 D.C.: Brookings Institution Pres.

Saez, E. and S. Stantcheva (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review 106*(1), 24–45.

Slemrod, J. and W. Kopczuk (2002). The optimal elasticity of taxable income. *Journal of Public Economics 84*(1), 91–112.

Stantcheva, S. (2021). Understanding tax policy: How do people reason? *The Quarterly Journal of Economics 136*(4), 2309–2369.

Stantcheva, S. (2022). How to run surveys: A guide to creating your own identifying variation and revealing the invisible. Technical report, National Bureau of Economic Research.

Tax Policy Center (2019). What are tax expenditures and how are they structured? The Tax Policy Center's Briefing Book. a citizen's guide to the fascinating (though often complex) elements of the federal tax system, online at https://www.taxpolicycenter.org/sites/default/files/briefing-book/bb_full_2018_1.pdf.

Tran-Nam, B. (2000). Tax reform and tax simplicity: A new and 'simpler' tax system? *University of New South Wales Law Journal, The 23*(2), 241–251.

Tsankova, T., C. Imbert, M. Luts, and J. Spinnewijn (2019). 'how to improve tax compliance? evidence from countrywide experiments in belgium.

Vox (2017). Why democrats should support radically simpler taxes. Vox article, online at https://www.vox.com/policy-and-politics/2017/5/8/15442172/democrats-tax-plan-return-free-filing-trump-ambitious.

Wagner, F. W. (2006). Was bedeutet Steuervereinfachung wirklich? *Perspektiven der Wirtschaftspolitik 7*(1), 19–33.

Warskett, G., S. L. Winer, and W. Hettich (1998). The complexity of tax structure in competitive political systems. *International Tax and Public Finance 5*, 123–151.

Wartena, C. (2019). A probabilistic morphology model for german lemmatization. In *Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)*, pp. 40–49.

Weinzierl, M. (2014). The promise of positive optimal taxation: normative diversity and a role for equal sacrifice. *Journal of Public Economics 118*, 128 – 142.

Weinzierl, M. (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 S. S. Young (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*. John Wiley & Sons.

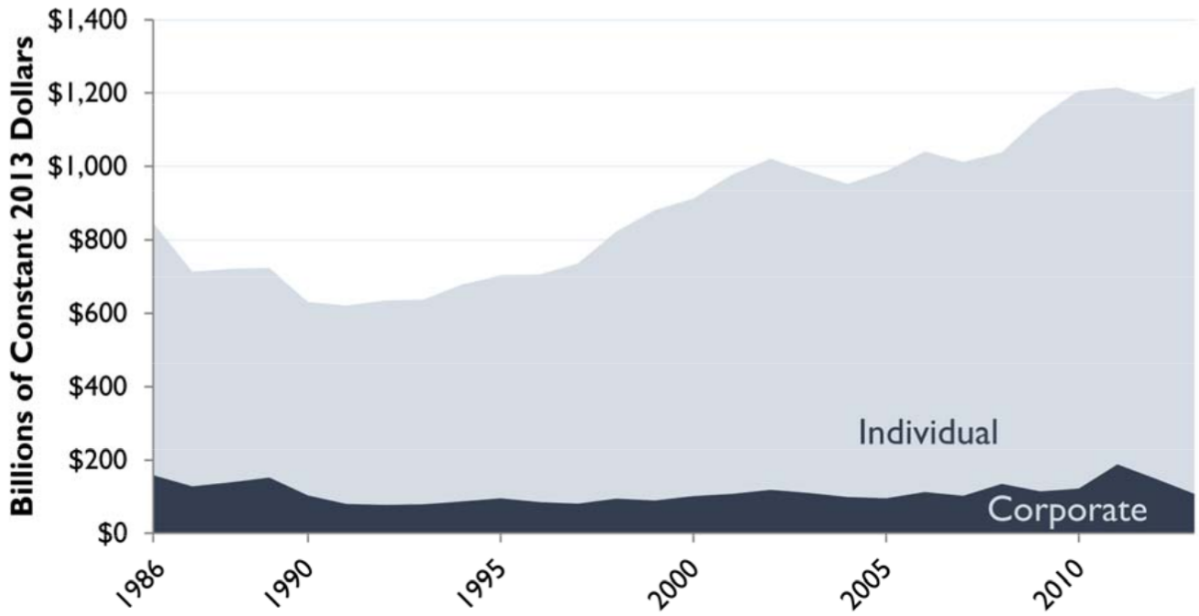
WiWo (2017). Pendler kosten den fiskus fuenf milliarden euro. WirtschaftsWoche (WiWo) article, online at `https://www.wiwo.de/politik/deutschland/berufspendler-pendler-kosten-den-fiskus-fuenf-milliarden-euro/20654060.html`.

Young, A. (2018). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics 134*(2), 557–598.

Zwick, E. (2020). The costs of corporate tax complexity. *American Economic Journal: Economic Policy*. forthcoming.

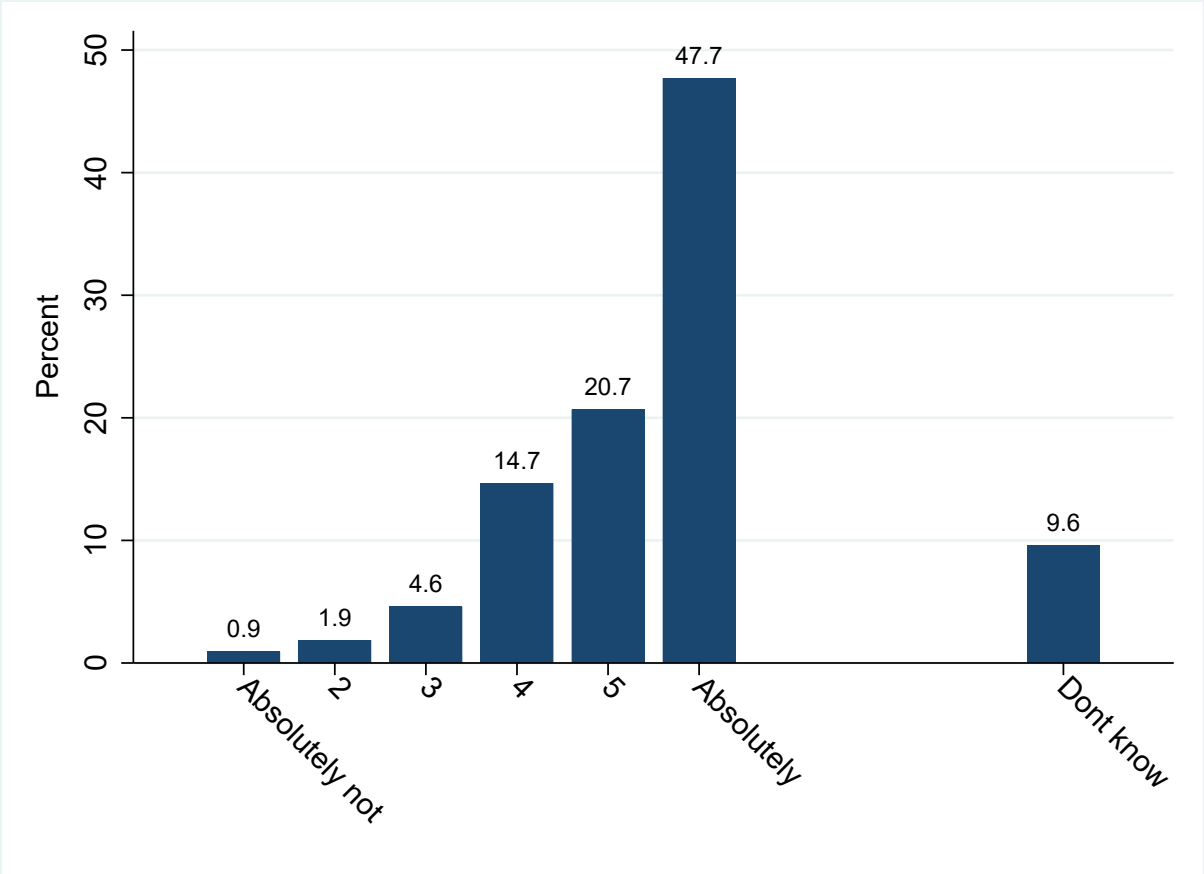
Main Figures and Tables

Figure 1: Growth of Tax Expenditures over Time in the US



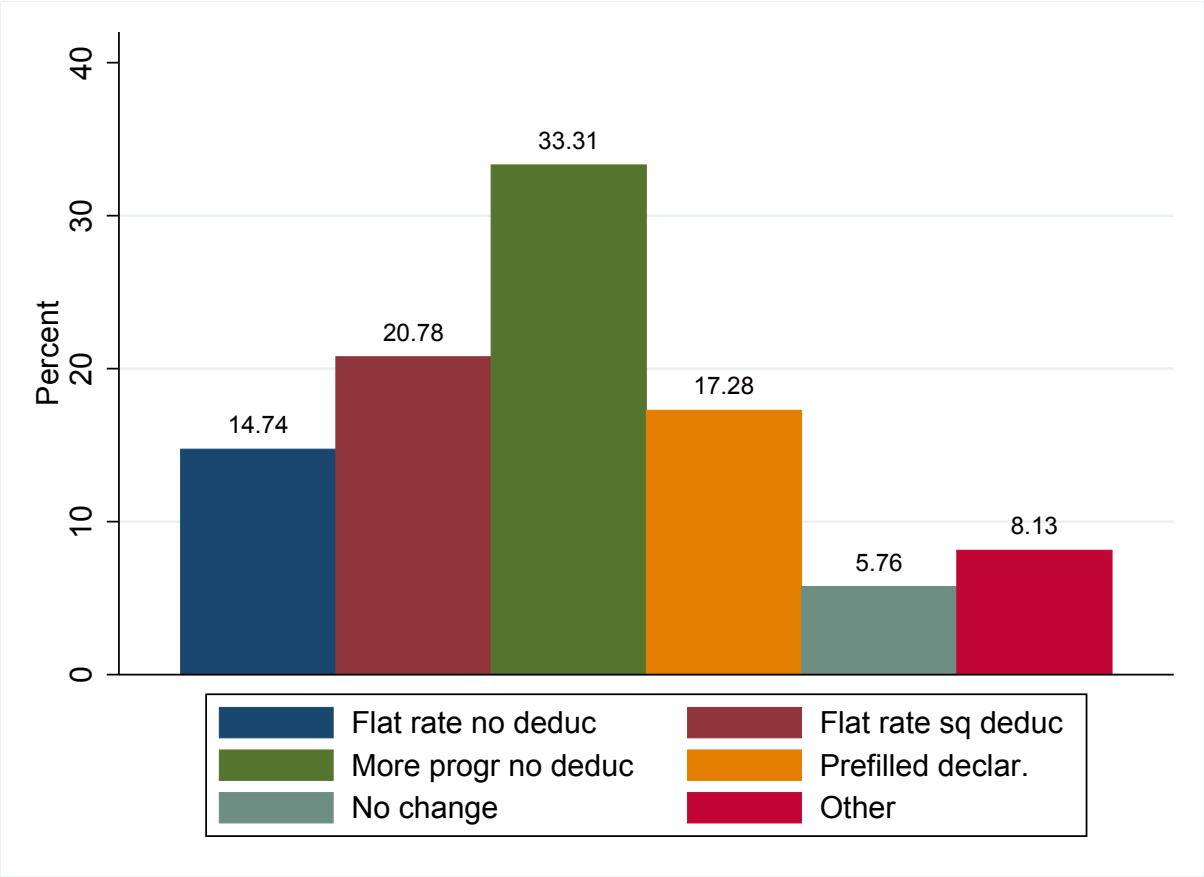
Notes: US Treasury estimates of tax expenditures, 1986-2013, adjusted for inflation to 2013 dollars. Source: Tax Foundation, Fiscal Fact, A Brief History of Tax Expenditures. Available online: `https://files.taxfoundation.org/legacy/docs/ff391.pdf`.

Figure 2: Preferences for Tax Simplification



Notes: This figure depicts the percentage share of respondents in the respective categories of the question "Do you generally think that the income tax system in Germany needs to be simplified?" Respondents could pick one of the following categories: 1 Absolutely not; ... ; 6 Absolutely; I do not know. The figure is based on 2,423 non missing observations. The mean answer is 5.16. Source: Own calculations based on German Internet Panel.

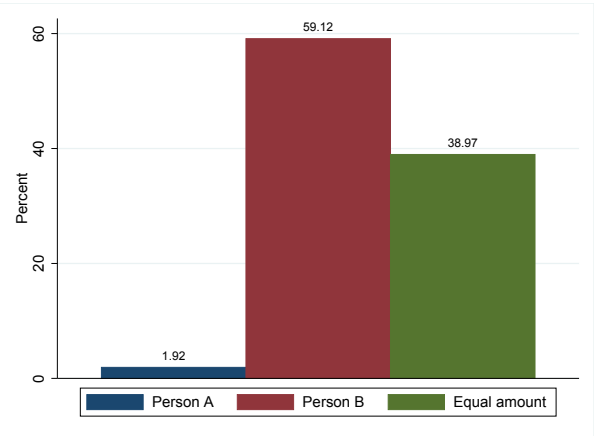
Figure 3: Which Revenue-Neutral Reform?



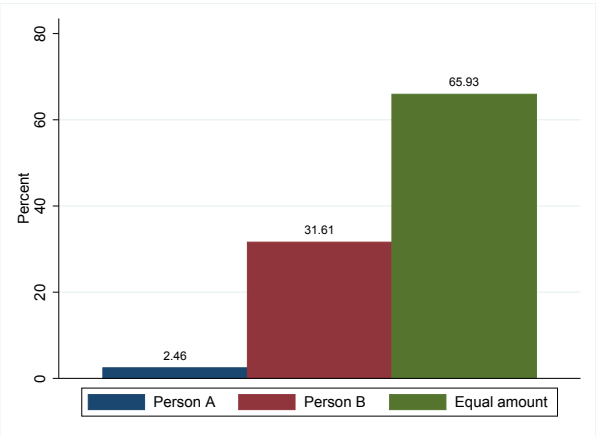
Notes: This figure depicts the percentage share of respondents in the respective categories of the question: "Which of the following measures to simplify the income tax system would you like the most? Assume the proposed measures will lead to unchanged tax revenues in each case." Respondents could pick 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 do not know. The figure is based on 1,771 non missing observations. Source: Own calculations based on German Internet Panel.

