Journal of Economic Psychology 35 (2013) 1–16

Contents lists available at SciVerse ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Decision heuristics and tax perception – An analysis of a tax-cut-cum-base-broadening policy

Kay Blaufus [a,*], Jonathan Bob [a], Jochen Hundsdoerfer [b], Dirk Kiesewetter [c], Joachim Weimann [d]

[a] Leibniz Universität Hannover, School of Economics and Management, Königsworther Platz 1, D-30167 Hanover, Germany
[b] Freie Universität Berlin, School of Business & Economics, Garystr. 21, D-14195 Berlin, Germany
[c] Julius Maximilian University of Würzburg, Faculty of Economics and Management, Sanderring 2, D-97070 Würzburg, Germany
[d] Otto-von-Guericke University Magdeburg, Faculty of Economics and Management, Universitätsplatz 2, D-39016 Magdeburg, Germany

ARTICLE INFO

Article history:
Received 18 June 2012
Received in revised form 13 December 2012
Accepted 16 December 2012
Available online 7 January 2013

JEL classification:
D03
G11
H20
H30
K34
M41

PsycINFO classification:
2340
2960

Keywords:
Behavioral public finance
Decision heuristics
Framing effects
Perceived tax burden
Tax-cut-cum-base-broadening

ABSTRACT

In this paper, both a conjoint analysis and a lab experiment are conducted to analyze the influence of changes in the tax rate and the tax base on the perceived tax burden. Our results show that the majority of individuals do not make rational tax decisions based on the actual tax burden but rather use simple decision heuristics. This leads to an irrationally high impact of changes in nominal tax rates on the perceived tax burden. Taxpayers favor tax options that apply a lower tax rate on their gross income over a higher tax rate applied on their net income despite the lower actual tax burden of the latter option. This result suggests that politicians could combine increasing fiscal revenues and decreasing subjects' tax perception. Furthermore, overestimation of tax rate changes increases considerably when information on tax rate is considered first (framing effect).

1. Introduction

Following the United Kingdom and the United States, in recent decades, many countries pursue a policy of tax-cut-cum-base-broadening (OECD, 2010: 11), i.e., they reduce the nominal tax rate and simultaneously increase the tax base. One recent example is the introduction of a final withholding tax on interest income in many OECD countries. These countries cut the nominal tax rate on capital income but abolish the possibility to deduct capital income-related expenses (Genser & Reutter, 2007). Moreover, several countries provide explicit tax options that reflect the above mentioned tax-cut-cum-base-broadening policy. Exemplarily, small corporations in Russia may choose between a lower tax rate on their gross

* Corresponding author. Tel.: +49 511 7625660; fax: +49 511 76219858.
E-mail address: blaufus@steuern.uni-hannover.de (K. Blaufus).

0167-4870/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.joep.2012.12.004

income and a higher tax rate on their net income. In France and Germany taxpayers may—under certain conditions—choose whether their dividends are taxed under a progressive tariff based on their net income or under a flat tax based on their gross income. The aforementioned tax reforms are usually reasoned by expected efficiency gains and tax simplification (e.g., Kopczuk, 2005). We propose a behavioral approach as an additional explanation as to why politicians pursue such a policy of tax-cut-cum-base-broadening: the use of decision heuristics that lead to an irrationally high impact of nominal tax rate reductions on taxpayers' perceived burden.

Previous economic and psychological research already shows that attitudes toward tax policies are subject to several biases (McCaffery & Baron, 2004). As early as 1960 Schmölders finds evidence that a high percentage of entrepreneurs overestimate their tax burden (Schmölders, 1960: 84f). Especially, if tax complexity is high or if taxes are non-salient, subjects make substantial decision errors (Boylan & Frischmann, 2006; Chetty, Looney, & Kroft, 2009; Rupert, Single, & Wright, 2003; Rupert & Wright, 1998). In addition, research shows that price complexity influences buyers' price perceptions and demands. Breaking down the price into several components (e.g., base price and shipping costs) leads to a decrease in the perceived price and an increased demand for the corresponding commodity (see Morwitz, Greenleaf, & Johnson, 1998). Krishna and Slemrod (2003) discuss the potential meaning of these price research results for tax policy and Chetty et al. (2009) find that using prices plus sales tax instead of the net amounts leads to a significant reduction in demand. The following article is also based on this idea and is, to our knowledge, the first to examine whether the perceived tax burden is dependent upon which price component (tax rate or tax base) is changed.

The calculation of the tax burden is regarded as a complex task (Kirchler, 2007: 32). As we know from previous research decision heuristics are used in such complex tasks, particularly. Therefore, it seems reasonable to assume that subjects will use heuristics in our tax framing. Even though there are papers on the use of decision heuristics in tax contexts (e.g., McCaffery & Baron, 2003), a paper on the relationship between decision heuristics and a tax-cut-cum-base-broadening policy is missing.

The main purpose of this study is to examine whether the use of decision heuristics leads to a misperception of changes in nominal tax rates compared to changes in tax deductions. We analyze (i) whether decision heuristics are used in choosing between tax systems that differ, exclusively, in the nominal tax rate and the amount of tax deductible expenses, (ii) whether the use of decision heuristics will induce subjects to overemphasize the importance of nominal tax rate changes relative to changes in the amount of tax deductible expenses, and (iii) whether framing effects drive the misperception of nominal tax rate changes compared to tax base changes.

To this aim, we conduct (i) a survey study (conjoint analysis) with a sample of German working individuals that represents the population in terms of age, education, gender, and income and (ii) a lab experiment with working individuals who are compensated according to their experimental performance. By combining these two methods we are able to benefit from the respective strengths of both procedures. First of all, conjoint analysis allows us to draw a large (representative) sample. Moreover, we are able to measure preferences on the individual level. This incorporates the advantage of conjoint analysis to identify whether individuals choose a decision heuristic over the rational ranking. On the other hand, conjoint analysis is a mere preference measurement whereas a lab experiment sets monetary incentives. Besides, while in conjoint analysis the task is performed just once, the lab experiment enables us to execute the task multiple times.