Figure 4: Who should pay more taxes?

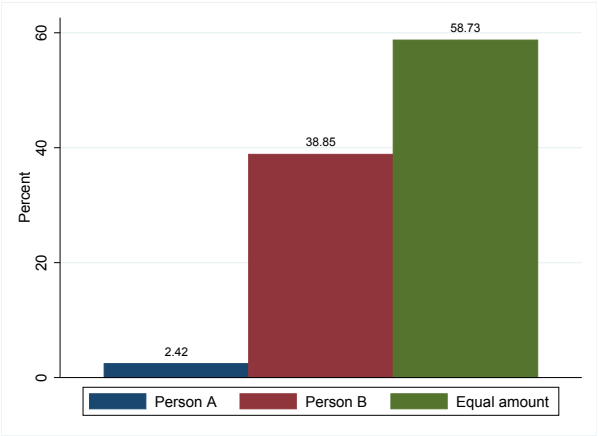
(a) A. Elderly Care



(b) B. Donations

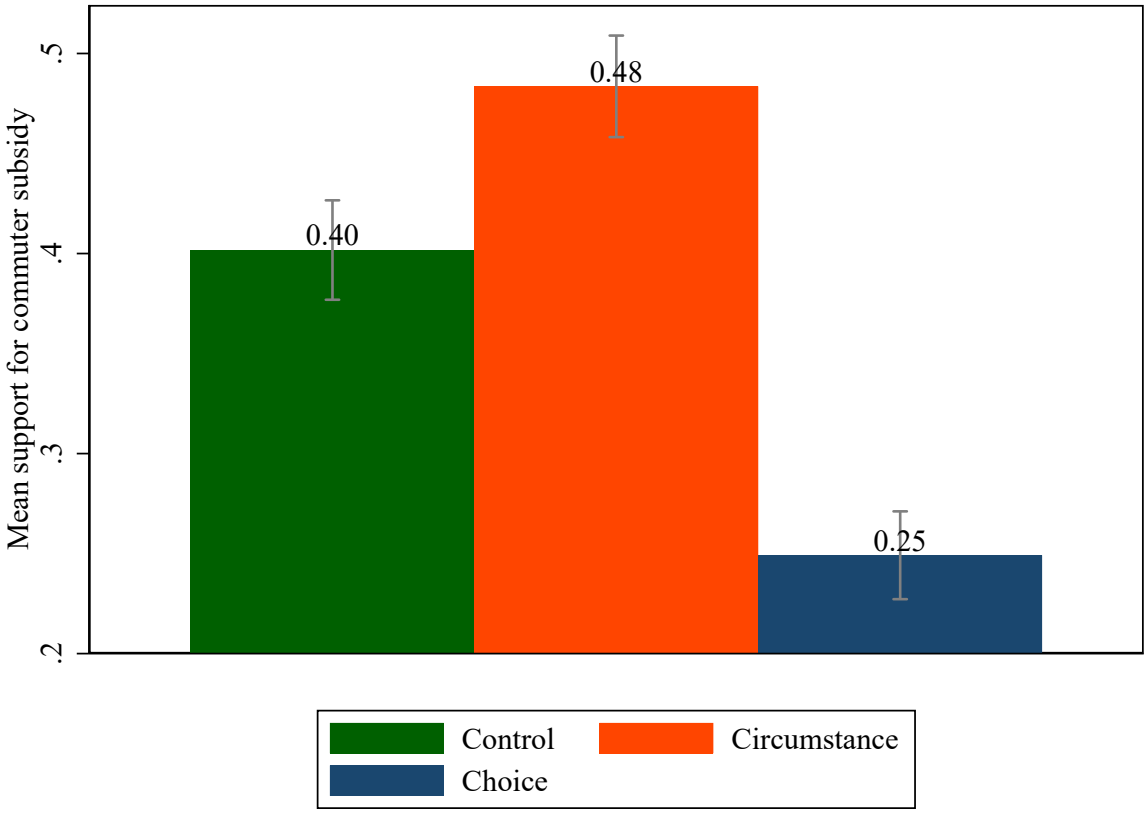


(c) C. Commuting To Work



Notes: This figure depicts the percentage share of respondents in the respective categories of the question for panel 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 B have the same gross income and are very similar in all other respects." Panel B: "Person A spends a considerable amount of her income on charitable giving. Person B does no such thing. Both Person A and B have the same gross income and are very similar in all other respects." Panel C: "Person A has to travel a considerable distance to work. Person B lives very close to work. Both Person A and B have the same gross income and are very similar in all other respects." Respondents could pick one of the following categories (order of answer categories was randomized): Person A should pay higher taxes; Person B should pay higher taxes; Person A and B should pay equal tax amounts. For Panel A the figure is based on 2,397 non missing observations, panel B is based on 2,398 non missing observations and panel C is based on 2394 non missing observations. Source: Own calculations based on German Internet Panel.

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 which 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 Section 5.1 about preferences about different tax burdens for taxpayers in alternate living situations. Specifically, we augment Q8 from wave 36 of the GIP in which our main survey is implemented by two treatment conditions which 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 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 to travel a considerable distance to work ever since, while Person B lives very close to work. Both Person A and B have the same gross income and are very similar in all other respects.". The *Choice* treatment group receives the following statement: "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 for a job with a very long way to work, while Person B lives very close to work. Both Person A and 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 B should pay equal tax amounts. The figure now 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 are displayed with 95% confidence bars. Total number of observations is 4,480 with 1,496, 1,491 and 1,493 observations for the control, *Circumstance* and *Choice* group, respectively. Source: Own calculations based on German Internet Panel.

Table 1: Exp 1: Effect on Preferences for Tax Simplification

	(1)	(2)	(3)	(4)	(5)
Experimental Group Reference category: Control					
Redistribution	-0.115**	-0.123**	-0.133**	-0.134**	-0.133**
	(0.058)	(0.058)	(0.055)	(0.055)	(0.055)
Avoidance	-0.042	-0.039	-0.032	-0.032	-0.029
	(0.057)	(0.056)	(0.053)	(0.053)	(0.053)
Constant	5.215***	4.610***	3.453***	3.469***	3.456***
	(0.040)	(0.149)	(0.242)	(0.248)	(0.249)
N	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. 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 deduction possibilities and allowances has a social-policy and redistributive compensation 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 deduction possibilities and allowances offers more scope for tax avoidance and tax adjustment. For example, tax deductions can be used to reduce one's own tax burden through better knowledge of the tax system or through unjustified specifications in the tax return." 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): (2) plus perceived difficulty to declare taxes, (4): (3) plus net household income, (5): (4) plus political preferences. Robust The scale of the outcome variable is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Exp 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)
N	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. 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 experiment. 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. 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 against 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 " Columns (1)-(5) all include a full set of interactions of the treatment groups of the first and second experiment, 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): (2) plus perceived difficulty to declare taxes, (4): (3) plus net household income, (5): (4) plus political preferences. The scale of the outcome variable is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Support for tax simplification: Difference between Q9 and Q2, by treatment combination

Exp 2 Exp1:	Control	Redistribution	Avoidance
Control	-0.193***	-0.114**	-0.136**
	(0.055)	(0.051)	(0.059)
N	218	246	242
Economic efficiency	-0.306***	-0.248***	-0.264***
	(0.063)	(0.051)	(0.058)
N	219	214	250
Special interest	-0.188***	-0.081*	-0.224***
	(0.054)	(0.049)	(0.060)
N	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 (s.e. = 0.019, N = 2108). We only include respondents who have given a preference in Q2 and Q9. Preference for tax simplification is measured on a 6 point scale based on question 2 "Do you generally think that the income tax system in Germany needs to be simplified?" and question 9: " 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 of the outcome variables is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Effects of tax information on non-response to open-text question in follow-up

Experimental Group Reference category: Control	
Redistribution	-0.034*
	(0.019)
Avoidance	0.011
	(0.020)
Economic Efficiency	-0.040**
	(0.020)
Special Interest	-0.032
	(0.020)
Constant	0.201***
	(0.018)
test	0.104
N	2160
r2	0.005
F	2.584
p	0.035

Notes: The table presents the effects of the randomized treatment interventions of the first and second experiment in wave 36 in July 2018 on the likelihood of not-answering an open-ended question on the main considerations of tax deductions and exemptions in the follow-up wave 46 (March 2020) of the GIP. Specifically, the question read as "*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?*". This exercise analyzes the persistence of information effects from our randomized information provision 20 months after the main intervention. The effects are estimated by OLS regressions of a dummy of not answering the open-ended question on treatment dummies. The experimental groups are: Control group, Redistribution, Avoidance, Economic efficiency group and Special interest group. Control is omitted, implying that the effects are relative to subjects which were part of the Control Group in both Experiment 1 and 2. The line test presents the p-value from F-test which compares whether the regression coefficients for the individual treatment groups are different from each other. The F-value and corresponding p-value at the bottom of the table analyze whether the treatment dummies can independently predict the outcome of non-response to the open-ended question. Effects are estimated without the inclusion of covariates. Altogether, 2160 out of 4692 participants in wave 46 already participated in wave 36 (i.e. 46.04%) and participation in wave 46 is not predicted by treatment status. Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Appendix

Table of Contents

A Summary Statistics and Balancedness Across Experimental Groups

A.1 Summary Statistics

Table A.1 provides descriptive statistics with respect to the demographics of our survey participants.[28] Most demographics in Table A.1 were not surveyed in the context of our specific survey block on tax simplification, but in other parts of the same wave or in other waves (some variables that do not change over time are linked to the current survey wave through the panel character of the GIP). The descriptive results for the questions of our survey block are not in this table, but are instead presented further below in the results section 4.

The table shows that we have a survey sample with balanced gender composition (48% female) and that we cover all age categories (with 36% of the participants being older than 58, and 23% retired). 61% of the respondents are married. 17% of the sample participants live in single households, 46% in 2-person households and 18% in households with three people. The distribution of education levels is also very reasonable. We split participants in different income categories and see that 11% are quite poor (net monthly income less than 1500 EUR) and 15% are relatively rich (net monthly income greater than 4500 EUR). The share of people in the three income classes in between poor and rich are quite balanced.

Corresponding with low current unemployment rates in Germany, only about 2% of the survey participants are unemployed. In terms of political affiliations, we see that about 38% of the sample are in the rather conservative political spectrum and 47% are rather left-wing. 8% indicate that they do not have any partisan preferences (left-right preferences are elicited on a 11-point scale from right to left, where we classify 'conservative' as <= 5 on this scale).

[28] Note that the GIP is designed to be representative of the German population (see Blom et al. 2015 for more details on the GIP's representativeness).

Table A.1: Descriptive Statistics

	N	Mean	Std.Dev.	Min	Max
Experiment 1					
Redistribution	2424	0.33	0.47	0	1
Avoidance	2424	0.33	0.47	0	1
Control	2424	0.33	0.47	0	1
Experiment 2					
Efficiency	2419	0.33	0.47	0	1
Special interest	2419	0.33	0.47	0	1
Control	2419	0.33	0.47	0	1
Demographics					
Single households	2463	0.17	0.38	0	1
2	2463	0.46	0.50	0	1
3	2463	0.18	0.38	0	1
4	2463	0.14	0.35	0	1
5+	2463	0.05	0.22	0	1
Age <=28	2461	0.09	0.29	0	1
Age 29-38	2461	0.15	0.36	0	1
Age 39-48	2461	0.15	0.36	0	1
Age 49-58	2461	0.24	0.43	0	1
Age >=59	2461	0.36	0.48	0	1
Married	2464	0.61	0.49	0	1
Female	2463	0.48	0.50	0	1
Unemployed	2463	0.02	0.13	0	1
Retired	2463	0.23	0.42	0	1
Low education	2401	0.03	0.17	0	1
Low-med education	2401	0.45	0.50	0	1
High-med education	2401	0.23	0.42	0	1
High education	2401	0.30	0.46	0	1
Difficulty in declaring taxes					
No difficulty	2381	0.03	0.17	0	1
2	2381	0.11	0.32	0	1
3	2381	0.18	0.38	0	1
4	2381	0.20	0.40	0	1

Very difficult	2381	0.12	0.32	0	1
No taxes declared	2381	0.09	0.28	0	1
Not self declared	2381	0.27	0.45	0	1
Household net income					
Poor	2464	0.11	0.32	0	1
2	2464	0.19	0.39	0	1
3	2464	0.20	0.40	0	1
4	2464	0.16	0.37	0	1
Rich	2464	0.15	0.36	0	1
No income stated	2464	0.11	0.32	0	1
Political orientation					
Conservatives	2464	0.38	0.48	0	1
Left-wing	2464	0.47	0.50	0	1
Non partisans	2464	0.08	0.27	0	1

Notes: The table depicts the summary statistics for all treatment group dummies and all variables in our tailored survey block on tax complexity. Variables are defined as follows: experiment 1 and experiment 2 realizations represent the respective group allocations of respondents in either experiment; household size comprises single households and household with 2, 3, 4 and 5+ members; age categories are $\leq$28, 29-38, 39-48, 49-55 and $\geq$59; Married equals 1 if respondent is married, 0 otherwise; Female equals 1 if respondent is female, 0 otherwise; Unemployed equals 1 if respondent is unemployed, 0 otherwise; Retired equals 1 if respondent is retired, 0 otherwise; education categories comprise low (secondary schooling, no job training), low to medium education (upper secondary schooling or finished job training), high to medium education (upper secondary schooling and finished job training) and high eduction (tertiary education); household income variables define net monthly household incomes on a 5-point scale from poor, i.e. 1 ($\leq$1500 Euro), 2 (1500$\geq x<$2500 Euro), 3 (2500$\geq x<$3500 Euro), 4 (3500$\geq x<$4500 Euro) to 5 being rich ($\geq$4500 Euro) as well as a dummy for no answers (No income stated) and a dummy for those observations which had not been in the GIP wave where the income question was asked; conservatives equals $\leq$5 on a 11-scale left-right placement variable, for $>$ 5 left-wing equals 1. Non partisans did not report a score for the left-rich placement variable and a dummy for those observations which had not been in the GIP Wave where the political preference question was asked (omitted from this table). Data comes from German Internet Panel (GIP) wave 36, except for the following items: political preferences derived from variable left-right placement (wave 31) as well as household incomes (wave 31).

A.2 Balancedness Across Experimental Groups

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 OLS regressions of the form $y_i = \beta Covariate_i + \epsilon_i$, 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. As a result of 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 predict treatment status. Out of 60 estimated coefficients, only 5 are significant at the 10% level and only one is statistically significant at the 5% and 1% level, respectively. This is well in line with these coefficients being significant by chance within their margin of error.[29]

The equivalent strategy was used for testing balancedness of the second experiment; see 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 5.2 for a more detailed explanation). The results are again quite reassuring. Out of 60 coefficients, 7 are significant at conventional levels of significane (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. Further below in our regressions, we present specifications that condition on all observable covariates to mitigate all remaining concerns regarding balancedness.

[29] With 60 estimated coefficients, one would expect six coefficients with a significance level of 10% even in the absence of any systematic difefrences between groups.

Table A.2: Balancing Tests First Experiment

Variable	Redistribution	Avoidance	Control
Gender: Reference category Male			
Sex	0.022	-0.029	0.002
	(0.019)	(0.019)	(0.019)
Marital status: Reference category: Not married			
Married	0.013	0.029	-0.036*
	(0.019)	(0.019)	(0.019)
Unemployed: Reference category: Employed			
Unemployed	-0.011	-0.061	0.063
	(0.074)	(0.074)	(0.074)
Retirement status: Reference category: Not retired			
Retired	0.042*	0.016	-0.042*
	(0.023)	(0.023)	(0.023)
Household size:			
HH-size	-0.007	0.012	-0.008
	(0.009)	(0.009)	(0.009)
Education:			
Education	-0.010	-0.011	0.002
	(0.011)	(0.011)	(0.011)
Household net income: Reference category poor			
2	-0.006	0.031	-0.016
	(0.036)	(0.035)	(0.036)
3	-0.030	0.065*	-0.032
	(0.036)	(0.035)	(0.035)
4	-0.032	0.047	-0.010
	(0.037)	(0.037)	(0.037)
Rich	-0.020	0.041	-0.014
	(0.038)	(0.037)	(0.037)
No income stated	-0.033	0.053	-0.001
	(0.040)	(0.040)	(0.040)
Age category:			
Age	0.017**	-0.004	-0.007

	(0.007)	(0.007)	(0.007)
Political orientation: Reference category: Left-wing			
Conservative	-0.022	-0.056***	-0.033
	(0.021)	(0.021)	(0.021)
Non partisans	-0.051	-0.003	0.050
	(0.037)	(0.037)	(0.037)
Difficulty in declaring taxes: Reference category: No difficulty			
2	0.022	0.099	-0.022
	(0.063)	(0.063)	(0.063)
3	-0.066	0.116*	-0.050
	(0.060)	(0.060)	(0.060)
4	-0.059	0.064	-0.005
	(0.060)	(0.060)	(0.060)
Very difficult	-0.012	0.020	-0.009
	(0.063)	(0.063)	(0.063)
No taxes declared	-0.055	0.086	-0.031
	(0.065)	(0.065)	(0.065)
Taxes not self declared	-0.030	0.089	-0.059
	(0.059)	(0.059)	(0.059)

Notes: Randomization checks for the first experiment. The table shows the coefficients and standard errors (in parentheses) from a series of regressions of the form: $y_i = \beta Covariate_i + \epsilon_i$. Where $Covariate_i$ is the respective covariate listed above. In Column (1) y_i equals 1 if participant i is in the redistribution group and 0 otherwise. In Column (2),y_i equals 1 if participant i is in the avoidance group and 0 otherwise. In Column (3), y_i equals 1 if participant i is in the control group and 0 otherwise. Standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Balancing Tests Second Experiment

Variable	Economic efficency	Special interest	Control
Gender. Reference category Male			
Sex	-0.025	0.021	-0.001
	(0.033)	(0.034)	(0.033)
Marital status: Reference category: Not married			
Married	-0.038	0.032	0.012
	(0.034)	(0.034)	(0.033)
Unemployed: Reference category: Employed			
Unemployed	0.302**	-0.165	-0.134
	(0.118)	(0.120)	(0.118)
Retirement status: Reference category: Not retired			
Retired	0.031	-0.002	-0.026
	(0.041)	(0.042)	(0.041)
Household size:			
HH-size	-0.01	0.009	-0.002
	(0.016)	(0.016)	(0.015)
Education:			
Education	-0.018	0.011	0.009
	(0.019)	(0.019)	(0.019)
Household net income: Reference category poor			
2	-0.069	0.027	0.043
	(0.061)	(0.062)	(0.062)
3	-0.012	-0.008	0.020
	(0.061)	(0.062)	(0.062)
4	-0.138**	0.063	0.075
	(0.063)	(0.064)	(0.063)
Rich	-0.092	-0.004	0.088
	(0.064)	(0.065)	(0.065)
No income stated	-0.014	0.090	-0.086
	(0.068)	(0.069)	(0.067)
Age category:			
Age	-0.004	0.014	-0.008

	(0.012)	(0.013)	(0.012)
Political orientation: Reference category: Left-wing			
Conservative	-0.007	-0.019	0.026
	(0.036)	(0.037)	(0.036)
Non partisans	0.141**	-0.160***	0.022
	(0.059)	(0.060)	(0.059)
Difficulty in declaring taxes: Reference category: No difficulty			
2	-0.025	-0.009	-0.125
	(0.104)	(0.106)	(0.103)
3	0.117	0.021	-0.145
	(0.101)	(0.102)	(0.100)
4	0.135	0.050	-0.191*
	(0.099)	(0.101)	(0.098)
Very difficult	0.206**	-0.091	-0.115
	(0.104)	(0.106)	(0.103)
No taxes declared	0.175	0.021	-0.197*
	(0.108)	(0.110)	(0.107)
Taxes not self declared	0.116	0.012	-0.128
	(0.098)	(0.100)	(0.097)

Notes: Randomization checks for the second experiment. The table shows the coefficients and standard errors (in parentheses) from a series of regressions of the form: $y_i = \beta Covariate_i + \epsilon_i$. Where $Covariate_i$ is the respective covariate listed above. Sample restricted to those participants who were in the control group in the first experiment. In Column (1) y_i equals 1 if participant i is in the economic efficency group and 0 otherwise. In Column (2),y_i equals 1 if participant i is in the special interest group and 0 otherwise. In Column (3), y_i equals 1 if participant i is in the control group and 0 otherwise. Standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

B Additional Results for Understanding Attitudes towards Tax Complexity

This Appendix presents more detailed results and discussions of further survey questions, most of which are briefly summarized in Section 4 of the main text.

Difficulty of Filing a Return. We survey the perceived difficulty of filing a tax return on a 5-point scale from 1 'Very Easy' to 5 'Very Difficult' (Q1). This question particularly allows us to investigate if the strong preferences for tax simplification correspond with the perceived difficulty to file a return. Recall that this question was asked before respondents were treated. It is therefore not important to report the control-group stats separately. The results for this survey question are summarized in Figure B.1, which presents the share of respondents in each reply category.

The right part of the Figure shows that more than 1/3 of the respondents did not indicate their perceived difficulty, either because they do not file a return or because they have their return filed by someone else (e.g., a tax preparer, spouse). Among all respondents who file a return, about 50% find it difficult to file their return (either category 5 or 4). The medium category 3 was checked by about 28%, and 22% of the respondents find it rather easy to file the return. The mean reply for this question is 3.41 (on a scale of 1-5) among all respondents who file a tax return.

General Need for Tax Reforms. We further surveyed if participants believe that the German tax system generally needs to be reformed (Q3). The question was asked on a scale from 1 'Absolutely not' to 6 'Absolutely'. Figure B.2 shows that a large fraction of 47% (control group: 46.7%) of question respondents (i.e., those who gave a non-missing and non-*I-don't-know* reply) think that the tax system 'absolutely' (reply category 6) needs to be reformed. Another 25.2% (control group: 27.7%) have a fairly strong preference for reforming the system (reply category 5). A negligible share of people do not see a need to implement reforms: only about 9.6% (control group: 8%) of the respondents checked reply categories 1, 2 or 3. The mean reply for this question is 5.06 (control group: 5.09).

Distributional Implications of Tax Expenditures. One frequently raised concern in the context of tax complexity is that the rich are able to exploit tax expenditures better than low-income taxpayers; for example, because they afford professional tax advisors or because they have income sources with more possibilities for tax planning. However, given that many tax expenditures also have a redistributive purpose, it is interesting to survey the public opinion in this context. We therefore survey beliefs about the distributional implications of tax expenditures. In particular, we ask if deductions and allowances contribute to equality or if high-income taxpayers tend to benefit from them (Q4). The reply categories are 1 'Equality' to 6 'High Incomes Benefit' and the question results are summarized in Figure B.3.

The majority of question respondents (i.e., those who gave a non-missing and non-*I-don't-know* reply) believe that allowances and deductions benefit high-income taxpayers, rather than

contributing to equality.[30] Only about 29% (control group: 30%) of all respondents checked reply categories 1, 2 or 3. The mean reply for this question is 4.33 (control group: 4.28).

Which Tax Expenditures are Used? We also survey which type of tax expenditures respondents use regularly (Q10). This survey question mainly serves the purpose of evaluating if survey answers about particular type of expenditures (see below section 5.1) are driven by self-interest. The question, however, is also interesting in itself and we therefore briefly summarize the results in Figure B.4. The Figure presents the share of people who use particular tax expenditures (note that multiple answers were possible so the shares do not add up to 100). The list of itemized deductions is, of course, not exhaustive. The most frequently used tax expenditures in our sample are the commuting-to-work allowance, the deduction of other type of work expenses (e.g., work-related costs for books, clothes, etc.), charitable donations, and deductible expenses for pension and retirement savings. Child allowances and so-called 'standard deductions'.[31] are also quite frequently used.

Anatomy of Preferences for Tax Simplification. We also study the 'anatomy' of simplification attitudes and investigate which groups are more likely to support tax simplification. For this purpose, we simply regress (using OLS) our measure of simplification preferences on a wide set of observable characteristics (and the experimental-group status). These characteristics comprise demographic factors, a measure of perceived tax difficulty, household income and political preferences. We report results with robust standard errors. The coefficients in this regression are (conditional) correlations and should not be given a causal interpretation.

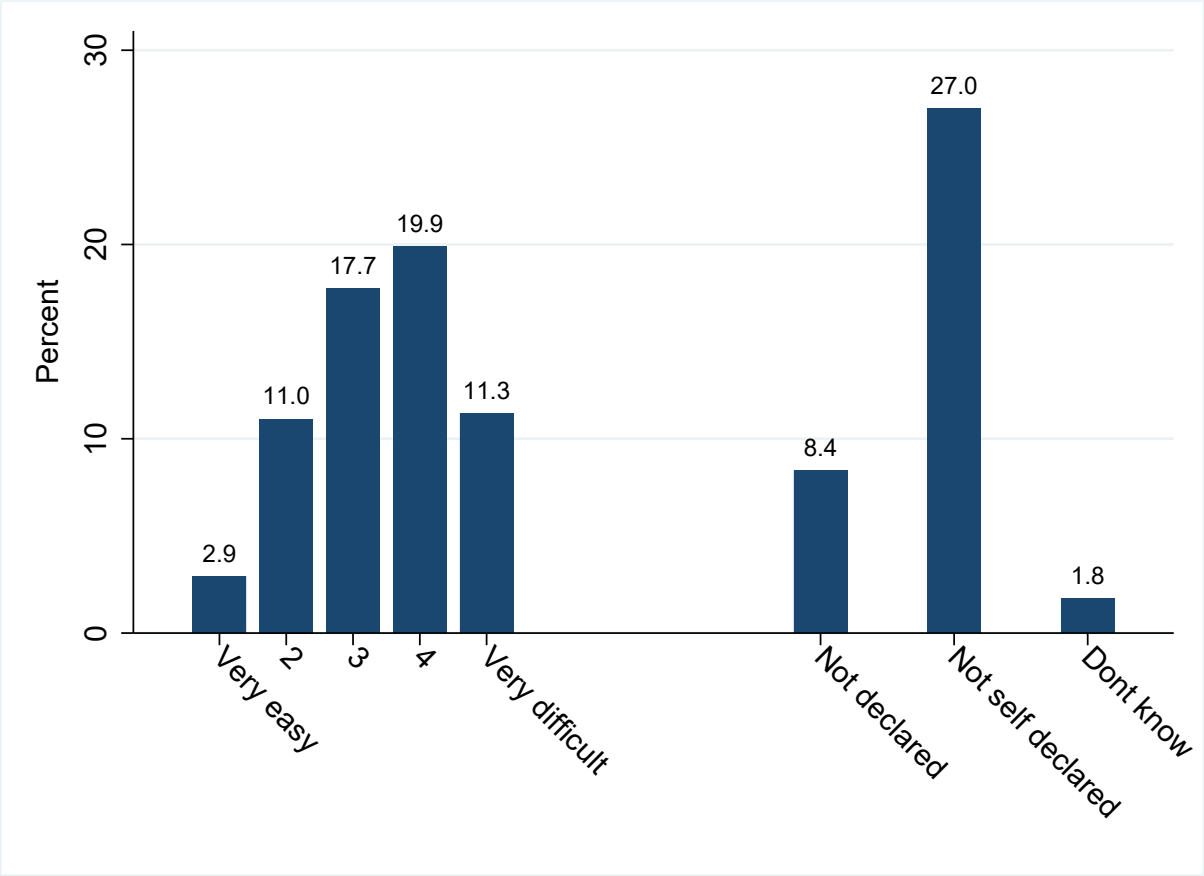
The results for this anatomy analysis are presented in Table B.1 below. Important demographic correlates of simplification preferences are age and gender. Older people tend to have stronger preferences for simplification, and women have weaker preferences. Another important correlate of simplification preferences is the perceived difficulty of filing a tax return. Respondents who find it easy to file a tax return have lower simplification preferences than respondents who find it difficult.[32] 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.