The remainder of this article is organized as follows. First, in Section 2 we discuss the effects that decision heuristics may have on the perceived importance of nominal tax rate changes and derive our hypotheses. The sample, method, and results are presented in Sections 3 and 4, "Conjoint Analysis" and "Experimental Validation", respectively. A discussion of the results and the implications for tax policy and research are carried out in Section 5.

2. Hypotheses

We consider hypothetical tax options and put subjects in the position of a choice between these alternatives that, for a given income, differ only in terms of the nominal tax rate and the allowed deduction of income-related expenses.[1] This setting mirrors the often proposed policy option to eliminate itemized deductions (e.g., Pitt & Slemrod, 1989; Slemrod, 1989). Opposition to itemized deductions is in line with public opinion as 54% of Americans would be willing to give up some deductions to make the tax system simpler (Tax Foundation, 2005). According to Elffers and Hessing (1997) subjects favor higher standard deductions over itemized deductions because taxpayers "prefer being lazy to becoming tired". However, our main purpose is to study whether the use of decision heuristics leads subjects to misperceive nominal tax rate changes compared to tax base changes. We are not interested in the preferences for tax simplification. To control for such preferences subjects are explicitly informed that each tax system from which they can choose leads to the same compliance effort. Therefore, rational decision makers choose among these alternatives based on the actual tax burden B_i, which can be written for the ith alternative as

$$B_i = \tau_i Y - \tau_i D_i \tag{1}$$

where τ_i is the tax rate, D_i is the deduction of income-related expenses, and $Y > D_i$ is taxable revenue (identical for all alternatives).

[1] It is important to note that we are not looking at tax reform models that differ such that certain economic activities are taxable in one alternative but not the other. As is well known from literature (see e.g., Willner & Granqvist, 2002) a base-broadening, rate-reducing policy that taxes previously untaxed opportunities while reducing the tax rate on all taxable opportunities could lead to efficiency gains even if tax payments remain the same. By contrast, we are studying the effect of taxing a single (already taxable) economic activity differently in each tax option.

Traditional theory assumes that economic subjects do not make arithmetic errors and that the calculations included in the decision-making process do not require a great deal of cognitive effort.

In contrast, psychological research shows that a human's cognitive ability is limited and that the calculation of decision variables causes cognitive strain, making it reasonable for individuals to use simple decision heuristics rather than exact calculations (see Tversky & Kahneman, 1974). Using a decision heuristic may lead to a suboptimal solution that is, however, individually satisfactory (Simon, 1990). Particularly in tax settings, it seems reasonable to assume that subjects will use heuristics. The calculation of the tax burden is regarded as a complex task and usually causes enormous cognitive strain. Kirchler summarizes: "People blame the complexity of the tax law for their feelings of tax incompetence." (Kirchler, 2007: 32). In the light of tax complexity the advantage of selecting a decision heuristic consists in lower cognitive strain and less time spent on making the decision.

This leads to our first hypothesis to be tested:

H1. The decision among alternatives is not based exclusively on the actual tax burden but rather on the use of decision heuristics.

One reasonable heuristic is what is known as the anchor heuristic. This heuristic has been observed empirically in several other contexts (Epley & Gilovich, 2006; McCaffery & Baron, 2003; Tversky & Kahneman, 1974). According to the anchor heuristic, individuals who estimate a certain value, such as the actual tax burden, begin with a starting value that serves as an anchor for finding the estimated value. The disadvantage of this heuristic is that the anchor is consistently overweighted, and additional information is not adequately included.

The anchor is often chosen by selecting either the information with which the individual is first confronted (see Hogarth & Einhorn, 1992) or the information which is considered most important (Yadav, 1994). The anchor value is then adjusted (inadequately) based on later information or information that is considered to be less important.

In our opinion, there are several reasons why the tax rate and not the amount of deductible expenses serves as the anchor.

First, information about the nominal tax rate τ_i is, in reality, more readily available than information about the deductibility of single expenditures. According to the availability bias, people overweight evidence that is easily available (Tversky & Kahneman, 1973: 211).

Second, in general, tax liability reacts more elastically to changes in the tax rate than to changes in deductions. A 1% increase in tax rate always leads to a 1% increase in tax liability. However, a 1% decrease in deductions, usually, leads to a tax increase of less than 1%.[2] Thus, the amount of income-related expenses to be deducted could be considered less important by some individuals.

Third, the effects of different tax rates on a given income can be easily recognized such that individuals determine the positive relation between tax rate and income without cognitive strain. By contrast, with the influence of the deduction of income-related expenses, there is a negative relation between the tax base and income-related expenses as well as a positive relation between the tax base and tax liability. The marginal necessary cognitive effort compared to the tax rate effects, thus, supports the assumption that the tax rate, and not income-related expenses, serves as an anchor.

In addition to the anchor heuristic, the use of a lexicographic heuristic is reasonable (e.g., Brandstaetter, Gigerenzer, & Hertwig, 2006). Individuals reduce their cognitive effort by first evaluating the alternatives based on only one criterion, and if no decision is possible with this criterion, they apply other criteria. Based on the high relevance of the tax rate presumed above, we assume that individuals who use a lexicographic heuristic first evaluate alternatives based on the tax rate; only if information regarding tax rates is equivalent they consider deductible expenses.

The use of the mentioned heuristics in combination with the assumption of the perceived high relevance of the tax rate leads to the following hypothesis:

H2. The effect of changes in the tax rate (changes in the tax base) is overestimated (underestimated) compared to the rational benchmark.

Extensive behavioral and psychological research shows that human behavior is subject to framing effects (McCaffery & Baron, 2004; Tversky & Kahneman, 1981). Presenting the same option in different ways can alter subjects' decisions. One example is that individuals tend to choose the piece of information with which they are first confronted as an anchor or as the primary criterion for the lexicographic heuristic (Blaufus & Ortlieb, 2009; Chrzan, 1994; Moran & Meyer, 2006). Thus, we hypothesize:

H3. Confronting subjects first with the tax rate increases the likelihood that subjects overestimate the importance of changes in the tax rate.

3. Conjoint analysis

3.1. Method

To test the hypotheses H1–H3 formulated in the previous section, we conduct a conjoint analysis. Conjoint analysis is based on Luce and Tukey (1964). The aim of this method is to derive the influence of attributes (and their levels) on the total

[2] The elasticity equals $\frac{D_i}{Y-D_i}$ and is always less than one percent if $Y > 2D_i$, which is typically the case.

utility of a combination of attributes (stimuli). For this purpose, subjects are given various stimuli to evaluate. Conjoint analysis is a decomposition method in which the estimation of influence (part-worth utilities of attribute levels, relative importance of attributes) is based on empirically collected total utility of the respective stimuli (see Hair, Black, Babin, & Anderson, 2008).

The subjects' task in conjoint analysis is to rank several of these stimuli according to their personal preferences.[3] The subjects assign the lowest (highest) rank to the stimulus with the highest (lowest) preference. The total utility of the stimuli is derived from the individual ranking by each subject. For this purpose, the stimulus with the lowest (highest) rank is assigned the highest (lowest) utility. Metric part-worth utility for the attribute levels is determined using the calculated total utility and the ordinary least squares method.[4] As is standard in conjoint analysis, we assume an additive model for the relation between total utility and part-worth utility, i.e., the sum of the part-worth utilities of a stimulus corresponds to its total utility. This follows from the basic assumption that the explanatory variables do not interact (see Hair et al., 2008).[5] This leads to the following relationship:

$$U_{i,j} = \mu_i + \sum_{k=1}^{K} \beta_{k,i} x_{k,j} + \varepsilon_j \tag{2}$$

where $U_{i,j}$ represents the total utility of the jth stimulus for the ith subject. μ is the constant, and $\beta_{k,i}$ are the part-worth utilities of the attribute levels. The dummy variables $x_{k,j}$ take on a value of one if the observed stimulus contains the respective attribute level. ε_j is the error term.

As can be seen in Eq. (2), the part-worth utilities are estimated on the subjects' level. To compare and aggregate the part-worth utilities among subjects, they have to be standardized (Green & Krieger, 1985, 3f). For this purpose, the highest total utility is set to the value of one while the lowest total utility is set to zero. Hence, the sum of the most preferred level of each attribute is one, whereas standardized part-worth utilities of the fewest preferred levels are all set to zero. The relative importance of each attribute equals the part-worth utility of the most preferred level for this attribute. The more the total utility of a stimulus changes when the level varies for a given attribute, the higher is the relative importance.

Conjoint analysis is predominantly used in marketing research. However, conjoint analysis has also been used to measure tax effects (e.g. Blaufus & Ortlieb, 2009; Hundsdoerfer & Sichtmann, 2009; Milliron & Toy, 1988; O'Neil, 1982). The idea is to define tax characteristics (e.g. tax rate and deductible expenses) as attributes of products. Traditional conjoint analysis[6] allows to estimate the relative importances of these tax characteristics at the subject level (see Green & Srinivasan, 1978: 104) and compare the measured importances with the importances for a "rational" taxpayer. For this purpose, the part-worth utility of a tax characteristic for a "rational" tax payer will be compared with the actual measured part-worth utility. Thus, we can identify whether subjects misperceive the impact of changes in specific tax characteristics like the nominal tax rates.

A further advantage of conjoint analysis is the simultaneous evaluation of the attributes. Alternatively, one could directly ask subjects for the value they attach to an attribute. This sequential evaluation has the disadvantage that subjects tend to neglect the trade-off effects. All attributes are considered to be very important and hence, the importance of the individual attribute is overrated. Contrary, because of the simultaneous evaluation of the attributes in conjoint analysis subjects must keep in mind the trade-off effects between attributes which also exist in reality.

3.2. Sample

A total of 467 working individuals are interviewed who match the population in terms of the following attributes: gender, age, education, and monthly net income. The selection of working individuals ensures that subjects have experience with income taxation.

Trained interviewers conduct standardized face-to-face interviews that last an average of 20 min. In addition to conjoint analysis, subjects are asked questions regarding demographic attributes, general attitudes toward tax policy, current German income tax law, and tax complexity.

The sample is drawn based on a quota schedule,[7] as a pure random sample was not feasible financially. The quota parameters are based on the following four attributes: age, gender, education, and monthly net income. The corresponding frequency

[3] This ranking is the most common valuation procedure after the rating scale. An overview of various procedures is given in Green and Srinivasan (1978, 104).

[4] The subjects' preference judgments, expressed through their rank ordering, have an ordinal measurement level. Hence, one can apply a monotone variance analysis. However, the least squares method has proven to be very robust in the estimation of part-worth utility values also in the case of ordinally scaled dependent variables (see e.g. Green & Krieger 1993, 478).

[5] Eq. (1) leads to an (at least theoretical) interaction between the two attributes tax rate and income-related expenses. However, by assuming the rational valuation of a homo economicus, this is a no-crossover interaction (see Green & Devita, 1974, 56). Thus, in line with previous research, the interaction can be neglected in the following analysis (see Carmone & Green, 1981, 93).