[30] 33.1% were in the corner category 6 (control group: 32.8%) and another 20.4% (control group: 19%) are in the second-largest reply category 5.

[31] This 'standard deduction' represents the lump sum deduction amount for taxpayers who do not exceed the thresholds in other deduction categories.

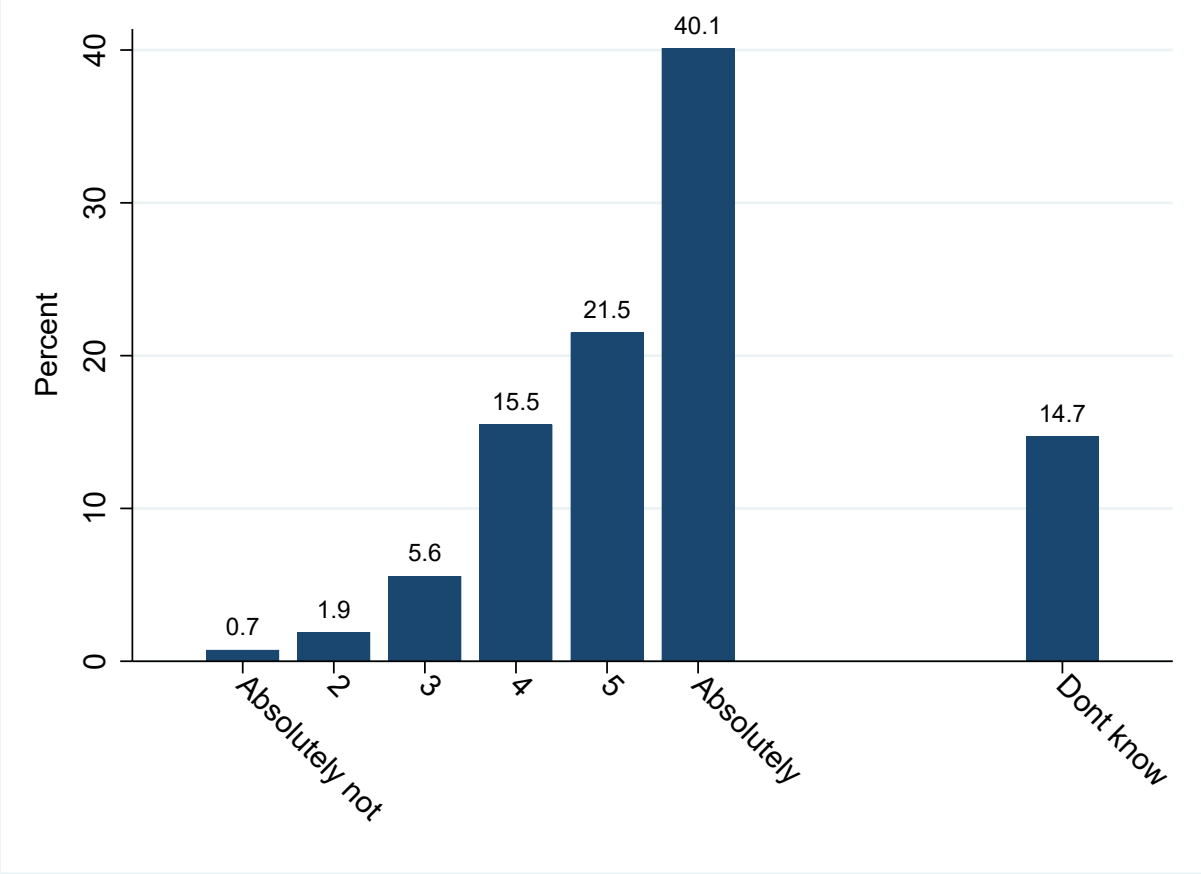
[32] Age and gender differences are further investigated in Figures B.6a and B.6b which illustrate unconditional differences between age groups and between men and women, respectively. Figure B.6a shows that the average support for simplification steadily increases in age; the support is 16% (control group: 12.3%) higher among respondents older than 58, relative to respondents younger than 29. As illustrated in Figure B.6b, the support for simplification among men is roughly 3% (control group: 4.5%) greater than among women. The unconditional relationship between simplification preferences and perceived filing difficulty are displayed in Figure B.6c. The Figure confirms the intuitive result that the perceived difficulty to declare income taxes is positively associated with support for tax simplification.

Figure B.1: Perceived Difficulty of Filing a Tax Return



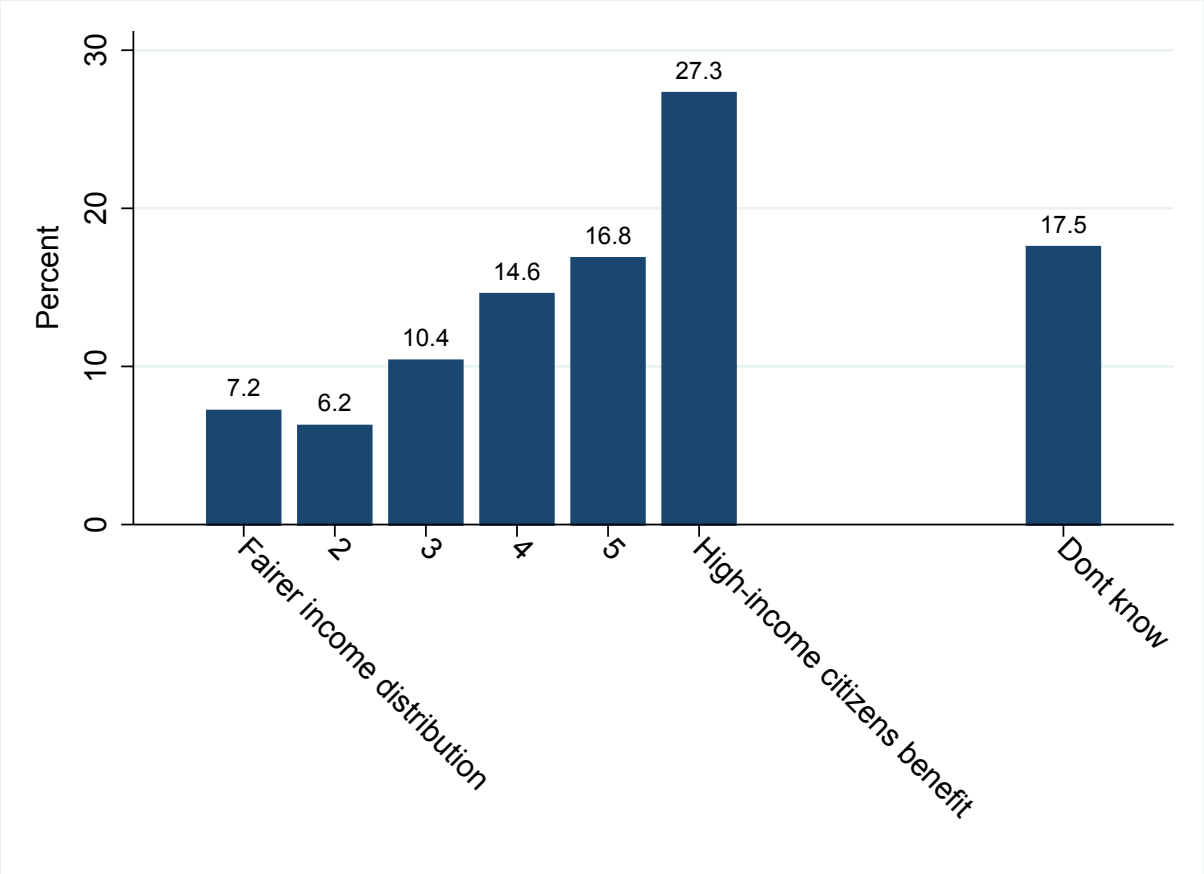
Notes: This figure depicts the percentage share of respondents in the respective categories of the question "How difficult is it for you to fill out your tax return?" Respondents could pick one of the following categories: 1 Very easy; ... ; 5 Very difficult; I do not know because no taxes are declared in my name; I do not know because I do not declare taxes myself (rather, my partner or a tax consultant, etc. does this); I do not know. The figure is based on 2,424 non missing observations. The mean answer for categories 1 to 5 is 3.41. Source: Own calculations based on German Internet Panel.

Figure B.2: Need for Tax Reform



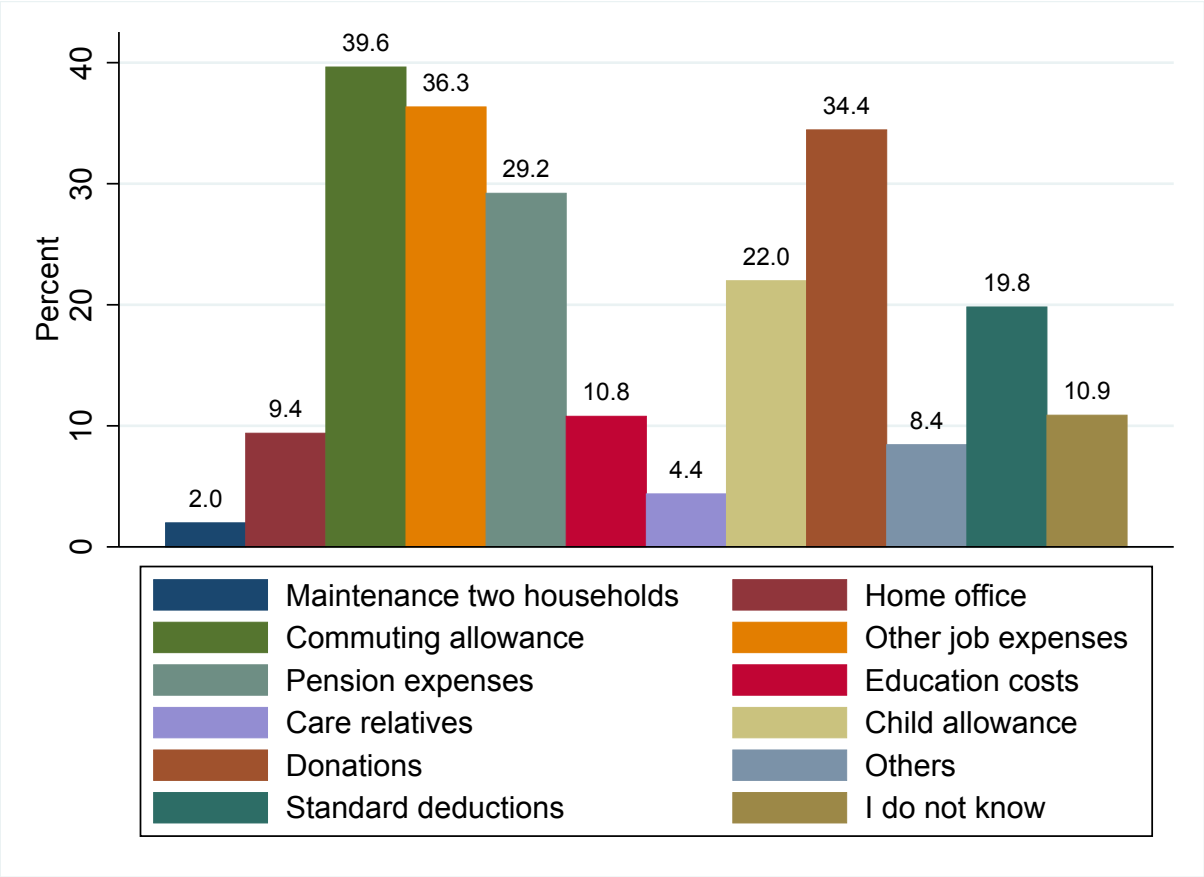
Notes: This figure depicts the percentage share of respondents in the respective categories of the question "Do you generally believe that the income tax system in Germany is in need of reform?" Respondents could pick one of the following categories: 1 Absolutely not; ... ; 6 Absolutely; I do not know. The figure is based on 2,423 non missing observations. The mean answer is 5.06. Source: Own calculations based on German Internet Panel.

Figure B.3: Perceived Distributional Implications of Complexity



Notes: This figure depicts the percentage share of respondents in the respective categories of the question "Do you think that numerous deductions and allowances contribute to a fairer distribution of income, or do you believe that high-income citizens benefit more from these deductions and allowances?" Respondents could pick one of the following categories: 1 They contribute to fairer income distribution; ... ; 6 High-income citizens benefit; I do not know. The figure is based on 2,423 non missing observations. The mean answer is 4.33. Source: Own calculations based on German Internet Panel.

Figure B.4: Which Deductions and Allowances are used?



Notes: This figure depicts the percentage share of respondents in the respective categories of the question: "Which of the following deductions and/or allowances do you usually use when filing your income tax?" Respondents could pick one of the following categories: Maintenance of two households; Home office; Commuting allowance; Other job related expenditures; Pension expenses; Education costs; Care relatives; Child allowance, childcare; Donations; Others [insert text]; No deductions; I do not know. The figure is based on 2,215 non missing observations. Note shares do not add up to one because respondents could check multiple items. Source: Own calculations based on German Internet Panel.