[6] Other conjoint analysis procedures (e.g. Choice Based Conjoint Analysis) do not allow for an estimation of preferences on the individual level and are therefore unsuitable for the purpose of our study. Related models for preference measurement, such as the Rank-Ordered Logit, allow for estimation on the individual level but require subject specific attributes in addition to the attributes of the stimuli, which distinguish the evaluation of the stimuli attributes by the subjects. See Allison & Christakis, 1994, 202.

[7] Quota samples do not strictly fulfill the requirements of a pure random selection. Nevertheless, it is the most widely used procedure in marketing research and continuously yields good results in comparative studies with pure random selection (see Green, Tull, & Albaum, 1988, 325–327).

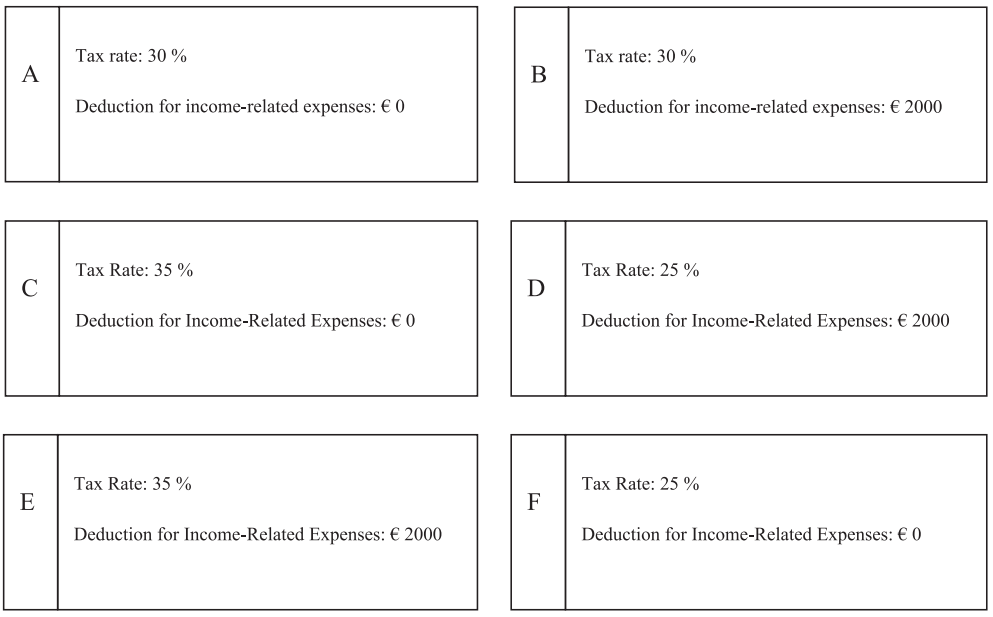

Fig. 1. Stimuli used in conjoint analysis.

in the population is taken from the Statistical Yearbook of the Federal Statistical Office in Germany, which covers the 37 million people who make up Germany's working population.

Compliance with the quota is statistically tested by a chi-squared test. With a margin of error of 5%, no significant difference between population and sample can be detected. In this respect, the sample can be seen as representative of the working population in Germany. The distribution of attributes in the sample is given in Table 11 Appendix A.

Of the 467 polled individuals, 33 favor high tax rates and non-deductibility of income-related expenses. Because the sincerity of this stated preference is doubtful,[8] these individuals are not included in the analysis. The adjusted sample, therefore, includes 434 individuals who also match the population in terms of gender, age, education, and monthly income.

3.3. Research design and operationalization of the hypotheses

Subjects are asked to rank various tax systems according to their individual preferences. For this purpose, subjects are asked to assume they receive taxable earnings (an interest payment) of € 10000 and bear income-related expenses of € 2000. The individual tax systems differ solely in terms of tax rate τ_i and allowable deductions of income-related expenses D_i and, therefore, in terms of tax burden B_i, which can be calculated using Eq. (1).

Interviewers explain this income should be considered as the respondents' own income, and that every tax system leads to the same compliance effort (''time required for the tax return'').

The stimuli are presented using the Full Profile Method, i.e., each stimulus exhibits a combination of the two attributes (tax rate and deduction of income-related expenses). Three levels are chosen to express the tax rate (low, medium, high), and two levels are selected for income-related expenses (no deduction, full deduction). This combination of levels yields a maximum of six (3×2) different stimuli.[9] The corresponding complete design is presented in Fig. 1.[10]

The individual stimuli are given to the subjects in the form of randomly ordered laminated cards. The random issuance of the stimuli ensures that the order of presentation has no influence on the valuation (regarding the ''sequencing effect'' see e.g., Tourangeau & Rasinski, 1988: 304).

Subjects are asked to arrange the cards on a magnetic board in previously numbered positions according to their preferences. Before the subjects begin, the interviewer explains the terms ''tax rate'' and ''income-related expenses'' as well as their

[8] For instance, some subjects simply ranked the stimuli alphabetically. An examination showed that the exclusion of these so called reversals had no influence on the presented results.

[9] In addition to the two attributes of tax rate and tax base, the stimuli contained a third attribute (''time required for the tax return''), which is not relevant in the current study, and which only serve as a control in the current study. No significant differences can be detected between the stimuli used and the control stimuli.

[10] Two different stimuli sets are used to test whether the amount of tax rates and income-related expenses influences subject's decision. These two sets differ in the differences between the particular tax rates and expenses. The values on the second stimuli set amount to 25%, 27%, and 29% (tax rate) as well as € 0 and € 1000 (income-related expenses). Of course, rational rankings stay unchanged.

Table 1
Tax burden.

Stimulus	Tax burden (€)
A	3000
B	2400
C	3500
D	2000
E	2800
F	2500

effects on the tax burden using the numbers given in the actual decision task. Furthermore, subjects are made aware of the "objectively best tax alternative", which is shown on card D (see Fig. 1).[11] They then have to rank only the remaining alternatives according to their given income. After the subjects conduct the rankings, the interviewer offers the subjects a chance to review their choices. After the subjects made their final choices, the interviewer records the final preference ranking.