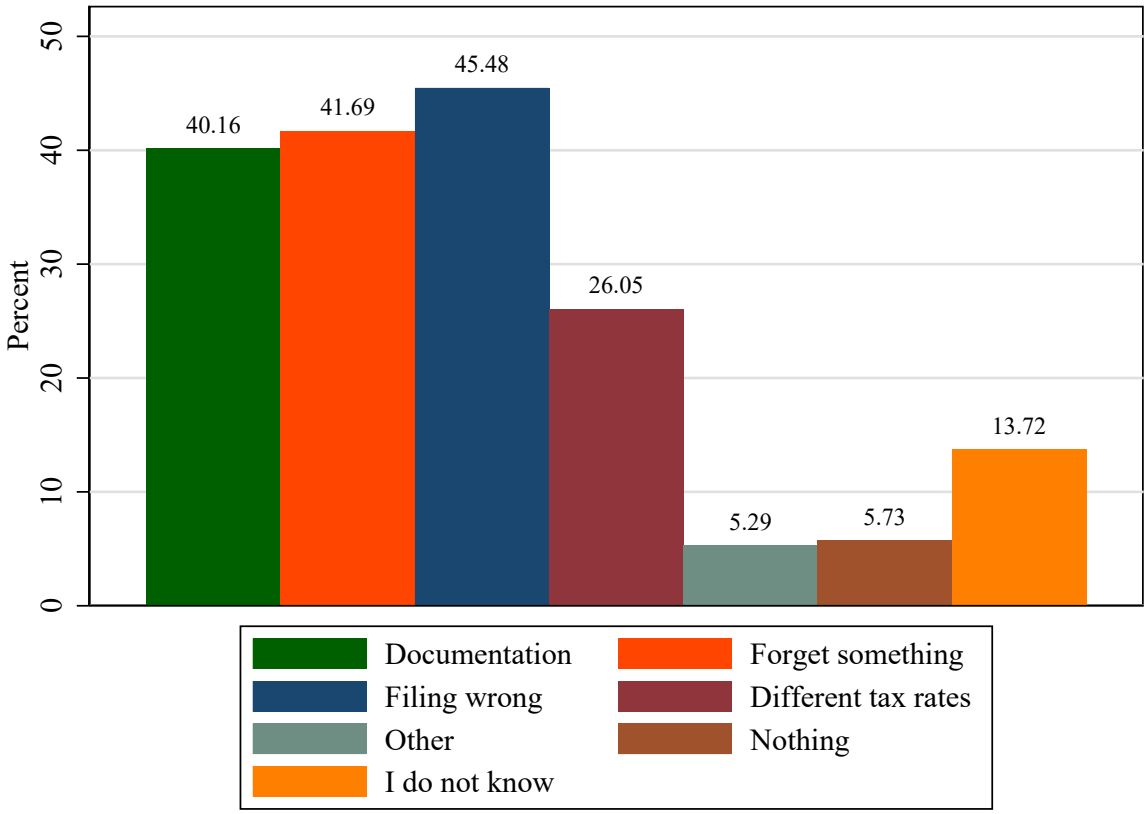
Table B.1: Anatomy of tax simplification preferences

	(1)	(2)	(3)	(4)
Household size	0.030	0.022	0.021	0.021
	(0.026)	(0.024)	(0.024)	(0.024)
Age	0.181***	0.170***	0.170***	0.169***
	(0.025)	(0.024)	(0.024)	(0.024)
Married	-0.053	-0.086	-0.091	-0.093
	(0.058)	(0.056)	(0.057)	(0.057)
Female	-0.126***	-0.133***	-0.132***	-0.138***
	(0.047)	(0.045)	(0.045)	(0.046)
Unemployed	0.142	-0.038	-0.031	-0.034
	(0.198)	(0.199)	(0.203)	(0.203)
Retired	0.002	0.028	0.030	0.030
	(0.064)	(0.063)	(0.063)	(0.063)

Education	-0.016	0.007	0.006	0.005
	(0.027)	(0.026)	(0.027)	(0.027)
Difficulty 2		0.684***	0.680***	0.675***
		(0.213)	(0.214)	(0.214)
Difficulty 3		1.029***	1.028***	1.028***
		(0.205)	(0.205)	(0.205)
Difficulty 4		1.370***	1.367***	1.364***
		(0.204)	(0.204)	(0.204)
Very difficult		1.724***	1.721***	1.720***
		(0.205)	(0.206)	(0.206)
No taxes declared		1.249***	1.250***	1.242***
		(0.216)	(0.216)	(0.216)
Taxes not self declared		1.207***	1.204***	1.203***
		(0.203)	(0.204)	(0.203)
Income gr 2			-0.034	-0.034
			(0.100)	(0.100)
Income gr 3			-0.022	-0.019
			(0.100)	(0.100)
Income gr 4			0.055	0.056
			(0.102)	(0.102)
Rich			-0.018	-0.014
			(0.110)	(0.110)
No income stated			0.033	0.032
			(0.108)	(0.109)
Conservative				0.039
				(0.048)
Non partisans				0.064
				(0.102)
Redistribution	-0.123**	-0.133**	-0.134**	-0.133**
	(0.058)	(0.055)	(0.055)	(0.055)
Avoidance	-0.039	-0.032	-0.032	-0.029
	(0.056)	(0.053)	(0.053)	(0.053)
Constant	4.610***	3.453***	3.469***	3.456***
	(0.149)	(0.242)	(0.248)	(0.249)
N	2132	2109	2109	2109
R2	0.048	0.149	0.150	0.150

Notes: The table presents the determinants of Preferences for Tax Simplicity using OLS regressions of preferences for tax simplicity on various covariates. Each column (1)-(5) presents the results of one regression with different sets of covariates. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Variables are defined as follows: experiment 1 and experiment 2 realizations represent the respective group allocations of respondents in either experiment; household size comprises single households and household with 2, 3, 4 and 5+ members; age categories are $\leq$28, 29-38, 39-48, 49-55 and $\geq$59; Married equals 1 if respondent is married, 0 otherwise; Female equals 1 if respondent is female, 0 otherwise; Unemployed equals 1 if respondent is unemployed, 0 otherwise; Retired equals 1 if respondent is retired, 0 otherwise; education categories comprise low (secondary schooling, no job training), low to medium education (upper secondary schooling or finished job training), high to medium education (upper secondary schooling and finished job training) and high eduction (tertiary education); household income variables define net monthly household incomes on a 5-point scale from poor, i.e. 1 ($\leq$1500 Euro), 2 (1500$\geq x<$2500 Euro), 3 (2500$\geq x<$3500 Euro), 4 (3500$\geq x<$4500 Euro) to 5 being rich ($\geq$4500 Euro) as well as a dummy for no answers (No income stated) and a dummy for those observations which had not been in the GIP Wave where the income question was asked; conservatives equals 1 if $\leq$5 on a 11-scale left-right placement variable, for $>$ 5 left-wing equals 1. Non partisans did not report a score for the left-right placement variable and a dummy for those observations which had not been in the GIP Wave where the political preference question was asked. Data comes from German Internet Panel (GIP) wave 36, except for the following items: political preferences derived from variable left-right placement (wave 31) as well as household incomes (wave 31).

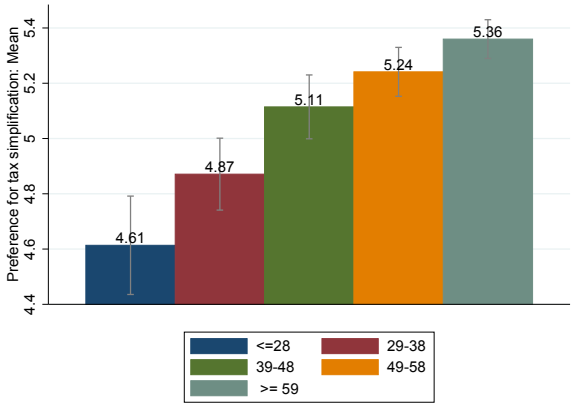
Figure B.5: Follow-up Survey: Which factors make the tax system complicated?



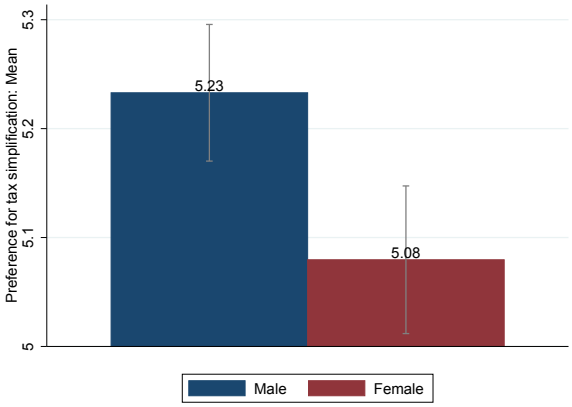
Notes: The figure depicts the frequencies for various factors which potentially make the German income tax complicated. These beliefs were elicited in a follow-up survey that was part of of GIP wave 46 in March 2020. The question reads: "In your opinion, which of the factors listed here makes income tax particularly complicated?". Respondents can choose multiple answers from the following answer categories: the scope of documentation requirements ("Documentation"), the fear of forgetting something important when filing your tax return ("Forget something"), the fear of filing something wrong in the tax return ("Filing wrong"), different tax rates in the income tax schedule ("Different tax rates"), other reasons ("Other reasons"), none of the options listed ("Nothing") and the category "I do not know". The figure uses 4,534 non missing observations. Source: Own calculations based on German Internet Panel.

Figure B.6: Preferences for Tax Simplification by groups

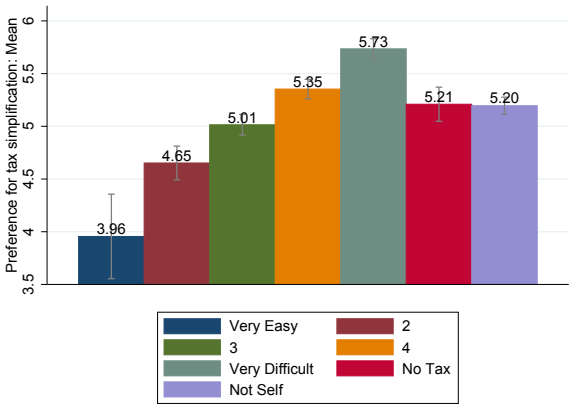
(a) A. Age Categories



(b) B. Gender



(c) C. Perceived Difficulty to File a Return



Notes: The outcome variable is the survey-based preference for tax simplification as described in Section 3. Panel A is based on 2,189 non missing observations, panel B is based on 2,190 non missing observations and panel C is based on 2,164. Source: Own calculations based on German Internet Panel.

C Additional Results Relating to the Randomized Survey Experiment

Table C.1: Exp 2: Effects of the Second Experiment by Treamtent Group Status in Exp 1.

	(1)	(2)	(3)	(4)	(5)
Experimental Group Reference category: Control in Experiment 2					
Effect of Exp 2 for Redistribution Group in Exp 1					
Economic Efficiency	-0.052	-0.066	-0.087	-0.091	-0.104
	(0.107)	(0.103)	(0.099)	(0.099)	(0.099)
Special Interest	-0.003	-0.044	-0.054	-0.063	-0.080
	(0.105)	(0.102)	(0.100)	(0.100)	(0.099)
Effect of Exp 2 for Avoidance Group in Exp 1					
Economic Efficiency	0.013	-0.003	-0.004	0.006	0.007
	(0.105)	(0.102)	(0.100)	(0.101)	(0.101)
Special Interest	0.179*	0.113	0.075	0.006	0.095
	(0.103)	(0.101)	(0.099)	(0.099)	(0.099)
Effect of Exp 2 for Control Group in Exp 1					
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)
N	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 by treatment group status in the first experiment. 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 experiment. 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. 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 against 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 " Columns (1)-(5) all include a full set of interactions of the treatment groups of the first and second experiment, 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): (2) plus perceived difficulty to declare taxes, (4): (3) plus net household income, (5): (4) plus political preferences. The scale of the outcome variable is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.2: Exp 1: Effect on Perceived Distributional Effects of Complexity

	(1)	(2)	(3)	(4)	(5)
Experimental Group Reference category: Control					
Redistribution	0.009	0.005	0.001	0.009	0.007
	(0.090)	(0.090)	(0.090)	(0.090)	(0.089)
Avoidance	0.118	0.137	0.157*	0.158*	0.169*
	(0.089)	(0.090)	(0.091)	(0.091)	(0.090)
Constant	4.285***	3.506***	3.136***	3.127***	3.126***
	(0.065)	(0.215)	(0.295)	(0.315)	(0.314)
N	1998	1946	1931	1931	1931
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 believes about whether people think that deductions work in favor of the rich. This is estimated by OLS regressions of beliefs on treatment dummies. The outcome is measured on a 6 point scale based on the question: "Do you think that numerous deductions and allowances contribute to a fairer distribution of income, or do you believe that high-income citizens benefit more from these deductions and allowances?" 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 possible deductions 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-policy role, particularly in relation to income redistribution. 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 . 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): (2) plus perceived difficulty to declare taxes, (4): (3) plus net household income, (5): (4) plus political preferences. The scale of the outcome variable is 1 (add to a fair income distribution) to 6 (higher incomes benefit). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.3: Exp 1: Effect on Further Survey Variables

	Q3	Q5	Q6	Q7	Q8
Redistribution	-0.021	-0.016	-0.034	-0.035	-0.032
	(0.060)	(0.029)	(0.024)	(0.024)	(0.025)
Avoidance	-0.074	-0.012	-0.025	-0.006	0.032
	(0.060)	(0.029)	(0.024)	(0.024)	(0.025)
Constant	5.088***	0.448***	0.409***	0.673***	0.588***
	(0.043)	(0.021)	(0.017)	(0.017)	(0.017)
N	2067	1771	2397	2398	2394

Notes: The table presents the effects of the randomized treatment interventions on Q3:"Do you generally believe that the income tax system in Germany is in need of reform?" Q5: "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." Q6-8: "Imagine two persons, A and B. Which person do you think should pay more taxes in the following situation?" For a better readability we report for Q5 not a coefficient for a regression of treatment status on each single category (they are all insignificant). Instead we report regression results for a dummy indicating whether respondents did chose an option with status quo deductions. The experimental groups are: Control group, Redistribution group and Avoidance group. Control is omitted, implying that the effects are relative to the Control Group. No further control variables were included. 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 possible deductions 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-policy role, particularly in relation to income redistribution. 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 . 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." For detailed information on Q3,Q5 and Q6-8 see section G.1. Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.4: Exp 1: Effect on Preferences for Tax Simplification. Ordered Probit