The subjects' rankings serve to test whether the individuals conduct the rankings rationally—according to the actual tax burden—or whether they use decision heuristics (hypothesis H1).

To test whether changes in tax rates are overestimated (hypothesis H2), the ranking is used to estimate part-worth utilities and relative importances. To this aim, we compare the results of our sample with the rational benchmark of the "homo economicus" and consider the relative importance to be overrated if and only if it exceeds the rational benchmark significantly.

For testing hypothesis H3, we use a between-subject design with the independent variable "framing effect". In the first group the tax rate is the first information mentioned, while in the second group it is mentioned last. The stimuli used in the first of the two groups are given in Fig. 1 (tax rate mentioned first). The stimuli used in the second group are the same except for the permutation of the item order.

Subjects are randomly divided into two groups. The 190 subjects in the first group are first presented with the attribute "nominal tax rate", the 244 subjects in the second group are first presented with the attribute "tax deductible expenses".

3.4. Results

3.4.1. Analysis of ranking behavior (Hypothesis H1)

To test hypothesis H1, we analyze whether the empirically observed rank order of the stimuli matches the prediction of the rational model. "Calculating" rational investors minimize their tax burden. Tax burdens of the six stimuli can be calculated using Eq. (1) and are shown in Table 1.

The rational ranking is given as follows:

$$D \succ B \succ F \succ E \succ A \succ C.$$

To reduce cognitive strain, the individuals may use a simple heuristic instead of computation. In Section 2 the anchor heuristic and lexicographic heuristic are highlighted. If individuals rank lexicographically first according to tax rate and subsequently according to income-related expenses, then the following ranking, which differs from the rational ranking, results:

$$D \succ F \succ B \succ A \succ E \succ C.$$

Although we predict that the lexicographic ranking uses the tax rate as the first criterion, it should be tested whether there are individuals whose lexicographic ranking uses income-related expenses as the primary decision criterion. If the stimuli are lexicographically ranked first according to deduction of income-related expenses and then to tax rate, we would expect the following ranking results:

$$D \succ B \succ E \succ F \succ A \succ C.$$

We choose appropriate values permitting a clear distinction between rational and lexicographic ranking. An overview of the proportion of rankings in the sample is shown in Table 2.

From Table 2, it is apparent that a surprising number, more than 90% of the subjects, decide against the rational "homo economicus" ranking. Only 9.4% (41 individuals) follow the predictions based on the model of rational net income maximization.[12] Hypothesis H1 is, therefore, confirmed.

Using a logistic regression, we test for demographic factors as age, gender, education, tax knowledge, and income to distinguish between the different groups of subjects, i.e. those subjects who rank their stimuli rational, those who use a lexicographic heuristic and those who prefer other rankings. Tax knowledge was measured in three different ways. First,

[11] The indication of the objectively best alternative, and in the case of the six control cards (see Fn. 10) also of the objectively worst alternative, served to reduce the work of the subjects.

[12] The probability of randomly achieving a rational sequence is 0.83%. A binomial test shows that the percentage of the rational sequence cannot be the result of pure random selection ($p < 0.01$).

Table 2
Percentage of subjects per ranking.

Sequence	Subjects (%)
Rational	9.4
Lexikographic	
1st tax rate, 2nd income-related expenses	35.0
1st income-related expenses, 2nd tax rate	21.7
Other sequences	33.9
Total	100.0

Table 3
Logistic regression: determinants of a rational ranking.

Dependent variable: rational sequence (dummy)	β coef.	Std. error	β coef.	Std. error	β coef.	Std. error
Constants	−2.217***	0.709	−2.144***	0.688	−2.136***	0.692
With university-entrance qualification	0.193	0.356	0.187	0.358	0.28	0.352
Above average knowledge of tax law (self-assessment)					0.190	0.523
Very good knowledge of tax law (income tax rates)			1.236	0.565		
Tax return self-prepared	0.806**	0.374				
Gender	−0.153	0.344	−0.073	0.341	−0.075	0.346
Age	−0.011	0.015	−0.008	0.014	−0.007	0.014
Self-employed	−0.046	0.531	−0.033	0.537	−0.133	0.532
Net income < 1000 €/month	0.422	0.377	0.233	0.367	0.260	0.365
N	430		430		430	
Nagelkerke's R2	0.031		0.015		0.012	

β coef., β coefficients; Std. error, Standard error.
** Indicates significance on the 5% level.
*** Indicates significance on the 1% level.

subjects' were asked to self-assess their tax knowledge. Second, we asked subjects questions on the current actual income tax tariff. Third, we asked whether subjects self-prepare their tax return.[13] As it can be seen from Table 3, only tax knowledge could explain the subject's decision. Participants who self-prepared their last tax return (who have very good knowledge of tax law) choose around twice (three times) as likely the rational ranking. All other demographic variables are not significant.

Table 2 also shows that more than half of the subjects rank the stimuli lexicographically. One-third of all subjects order the tax systems lexicographically with tax rate the primary criterion. Contrary to our assumption, 21.7% of the individuals use a lexicographic heuristic in which income-related expenses are the dominant criterion. The probability of randomly achieving one of the two lexicographic sequences is 1.7%. Of the group of 434 subjects, seven could have arrived at a lexicographical sequence by randomly ordering the stimuli. Thus, one can assume that the two lexicographical heuristics are consciously chosen.[14]

These explanations point out that we can easily explain the behavior of two-thirds of our respondents. However, one-third chose a ranking that was neither rational nor lexicographical. In the next subsection, we show that the use of the anchor heuristic seems to be the most reasonable explanation for the remaining rankings.