	(1)	(2)	(3)	(4)	(5)
Experimental Group Reference category: Control					
Redistribution	-0.109*	-0.123**	-0.144**	-0.146**	-0.144**
	(0.060)	(0.061)	(0.062)	(0.062)	(0.062)
Avoidance	-0.041	-0.038	-0.024	-0.024	-0.019
	(0.059)	(0.060)	(0.061)	(0.061)	(0.061)
Constant	-0.120***	0.482***	1.537***	1.499***	1.516***
	(0.043)	(0.150)	(0.216)	(0.227)	(0.228)
N	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. Estimated by an ordered probit 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 deduction possibilities and allowances has a social-policy and redistributive compensation 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 deduction possibilities and allowances offers more scope for tax avoidance and tax adjustment. For example, tax deductions can be used to reduce one's own tax burden through better knowledge of the tax system or through unjustified specifications in the tax return." 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): (2) plus perceived difficulty to declare taxes, (4): (3) plus net household income, (5): (4) plus political preferences. Robust The scale of the outcome variable is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.5: Exp 1: Effect on Distributional Effects of Complexity. Ordered Probit

	(1)	(2)	(3)	(4)	(5)
Experimental Group Reference category: Control					
Redistribution	-0.004	-0.007	-0.010	-0.004	-0.004
	(0.058)	(0.059)	(0.059)	(0.059)	(0.060)
Avoidance	0.079	0.092	0.109*	0.110*	0.120**
	(0.059)	(0.060)	(0.061)	(0.061)	(0.061)
Constant	0.463***	0.976***	1.216***	1.229***	1.238***
	(0.046)	(0.142)	(0.195)	(0.211)	(0.211)
N	1998	1946	1931	1931	1931
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 beliefs about whether people think that deductions work in favor of the rich. Estimated by an ordered probit Regressions of believes on treatment dummies. The outcome is measured on a 6 point scale based on the question: 'Do you think that numerous deductions and allowances contribute to a fairer distribution of income, or do you believe that high-income citizens benefit more from these deductions and allowances?" 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 possible deductions 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-policy role, particularly in relation to income redistribution. 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 . 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): (2) plus perceived difficulty to declare taxes, (4): (3) plus net household income, (5): (4) plus political preferences. Robust The scale of the outcome variable 1 (add to a fair income distribution) to 6 (higher incomes benefit). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.6: Exp 2: Effect on Preferences for Tax Simplification. Ordered Probit

	(1)	(2)	(3)	(4)	(5)
Experimental Group Reference category: Control					
Economic Efficiency	-0.153	-0.182*	-0.210**	-0.220**	-0.224**
	(0.102)	(0.105)	(0.108)	(0.108)	(0.108)
Special Interest	-0.041	-0.037	-0.023	-0.038	-0.035
	(0.096)	(0.098)	(0.101)	(0.101)	(0.101)
Constant	0.048	0.527***	1.145***	1.066***	1.107***
	(0.068)	(0.155)	(0.205)	(0.218)	(0.219)
N	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. Estimated by an ordered probit regressions of preferences for tax simplification on treatment dummies and a full set of interactions of the treatment groups of the first and second experiment. 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. 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 against 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 " 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): (2) plus perceived difficulty to declare taxes, (4): (3) plus net household income, (5): (4) plus political preferences. The scale of the outcome variable is 1 (absolutely not) to 6 (absolutely). Robust standard errors are in parentheses *** p<0.01, ** p<0.05, * p<0.1.

D Randomization Tests and Multiple Hypothesis Testing.

For both survey experiments, 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). Note that the coefficients are not affected by the alternative types of statistical inference that we present in the following. We present and discuss the results for both of the experiments in the following.

Survey Experiment 1. First, we perform randomization tests following Young (2018). The Young (2018)-procedure performs exact tests which test the sharp null hypothesis that the effect of the (randomly provided) arguments provided in the treatment is zero for all individuals receiving our treatment. That is, it does not test whether the average treatment effect is zero (which is what we tested in our main analysis), but whether the treatment effects are zero across all repondents. The randomization-test procedure, which is in the spirit of Fisher (1935), is more conservative in computing standard errors: Young (2018) reports that, using his approach, the number of significant results of randomized experiments is considerably reduced relative to conventional tests of individual treatment effects. Compared to classical asymptotic-based testing procedures, these randomization tests have the advantage that they are robust against concentrated leverage and do not rely on sample size or the characteristics of the error (Young 2018).[33]

Using the Young (2018)-procedure with 5000 draws (to approximate the p-value of the Fisher distribution), the effect of the *Redistribution*-treatment on preferences for simplification yields a p-value of 0.015 (in our preferred specification with the full set of controls). The coefficients are thus statistically significant and the levels of significance of the classical testing method – as reported above – are confirmed. This stricter procedure for computing p-values also confirms the insignificant effect of our *Avoidance*-treatment on preferences for tax simplicity; the p-value for the *Avoidance*-dummy is computed to be 0.6 and thus far off conventional levels of statistical significance.

Overall, all p-values based on the randomization tests are very similar to the ones obtained by ordinary OLS with robust standard errors. This is reassuring and lends credibility to the inference used in our main analyses above (which used classical hypothesis testing). The similarity between p-values might be interpreted as an indication that the treatment effects in our setting are constant among individuals; as noted by Ding (2017), the sharp-null hypothesis and the null hypothesis of zero average causal effect are equivalent in the case of constant causal effects.

Second, we use the method proposed by Westfall and Young (1993), and for example recently applied by Blattman et al. (2017), to adjust standard errors for multiple comparisons. As Blattman et al. (2017), we take a rather conservative approach that adjusts for comparisons

[33] We implement the randomization tests using the ado file provided by Alwyn Young on his website; the exact testing procedure is described in (Young 2018). We report randomization-t tests since the author finds in practice "randomization-t to be superior to the -c".

across treatments and outcomes: in our first experiment, the combination of three outcome variables and two treatments implies that six hypothesis are tested (i.e., for each outcome variable, two treatment effects are tested relative to the control group). We tested the effect of the randomly provided arguments on the following three outcome variables: preferences for tax simplification (Q2), general need for tax reform (Q3), distributional implications of tax expenditures (Q4). Note that we only reported in detail the results for outcomes Q2 and Q4 because we did not find any effects of the treatment on Q3. However, since we initially intended to study the effect on all three outcome variables, the correct procedure here requires that we adjust standard errors to the case with three outcomes and two treatments.

Using the Westfall and Young (1993)-procedure to adjust standard errors for multiple comparisons, we find a standard error of 0.083 for the effect of the *Redistribution*-treatment on preferences for tax simplification (based on our preferred specification with the full set of control variables). The effect of the *Avoidance*-treatment on preferences for tax simplification is insignificant with a p-value of 0.84. We thus confirm the classical p-values regarding the treatment effects on our main outcome variable, preferences for tax simplification. The p-value of *Avoidance* on the perceived distributional implications, which is significant in the classical inference approach, stands at 0.24 with this method and therefore turns insignificant. All other hypotheses are insignificant with p-values greater than 0.7.

Survey Experiment 2. As with the first experiment, we again adjust standard errors using Young (2018)-type randomization tests and Westfall and Young (1993)-type tests for multiple comparisons. Note that we only have one outcome variable (preferences for tax simplicity) and two treatment groups here, implying that we test only two hypotheses in the context of this second experiment.

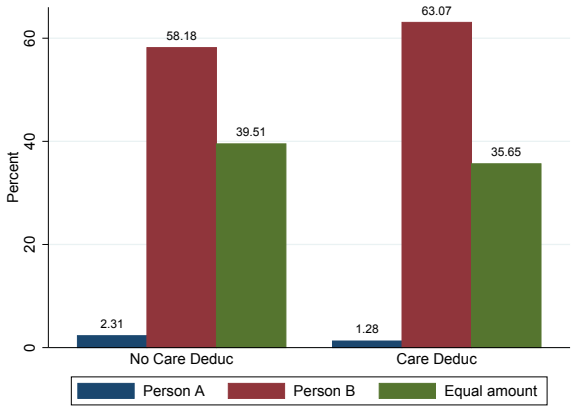
First, the randomization tests come with a p-value for the effect of the *Efficiency*-treatment on preferences for tax simplification of 0.019 (in our preferred specification with full set of controls). That is, the previously reported significance for the *Efficiency*-treatment is confirmed. The effect of the *Special interest group*-treatment remains insignificant with a p-value above 0.5. As with the first experiment, the p-values are remarkably similar to the p-values from classical testing methods. This is reassuring and again indicates that our treatment effects are constant across participants.

Second, the Westfall and Young (1993) method finds adjusted p-values of 0.043 for the *Efficiency*-treatment and 0.48 for the *Special interest group*-treatment (both in specifications with the full set of control variables). The procedure thus confirms the classical inference procedure that treatment *Efficiency* has a significant effect, while treatment *Special interest group* does not.

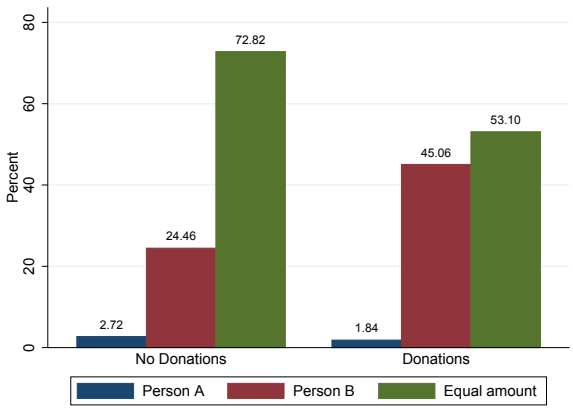
E Further Figures and Tables

Figure E.1: Driven by self interest?

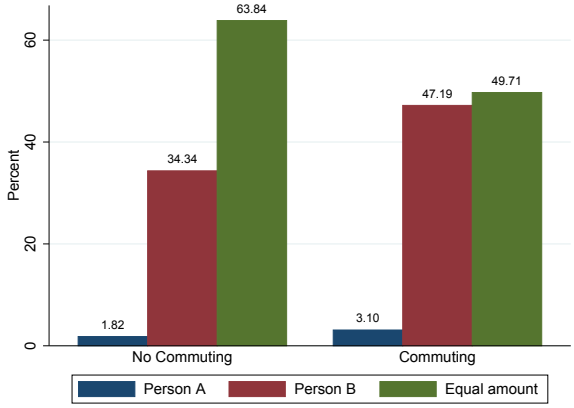
(a) A. Elderly Care



(b) B. Donations

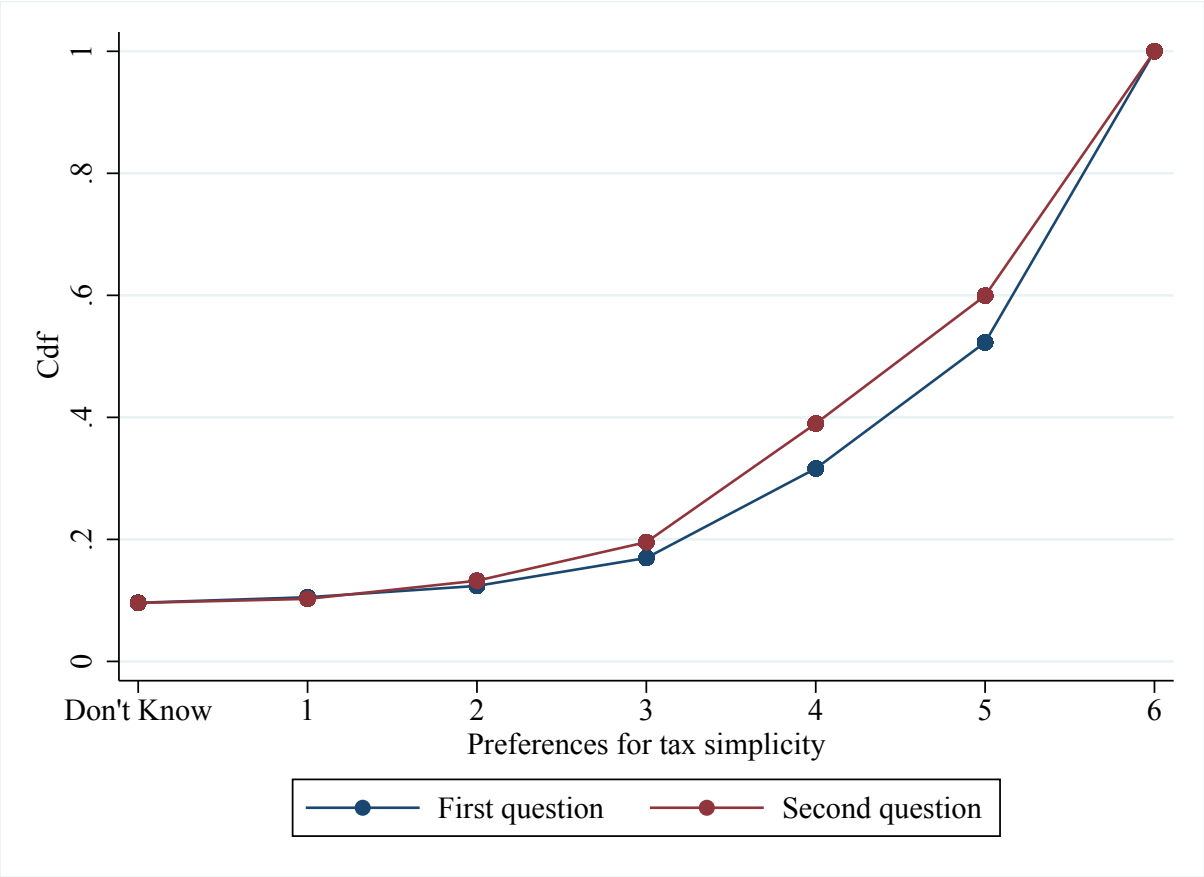


(c) C. Commuting



Notes: This figure depicts the percentage share of respondents in the respective categories by respondents claiming the respective deductions. Since only few people claim deductions for elderly care we consider the broader category of care deductions. Care deductions include deductions for elderly and child care. Panel A question: "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 B have the same gross income and are very similar in all other respects." Panel B question: "Person A spends a considerable amount of her income on charitable giving. Person B does no such thing. Both Person A and B have the same gross income and are very similar in all other respects." Panel C question: "Person A has to travel a considerable distance to work. Person B lives very close to work. Both Person A and B have the same gross income and are very similar in all other respects." Respondents could pick one of the following categories (order of answer categories was randomized): Person A should pay higher taxes; Person B should pay higher taxes; Person A and B should pay equal tax amounts. The left part shows replies for respondents who do not use the respective deduction. The right part shows replies of respondents who do use the respective deduction. TPanel A is based on 2,397 non missing observations, panel B is based on 2,398 non missing observations and panel C is based on 2,394 non missing observations. Source: Own calculations based on German Internet Panel.