3.4.2. Overestimation of the importance of changes in tax rate (Hypothesis H2)

According to hypothesis H2, the use of heuristics leads to an overestimation of the relative importance of the tax rate. Our approach is to estimate part-worth utilities by Eq. (2) using OLS and to compare the results of our sample with the rational benchmark of the "homo economicus"

$$\widehat{U}_{i,j} = \hat{\mu}_i + \hat{\beta}_{1,i}x_{1,j} + \hat{\beta}_{2,i}x_{2,j} + \hat{\beta}_{3,i}x_{3,j}.$$

$x_{1,j}$ to $x_{3,j}$ are dummy variables for the low tax rate, the medium tax rate and the full deduction of income-related expenses, respectively.

[13] The associated questions are:

1. How do you assess your own knowledge regarding tax law? Possible answers are "No knowledge", "Some basic knowledge", and "Good or very good knowledge".
2. How high do you rate the income tax burden in percent of the respective annual income? Respective incomes are: € 10000; € 40000; € 300000; and € 2000000. We compare subjects' answers with the actual tax burden according to the income tax tariff and compute the absolute value of subjects' error for every income category. The 5% of our subjects with the lowest total error about all four income categories are marked with "very good knowledge of tax law".
3. Who prepared your last tax return?Possible answers are: "On my own", "Someone else in the household", "A tax advisor", "Did not file a tax return", "Other".

[14] In this case, a binomial test also confirms that the percentage of lexicographical sequences cannot be the result of pure random selection ($p < 0.01$).

Table 4
Part-worth and relative importance (sample and rational sequence).

	Estimated part-worth		Relative importance	
	Rational	Sample	Rational	Sample
Low tax rate	0.5625	0.6331	0.5625	0.6331
Medium tax rate	0.2812	0.3166		
High tax rate	0.0000	0.0000		
High deduction of income-related expenses	0.4375	0.3669	0.4375	0.3669
Low deduction of income-related expenses	0.0000	0.0000		

The difference between the samples' relative importance of the tax rate (0.6331) and the rational value (0.5625) is significant as a Mann–Whitney-U test shows ($p < 0.01$).

Table 5
Relative importances of tax rate.

Tax rate named first	Tax rate named last
0.6996	0.6021

Both values are compared to the rational value of 0.5625. The differences are significant according to a Mann–Whitney-U test ($p < 0.01$).

The standardized part-worth utilities, as well as the relative importance that result from the sample of working individuals, are shown in Table 4, along with the part-worth utilities and relative importance of a "homo economicus" which serve as the rational benchmark (Hundsdoerfer, Sielaff, Blaufus, Kiesewetter, & Weimann, in press).

As it is shown in Table 4, the relative importance of tax rate changes is 12.5 percentage points higher than the importance of tax base changes. Therefore, tax rates have a stronger impact on the tax burden than tax deductions. However, by evaluating individuals' preferences we only consider the relative importance of the tax rate to be overrated if it exceeds the rational benchmark significantly. Table 4 shows that the part-worth utility of the low tax rate and the relative importance of the attribute nominal tax rate are higher in the sample. Hence, a change in the nominal tax rate results in a greater change of total utility. Whereas for a rational agent we observe a relative importance of 56% for the tax rate, the corresponding relative importance of the subjects is on average 63%. This difference is highly significant (Mann–Whitney U, $p < 0.01$).[15] Hypothesis H2 is, therefore, confirmed: the importance of changes in tax rates is overestimated, and the importance of changes in the tax base is underestimated.

However, it must be noted that the degree of overestimation can vary widely. The relative importance of the tax rate for individuals who conduct rankings lexicographically based primarily on the tax rate can amount to 80%, whereas the relative importance for the tax rate for other types of ranking only amounts to 63%. In addition, 32.5% of the individuals underestimate the relative importance of tax rates. This includes primarily those who rank the stimuli first according to income-related expenses and, therefore, attach a relative importance of only 40% to tax rates.

One should note the relative importance of the tax rate of 63% for the "other sequences" rankings. These individuals significantly overestimate the relative importance of the tax rate. As we pointed out in the last section, one approach to explain the "other sequences" rankings is the anchor heuristic. By thinking of individuals who use this heuristic and choose the tax rate as an anchor, one should expect a relative importance that is below the importance of lexicographic rankings (those who rank first according to the tax rate) but above the importance of the rational ranking. This is exactly what we observe in this case. Moreover, it becomes obvious that these subjects do not randomly rank the stimuli because the relative importance of the attributes differs significantly from 50% (sign test, $p < 0.01$).

3.4.3. Framing effects (Hypothesis H3)

To identify factors that explain the importance of the tax rate, we test whether framing effects increase the misperception of nominal tax rate changes. Therefore, we compute the relative importance for the two respective groups (tax rate mentioned first/tax rate mentioned last). The framing effect can be quantified as the difference in relative importance for the two settings "tax rate mentioned first" and "tax rate mentioned last". The resulting values are shown in Table 5 below.

As it can be seen from Table 5, the importance of the tax rate is overestimated and the importance of the expense deduction is underestimated in both groups. The deviations from the rational conclusion are highly significant (Mann–Whitney U, $p < 0.01$).

One can also see from Table 5 that the overestimation of the importance of the tax rate is significantly larger (Mann–Whitney U, $p < 0.01$) when the subjects are presented the tax rate first. The relative importance of the tax rate is 70% when

[15] In the following, solely the relative importance of the attribute "nominal tax rate" will be examined, as this can be used to derive all further values. The relative importance values add up to one. Hence, the relative importance of the attribute "allowable deduction of income-related expenses" is given by: 1 minus the relative importance of the "nominal tax rate." In addition, the largest standardized part-worth utility of an attribute level always corresponds to the relative importance of this attribute. The standardized part-worth utility of the middle tax rate can be calculated as half of the relative importance of the tax rate.

Table 6
Choice between tax options.

Position	Stimulus	Choices (%)	Tax burden (€)
1	F	55.10	2500
2	B	44.40	2400
3	E	0.50	2800
4	A	0.00	3000
5	C	0.00	3500

Total percentages according to the first choice model.

mentioned first and only 60% when mentioned last. Thus, hypothesis H3 is confirmed. The framing effect amounts to 10 percentage points. As we mentioned in section "Hypotheses" individuals tend to choose the piece of information with which they are first confronted as an anchor or as the primary criterion for the lexicographic heuristic. This explains the large framing effect. The percentage of subjects with a lexicographical ranking with tax rate as first criterion increases from 25% to 48% when the tax rate is mentioned first. In the same way the percentage of subjects with a lexicographical ranking with tax deductible expenses as first criterion increases from 13% to 28% when the tax rate is mentioned last.

3.4.4. Choice between tax options

The subjects are presented with the various alternatives as possible options for tax policy. Since we focus on tax changes only, we implicitly consider other political factors such as environmental policy, economic policy, or health care as constant among the alternatives. In our view, this assumption is particularly appropriate in situations where taxpayers can choose between different tax options as long as the choice does only affect themselves. Basically, in our setting this choice is whether to tax net income at a higher tax rate or gross income at a lower tax rate. This reflects for example the alternatives of smaller corporations in Russia. These corporations can choose between a tax rate of 6% on their sales (gross income) and a tax rate of 15% on their earnings (net income). Besides, there are several other countries as France and Germany that provide similar tax options with respect to dividend income. In these countries taxpayers can choose between the usual progressive tariff applied on their net dividend and a flat tax on their gross dividend.

To examine how taxpayers will choose between such tax options, a further assumption must be made about the relation between the ranking order (preference) of the individual and the individual's actual choice. We consider a deterministic model ("first choice"). Thus, the probability of choosing the most strongly preferred stimulus is one. All other stimuli have a choice probability of zero. The total percentage of a tax option corresponds to the number of subjects with first preference for this respective tax option divided by the total number of subjects. The percentages using the first choice model are shown in Table 6.[16] It shows that the absolute majority of choices are allotted to stimulus F even though this does not have the lowest actual tax burden. If the first choice model accurately describes choice behavior, then politicians could combine increasing tax revenues with a decrease in the perceived tax burden. This can be achieved by introducing tax options into tax law that allow taxpayers to decide between a higher tax rate on their net income and a lower tax rate on their gross income. Obviously, taxpayers could avoid wrong choices by consulting tax advisors. However, according to our data only 22% of the subjects actually consult a tax advisor for their last tax return.

4. Experimental validation

As we pointed out before, conjoint analysis comes with its own strengths and weaknesses. On one hand, we have the advantages of drawing a large (representative) sample, measuring preferences on the individual level and identify decision heuristics. On the other hand, conjoint analysis in practice does not allow for performance based incentives or repeated task execution. Hence, the results we reveal in the previous section are subject to these objections. Therefore, in the present section we validate our survey results by a lab experiment.

4.1. Method and sample

We apply a computer-based lab experiment that is programmed and conducted with the software z-Tree (Fischbacher, 2007). While the basic setting matches the conjoint analysis, there are two main differences that address the limitations of conjoint analysis. First, we introduce performance based compensation and second, we establish a learning environment.

The experiment took place at the European University Viadrina and the Freie Universität Berlin, Germany. Subjects are 56 non-academic university employees. In contrast to a convenient student sample this subject pool has the advantage that employees come with actual tax experience. Most of the non-academic employees at the European University Viadrina and the Freie Universität Berlin, Germany are women. This is also reflected in our sample as 50 out of 56 are females. Subjects are on average 44.1 years of age. This corresponds to the average age of the subjects that take part in conjoint analysis. The average education level is slightly higher in the lab experiment compared to the subjects that participated in the conjoint analysis.

[16] Stimulus D was eliminated from the analysis because it was given as the objectively best alternative. Thus, stimulus D could not be selected by the subjects.

Table 7
Aggregated ranking of tax options.

Stimulus	Tax rate (in %)	Tax deductible expenses (in %)	Tax Burden (in €)	Rational ranking	Sample ranking
1	25	33.3	2250	2	3
2	25	0	2500	4	2
3	30	100	2100	1	1
4	30	33.3	2700	5	5
5	30	0	3000	6	6.5
6	35	100	2450	3	4
7	35	33.3	3150	7	6.5
8	35	0	3500	8	8

Stimuli 5 and 7 receive the same number of choices. Therefore, they are allotted the same rank which is the average of rank 6 and rank 7.

4.2. Research design

According to conjoint analysis subjects have to choose between tax systems that differ only in terms of the nominal tax rate and the allowed deduction of income-related expenses for a given income.

While income is hypothetical in conjoint analysis, subjects that take part in the experiment have to perform a working task to receive income. This eliminates a potential house money effect.[17]

We design a simple decoding task where subjects have to decode letters into numbers according to a given table. Subjects receive 400 points for every correct entered number and are allowed to accomplish up to 25 successful decodings to receive at most 10000 points, which equals the earnings of € 10000 used in the conjoint analysis.

We limit the time subjects have to complete the task to a maximum of 10 min and charge subjects corresponding to the elapsed time. Thus, we are able to generate income-related expenses. There was a fix charge of 3000 points if subjects successfully earn 10000 points in the first five minutes. For every 30 s surpassing the first five minutes subjects are charged with additional 100 points.