Figure E.2: CDF comparison between support for simplification in Q2 and Q9



Notes: The figure shows the cumulative distribution functions (CDFs) for question 2 and question 9. The largest difference between the two distribution functions is 0.0765. The corresponding Kolmogorov-Smirnoff test rejects the hypothesis that the distributions are the same (p-value 0.000). Preference for tax simplification is measured on a 6 point scale based on question 2 "Do you generally think that the income tax system in Germany needs to be simplified?" and question 9: " 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 of the outcome variables is 1 (absolutely not) to 6 (absolutely).

F Details in Text Analysis Approach

In a first step, we prepare the text data from question Q2 of the follow-up survey, for example by removing punctuation, removing superfluous spaces, unifying the text to lower case letters and dropping answers that consist of vowels only or single letters. Then, we apply a spell correction based on the python package SymSpellpy and remove stop words based on a self-created list which is available on request. We don't rely on the stop word list provided by the python text analysis package nltk since it removes words as "keine" (not) which transports valuable information in our context. To reduce the number of distinct words we use the lemmatizer HannoverTagger as described in Wartena (2019) which translates words into their infinitive. An example for this would be the translation of the word *is* to its infinitive *be*. Additionally, we use the snowball stemming algorithm to unify the text further which reduces words to its stem (e.g., consulting and consultant becomes consult). In order to run machine learning predictions we create a training data set by manually assigning one or more of the following topics to 3046 of the answers. Based on reading multiple answers we identify 12 topics: *incentive*, *complexity*, *pro–rich*, *redistribution*, *special interest*, *tax rebate*, *tax equality*, *tax justice*, *simplification*, *special expenditures*, *don't know* and *miscellaneous*. In order to clarify which answers were assigned to the respective categories, we show ten randomly chosen answers by category in Table F.2. The table also shows how our data preparation changes the original answers in practice. For our machine learning text classification, we transform our data into a bag of word representation were the columns of the bag of words matrix are the words and the rows are the survey answers. We then apply a linear support vector machine algorithm to categorize answers into categories. In order to improve the performance of our linear support vector machine, we run a grid search and vary whether term frequencies or term frequency times inverse document frequency is used, the penalty alpha (we try values from $1e^{-2}$ to $1e^{-6}$) and whether single words or bigrams are used. We run predictions and grid search separately for each category which allows for the same answer to be categorized into more than one category. We hold out a random test sample of 20 percent from the training data in order to judge the quality of the prediction. We find accuracy rates in between of 92 for the category *miscellaneous* to 99 percent for *special interest* or *tax justice* category.

Table F.2: Randomly Selected Answers by Topic

Topic	**Orginal Answer**	**Answer after clearing and stemming**
Incentive	Konjunktur stärken	konjunktur stark
Incentive	Schaffung Arbeitsplätze Wohlstand sichern	schaffung arbeitsplatz wohlstand sichern
Incentive	Leistung soll sich lohnen	leistung lohnen

Incentive	Steuerung von Verhaltensweisen der Bürger	steuerung verhaltensweise
Incentive	Investitonen absetzen	investition absetzen
Incentive	Förderung Familien	förderung familie
Incentive	Stärkung der Wirtschaft	stärkung wirtschaft
Incentive	Felxibilität am Arbeitsmarkt	flexibilität am arbeitsmarkt
Incentive	Entsprechende Anreize zu schaffen	entsprechend anreiz
Incentive	Förderung eigener Vorsorge	förderung eigen vorsorge
Complexity	um das Steuersystem weiter kompliziert zu halten	steuersystem weiter kompliziert halten
Complexity	Wer kennt schon alle Möglichkeiten die Abzüge und Freibeträge auszunutzen?	kennen schön allen möglichst abzüge freibetrag ausnutzen
Complexity	Die Verwirrung der Bürger.	verwirrung
Complexity	bevorzugung und verwirrung stiften	bevorzugung verwirrung stiften
Complexity	Es kompliziert zu machen	kompliziert
Complexity	Verwirrung stiften	verwirrung stiften
Complexity	zu kompliziert	kompliziert
Complexity	Das Steuergesetz so kompliziert zu machen damit man ohne Steuerberater nicht auskommt.	steuergesetz kompliziert ohne steuerberater nicht auskommen
Complexity	Verwirrung	verwirrung
Complexity	Zuviel Bürokratie und Ausnahmeregelungen	zuviel bürokratie ausnahmeregelung
Pro rich	Familie wird nicht genug gefördert , nur wer viel Geld hat bekommt mehr Geld . Die kleinen Einkommen können nichts absetzen . Geld kommt zu Geld.	familie nicht genügend fördern nur viel geld bekommen mehr geld klein einkommen können nicht absetzen geld kommen geld
Pro rich	Um den Wohlhabenden noch mehr Möglichkeiten zu geben keine Steuern zu bezahlen.	wohlhabend noch mehr möglichst geben keine steuer bezahlen

Pro rich	Vorrangiges Ziel ist m.E., dass die, die viel Geld verdienen, viele Steuern sparen. Die jedoch, die wenig verdienen (Rente, kleines Zusatzeinkommen - wirkt bei Witwenrente - haben keine bzw. wenig Möglichkeiten um Steuer zu reduzieren.	vorrangig viel geld verdienen viel steuer sparen weniger verdienen rente klein zusatzeinkommen wirken witwenrente haben keine weniger möglichst steuer reduzieren
Pro rich	Begünstigung der oberen Schicht	begünstigung oberer schicht
Pro rich	Meiner Meinung nach dienen sie dazu höhere Gehälter zu entlasten.	meinung nächster dienen hoch gehalt entlasten
Pro rich	Dass Bürger mit hohem Einkommen die Steuerlast drücken	hoch einkommen steuerlast drücken
Pro rich	den kleinen zu fangen und die grossen laufen zu lassen	klein fangen groß laufen lassen
Pro rich	Betrug	betrug
Pro rich	Die reichen noch reicher zu machen	reichen noch reichen
Pro rich	Reichenentlastung	reichenentlastung
Redistribution	Eine Steuerentlastung der Geringverdiener.	steuerentlastung geringverdiener
Redistribution	Renten	rente
Redistribution	Entlastung der Bedurftigten.	entlasten bedürftig
Redistribution	Um niedrige Gehälter nicht zu belasten (Freibeträge) und zum Beispiel Pendler pauschal zu entlasten	niedrig gehalt nicht belastung freibetrag pendler pauschal entlasten
Redistribution	Geringverdiener steuerlich entlasten.	geringverdiener steuer entlasten
Redistribution	Existenzsicherung	existenzsicherung
Redistribution	Schutz von Bürgern mit geringem Einkommen.	schützen gering einkommen
Redistribution	Kinder	kind

Redistribution	soziale Gerechtigkeit	sozial gerecht
Redistribution	Entlastung von Geringverdienern	entlasten geringverdiener
Special interest	Entlastung für bestimmte Gruppen	entlasten bestimmen gruppe
Special interest	Gewisse Personengruppen zu bevorzugen.	gewissen personengruppe bevorzugung
Special interest	Bevorteilung von Familien mit Kinder	bevorteilung familie kind
Special interest	Entlastung bestimmter Gruppen von Bürgern	entlasten bestimmt gruppe
Special interest	Wahlen zu gewinnen	wahl gewinn
Special interest	Lobbyismus in klein-klein	lobbyismus klein klein
Special interest	Um bestimmte Arbeitnehmergruppen steuerlich zu entlasten.	bestimmen arbeitnehmergruppe steuer entlasten
Special interest	Entlastung einzelner Gruppen	entlasten einzeln gruppe
Special interest	Entlastungen für bestimmte Bevölkerungsgruppen	entlasten bestimmen bevölkerungsgruppe
Special interest	gewisse Gruppen zu entlasten	gewissen gruppe entlasten
Tax rebate	Steuerentlastungen	steuerentlastung
Tax rebate	steuerliche Entlastung für mich	steuer entlasten
Tax rebate	Steuersenkung für die Steuerpflichtigen	steuersenkung steuerpflichtig
Tax rebate	Geld sparen	geld sparen
Tax rebate	Steuern zu sparen	steuer sparen
Tax rebate	Steuererleichterungen	steuererleichterung
Tax rebate	Im Idealfall Entlastung des Steuerzahlers	idealfall entlasten steuerzahler
Tax rebate	Steuerlast zu senken	steuerlast senken

Tax rebate	das zu versteuernde Einkommen zu verringern	versteuernd einkommen verringern
Tax rebate	Entlastung	entlasten
Tax equality	Gleichmäßigkeit der Besteuerung	gleichmäßigkeit besteuerung
Tax equality	Finanzieller Lastenausgleich	finanziell lastenausgleich
Tax equality	Um Ungleichheiten und Benachteiligungen auszugleichen.	ungleich benachteiligung ausgleichen
Tax equality	Unterschiedliche Belastungen ausgleichen	unterschiedlich belastung ausgleichen
Tax equality	Ausgleich der Steuerschuld	ausgleichen steuerschuld
Tax equality	Ausgleich von unterschiedlichen Belastungen	ausgleichen unterschiedlich belastung
Tax equality	Finanzieller Ausgleich	finanziell ausgleichen
Tax equality	Ausgleich von schlechter und besser verdienenden	ausgleichen schlecht gut verdienst
Tax equality	Gleichbehandlung	gleichbehandlung
Tax equality	Ausgleich von Benachteiligung	ausgleichen benachteiligung
Tax justice	Gerechtigkeit	gerecht
Tax justice	Steuergerechtigkeit nach Einkommen	steuergerechtigkeit nächster einkommen
Tax justice	Gerechtigkeit	gerecht
Tax justice	Steuergerechtigkeit	steuergerechtigkeit
Tax justice	ungerechtigkeiten auszugleichen	ungerechtigkeit ausgleichen
Tax justice	Steuergerechtigkeit	steuergerechtigkeit
Tax justice	etwas Gerechtigkeit zu schaffen	gerecht
Tax justice	Gerechtigkeit	gerecht
Tax justice	Gerechter Ausgleich	gerecht ausgleichen
Tax justice	Ungerechtigkeiten auszugleichen.	ungerechtigkeit ausgleichen

Simplification	Geringerer administrativer Aufwand	gering administrativ aufwände
Simplification	Rechenvereinfachung	rechenvereinfachung
Simplification	weniger Nachweise, Dokumenten	weniger nachweis dokument
Simplification	Vereinfachung der Dokumentation.	vereinfachung dokumentation
Simplification	Die Steuererklärung leichter zu gestalten.	steuererklärung leicht gestalten
Simplification	Vereinfachung der Steuererklärung	vereinfachung steuererklärung
Simplification	Vereinfachung	vereinfachung
Simplification	Weniger Verwaltungsaufwand bei der Überprüfung der EKST Erklärung	weniger verwaltungsaufwand überprüfen ekst erklärung
Simplification	Vereinfachung der Nachweisfuehrung	vereinfachung nachweisführung
Simplification	Besser keine Ausnahmen !	gut keine ausnahme
Special expenditures	Steuerpflichtigen die Möglichkeir zu geben, aussergewöhnliche Belastunen (z.B. Arztrechnungen/Medikamente u.ä.) steuermindernd geltend zu machen	steuerpflichtig möglichst geben außergewöhnlich belastung arztrechnung medikament steuermindernde geltend
Special expenditures	Persönliche Belastungen anzugeben	persönlich belastung angeben
Special expenditures	Entlastung der Bürger für spezielle Belange	entlasten speziell belangen
Special expenditures	einer außergewöhnlichen Belastung mit Steuererleichterung entgegen zu wirken.	außergewöhnlich belastung steuererleichterung entgegen wirken
Special expenditures	Entlastung der Bürger/-innen in bestimmten Fällen.	entlasten innen bestimmen fall

Special expenditures	Steuerminderung durch Berücksichtigung zusätzlicher bestehender Ausgaben und Verpflichtungen (zb Kinderfreibetrag)	steuerminderung durch berücksichtigen zusätzlich bestehen ausgabe verpflichtung kinderfreibetrag
Special expenditures	Steuerentlastung wegen bestimmter Lebensumstände (Pendler, Kinder, etc.)	steuerentlastung weg bestimmt lebensumstand pendler kind
Special expenditures	Aufwendungen, die der Erlangung von Einkommen dienen und zusätzliche Belastungen zu vergüten.	aufwendung erlangen einkommen dienen zusätzlich belastung vergüten
Special expenditures	Belastungen zu berücksichtigen	belastung berücksichtigen
Special expenditures	Berücksichtigung von Belastungen und Ausgaben.	berücksichtigen belastung ausgabe
Don't know	Schwer zu sagen	schwer sagen
Don't know	kenne ich mich nicht mit aus	kennen nicht
Don't know	Keine Angabe	keine angabe
Don't know	Keine Ahnung	keine ahnung
Don't know	Weiß ich nicht	wissen nicht
Don't know	keine Ahnung	keine ahnung
Don't know	Weiß ich nicht	wissen nicht
Don't know	Da kenne ich mich nicht aus!	kennen nicht
Don't know	Keine Ahnung	keine ahnung
Don't know	Ich habe keine Freibeträge	keine freibetrag
Miscellaneous	Lohnfortzahlung bei Krankheiten (Covid-19)	lohnfortzahlung krankheit covid
Miscellaneous	Freibeträge	freibetrag
Miscellaneous	zu beruhigen	beruhigung
Miscellaneous	Renten beiträge	rente beitrag
Miscellaneous	Freibeträge	freibetrag
Miscellaneous	der bürger bekommt ein leckerli	bekommen leckerli
Miscellaneous	Attraktivitätsprinzip	attraktivitätsprinzip

Miscellaneous	Mehr Kosten den BürgerInnen auferlegen	mehr kosten bürgerinnen auferlegen
Miscellaneous	Um den Steuerzahler bei Laune zu halten	steuerzahler laune halten
Miscellaneous	Betreuungskosten	betreuungskosten

Notes: This table presents 10 randomly drawn answers per category. Beside the category it shows the original answers and the answer after cleaning and stemming the text data.

Table F.1: Perceived role of tax expenditures in open-end question

	Pure control group (Experiment 1 & 2 in wave 36)			**All respondents in wave 46**		
Topic	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Incentive	190	0.068	0.253	3851	0.058	0.235
Complexity	190	0.016	0.125	3851	0.028	0.164
Pro rich	190	0.042	0.201	3851	0.035	0.183
Redistribution	190	0.100	0.301	3851	0.113	0.317
Special interest	190	0.021	0.144	3851	0.035	0.183
Tax rebate	190	0.221	0.416	3851	0.251	0.433
Tax equality	190	0.053	0.224	3851	0.049	0.215
Tax justice	190	0.126	0.333	3851	0.104	0.305
Simplification	190	0.037	0.189	3851	0.041	0.199
Special expenditures	190	0.095	0.294	3851	0.074	0.262
Don't know	190	0.195	0.397	3851	0.185	0.389
Miscellaneous	190	0.153	0.361	3851	0.127	0.333

Notes: The table presents summary statistics of the topics mentioned by survey respondents. The table shows the frequencies at which respondents in either the pure control group (i.e. subjects who were neither in the treatment group in Experiment 1 nor Experiment 2 in the main survey of wave 36 and are therefore not influenced by information provision) or all respondents in wave 46 mentioned topics when being asked about the primary goal of tax deductions and exemptions in wave 46. Data is taken from the follow-up survey.

G Detailed Questionnaire

G.1 Detailed Questionnaire of the Main Survey

This Appendix section presents the translated survey questions including reply categories. The order of presentation and the numbering of the question corresponds with the description of the survey structure in section 3.2.

- **Introduction:** *In the following, we would like to ask you some questions about the tax system in Germany. We will focus in particular on rules surrounding the income tax and whether they are complicated or easy to understand. Whether a tax system is generally complicated or easy to understand depends in particular on the number of possible deductions and allowances.*

- **Q1:** How difficult is it for you to fill out your tax return?
  1 Very easy;...; *5 Very difficult*; *I do not know because no taxes are declared in my name*; *I do not know because I do not declare taxes myself (rather, my partner or a tax consultant, etc. does this)*; *I do not know*

- **Randomized Experiment 1**: See body of the text (section 3.3) and Appendix section H below.

- **Q2:** Do you generally think that the income tax system in Germany needs to be simplified?
  1 Absolutely not;...; *6 Absolutely*; *I do not know*

- **Q3:** Do you generally believe that the income tax system in Germany is in need of reform?
  1 Absolutely not;...; *6 Absolutely*; *I do not know*

- **Q4:** Do you think that numerous deductions and allowances contribute to a fairer distribution of income, or do you believe that high-income citizens benefit more from these deductions and allowances?
  1 They contribute to fairer income distribution;...; *6 High-income citizens benefit*; *I do not know*

- **Q5:** 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.
  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 do not know*

- **Introduction for Q6-8** Imagine two persons, A and B. Which person do you think should pay more taxes in the following situation?

- **Q6:** 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

B have the same gross income and are very similar in all other respects. (randomize order of answer categories)
Person A should pay higher taxes; *Person B should pay higher taxes*; *Person A and B should pay equal tax amounts*

- **Q7:** Person A spends a considerable amount of her income on charitable giving. Person B does no such thing. Both Person A and B have the same gross income and are very similar in all other respects. (randomize order of answer categories)
Person A should pay higher taxes; *Person B should pay higher taxes*; *Person A and B should pay equal tax amounts*

- **Q8:** Person A has to travel a considerable distance to work. Person B lives very close to work. Both Person A and B have the same gross income and are very similar in all other respects.(randomize order of answer categories)
Person A should pay higher taxes; *Person B should pay higher taxes*; *Person A and B should pay equal tax amounts*

- **Randomized Experiment 2**: See body of the text (section 3.3) and Appendix section H below.

- **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?
1 Absolutely not;...; *6 Absolutely*; *I do not know*

- **Q10:** Which of the following deductions and/or allowances do you usually use when filing your income tax?
Maintenance of two households; *Home office*; *Commuting allowance*; *Other job related expenditures*; *Pension expenses*; *Education costs*; *Care relatives*; *Child allowance, childcare*; *Donations*; *Others [insert text]*; *No deductions*; *I do not know*

G.2 Detailed Questionnaire of the Follow-up Survey

This Appendix section presents the translated survey questions including reply categories.

- **Q1:** Which of the following items specifically makes it complicated to file an income tax declaration? (multiple answers can be choosen)
Amount of documentation obligations; *Fear of forgetting something important*; *Fear of doing something wrong*; *Different tax rates in income tax tariff*; *Other*; *None of these*; *Don't know*

- **Q2:** In the German income tax system exist deduction possibilities and allowances. What is in your opinion the main goal of the existing deduction possibilities and allowances? Please insert your answer in the field below.

- **Q3:** If you could decide on what to use deduction possibilities and allowances in the German income tax system: To which end would you use these exemptions and deduction possibilities mainly? Please insert your answer in the field below.

- **Q4:** Please name all deduction possibilities and allowances in the German income tax system that you are aware of. It is fine if you do not know the official name of the deduction possibility or the allowance. In this case you can describe it in your own words. Please insert your answer in the field below.

- **Q5:** In the following we list five deduction possibilities and allowances in the German income tax system which are used often. Please order them depending on whether in your opinion they make the tax system more fair. Order the five cases in descending order and start with the cases that in your opinion make the income tax system most likely more fair. Therefore, enter a 1 for the deduction or allowance that you believe is most most important for a fair tax system. For the deductions or allowances that you consider to be the second, third, and fourth most important deductions or allowances, please enter a 2, 3 and 4. For the least important deduction or allowance, please enter a 5. (randomize order of answer categories)
  Expenses related to professional activity (e.g. Costs for the commute to work); *Donations for charitable purposes or political parties*; *Pension expenses for retirement and/or statutory pension insurance*; *Costs for the care of relatives*; *Child allowances*

- **Introduction for Q6:** Imagine two persons, A and B. Which person do you think should pay more taxes in the following situation?

- **Randomized Experiment (follow-up)**: One third of the respondents is randomly assigned to the *control* group, one third to the *choice* treatment group and one third to the *circumstance* treatment group.

- **Q6 - *control*:** Person A has to travel a considerable distance to work, while Person B lives very close to work. Both Person A and B have the same gross income and are very similar in all other respects.
  Person A should pay higher taxes; *Person B should pay higher taxes*; *Person A and B should pay equal tax amounts*

- **Q6 - *circumstance*:** Person A was relocated by his employer and has to travel a considerable distance to work ever since, while Person B lives very close to work. Both Person A and B have the same gross income and are very similar in all other respects.
  Person A should pay higher taxes; *Person B should pay higher taxes*; *Person A and B should pay equal tax amounts*

- **Q6 - *choice*:** 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 for a job with a very long way to work, while Person B lives very close to work. Both Person A and B have the same gross income and are very similar in all other respects.

Person A should pay higher taxes; *Person B should pay higher taxes*; *Person A and B should pay equal tax amounts*

- **Q7:** Which of the following deductions and/or allowances do you usually use when filing your income tax?
  Maintenance of two households; *Home office*; *Commuting allowance*; *Other job related expenditures*; *Pension expenses*; *Education costs*; *Care relatives*; *Child allowance, childcare*; *Donations*; *Others [insert text]*; *No deductions*; *I do not know*

H Illustration of Treatment Conditions

First experiment

Figure H.1: Control condition

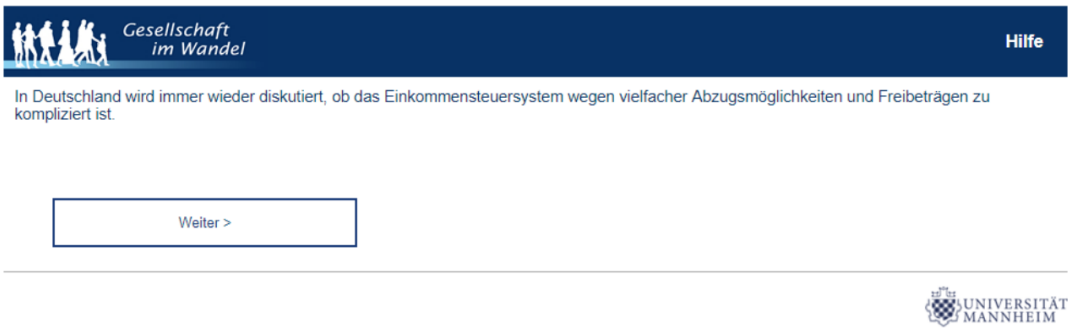


Figure H.2: Redistribution treatment

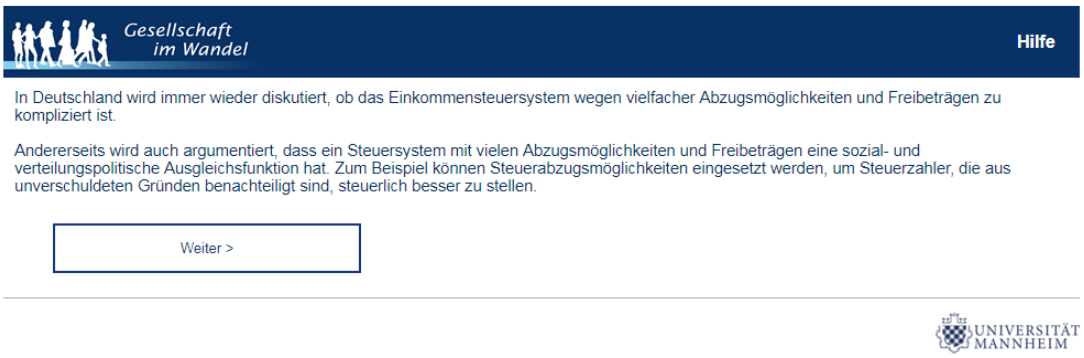
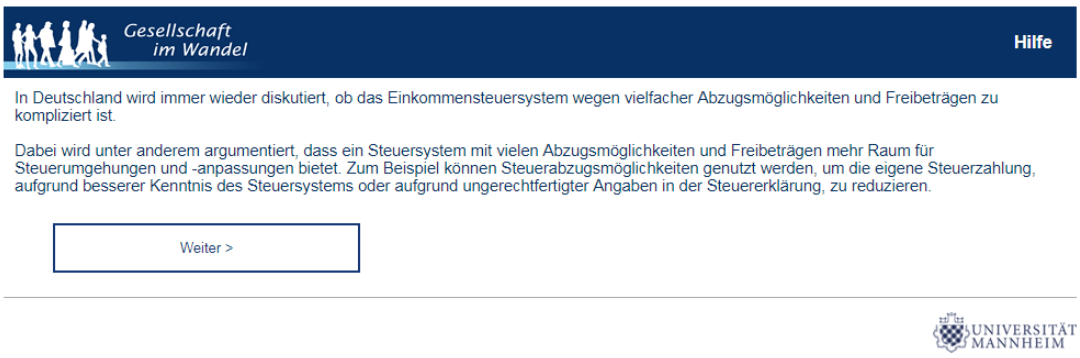


Figure H.3: Avoidance treatment



Second experiment

Figure H.4: Control condition

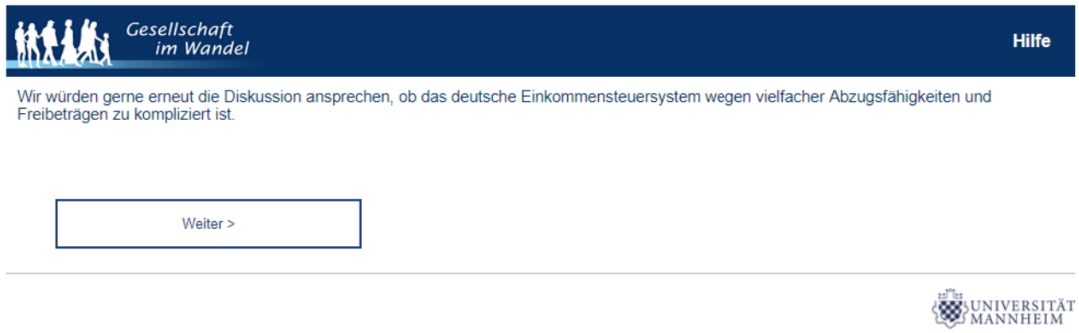


Figure H.5: Economic efficiency treatment

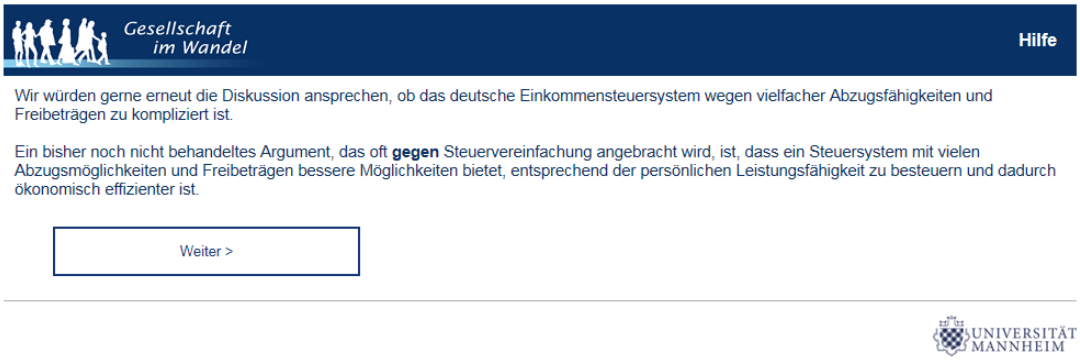


Figure H.6: Special interest group treatment