To avoid any income effects on the results, it is required that all subjects have the same income available. Pretests showed that almost everybody should be able to complete the decoding task within five minutes. With only one exception all subjects earn 10000 points and bore income-related expenses of 3000 points.[18]

After completing the task subjects are made aware of their working task income. Afterwards, they are presented eight different tax systems which differ only in tax rates (25%, 30%, and 35%) and deductions allowed (100%, 33%, and 0% of income related expenses).[19] Subjects are told to choose the tax option on the basis of which they want their own income to be taxed and that their decision does not affect the taxation of any other participant in the experiment.[20]

While subjects have to rank all tax options in conjoint analysis, they only have to choose one in the experiment. This is due to the fact that a performance based reward is required. To ensure that these rewards control for individuals' preferences saliency is needed. This means subjects must be aware of the relationship between decisions and incentives. Moreover, the relationship between decisions and incentives has to be understandable rather than too complex. Then and only then, subjects can recognize how their payoff is affected by their actions. At first glance, it seems to be pretty easy to find a performance rewarding payoff structure that leads to higher payoff for the rational ranking than for other rankings. However, one has to consider the 120 different ways to rank the stimuli from our conjoint analysis. There are too many issues to adapt a payoff structure that fulfills the demand for saliency. Therefore, subjects only choose their personal best tax option.

As in conjoint analysis the objectively best tax option is excluded, i.e., subjects choose only among the remaining eight options. Subsequent to their choice, subjects receive an extensive feedback. They are not only informed on their pre-tax, taxable, and net income, but even calculations are displayed. So the participants receive all the information needed to find the tax option with the lowest tax burden. The working task as well as the choice of the tax option is repeated three times in exactly the same manner. Screenshots of the experiment are displayed in Figs. 2 to 4 in Appendix B. The instructions are presented in Appendix C.

Subjects are compensated according to their net income. Net income is calculated as earned points less income-related expenses less taxes. Since the first two components are the same for all subjects, the choice of the tax option determines the payoff. To enforce the incentive, there are large differences in compensation between the respective tax options. Choosing the tax option with the lowest tax burden yields a return of € 8.00, while choosing the second best option is awarded with a return of € 5.00. For choosing any other than these two tax options subjects are paid € 2.00. Hence, choosing the tax option

[17] As it is shown in the literature subjects are more involved if their income is not given but earned. This is called the house money effect (Clark, 2002).

[18] Since results were not affected by this subject, we did not exclude her from our study.

[19] In conjoint analysis we used three levels for the attribute "tax rate" and two levels for the attribute "income related expenses". Some studies report that attributes with more levels achieve a higher relative importance (e.g. Wittink, Lakshman, & Reibstein, 1990). To ensure that our results are unaffected by such a "level effect" we decided to validate our conjoint results with a design that contains the same number of levels for both attributes.

[20] In a post experimental manipulation check we asked whether the subjects understand that their choice for a certain tax option does only affect their own experimental income and is independent from the other participants. On a scale from 1 (does not apply) to 7 (does apply) 40 out 56 participants name a 6 or higher. The average (median) answer is 5.63 (6) with a standard deviation of 1.75.

Table 8
Aggregated relative importance (sample and rational benchmark).

	Rational	Sample
Tax rate	0.4615	0.5600
Deduction of income-related expenses	0.5385	0.4400

Table 9
Impact of repeated execution.

	Decision in favor of tax system with lowest tax burden (in%)	
	Yes	No
Before first feedback	32.1	67.9
After first feedback	44.6	55.4

Above the percentages of tax burden minimizing decisions are displayed for the first choice and the following choices in the experiment. The complete overview of the experimental results is given in Table 12 (Appendix B).

with the lowest tax burden thrice leads to the maximum total payment of € 24.00. Since subjects are paid at least € 2.00 for every decision, the minimum total payment amounts up to € 6.00. On average subjects earned a total of € 15.67 (Std. Dev. € 5.86) for an experiment lasting on average 50 min.

The experiment is designed as a validation whether the conjoint analysis results regarding hypothesis H2 are robust with respect to performance based incentives and learning effects. To test whether performance based incentives affect the results from the conjoint analysis, we again use a conjoint analysis to look at the impact of changes in nominal tax rate and tax base.

To test whether the implementation of a learning environment changes the results we use a within-subject-design and apply generalized estimation equation to compare subjects' first choice with their subsequent choices.

4.3. Results

4.3.1. Overestimation of the importance of changes in tax rate (Hypotheses H1 and H2)

In the analysis of subjects' ranking behavior in Section 3, we point out that only 9.4% of our subjects rank all of the stimuli rationally. In the experiment subjects do not have to rank all of them but only have to choose one. Due to this change in our design an individual ranking cannot be obtained. Thus, to conduct a conjoint analysis we aggregate subjects' choices. We obtain the ranking by assigning the stimulus with the highest total percentage of choices the first rank, the stimulus with the second highest total percentage the second rank, and so on. To avoid any impact of learning effects we only consider choices from the first of the three periods. The number of choices for each stimuli is given in Table 12 in Appendix B.

As it can be seen from Table 7, the stimulus with the lowest tax burden is chosen most often. However, the aggregated sample ranking deviates from the rational ranking. In fact, only roughly a third of the subjects (18 out of 56) choose the tax option with the lowest tax burden. The result regarding hypothesis H1 is unaffected by the implementation of performance-based incentives. Even after the implementation of performance-based incentives the majority of the subjects does not choose their tax option based on the actual tax burden.

To test whether performance-based incentives change our results regarding hypothesis H2, we estimate part-worth utilities by Eq. (2) using OLS and compare the results of our aggregated ranking with the rational benchmark. The results are given in Table 8.

As it can be seen from Table 8, the relative importance of the nominal tax rate for the rational benchmark amounts to 46%, while the corresponding relative importance of the sample is 56%. The relative importance of the rate is 10 percentage points higher in the sample than for the homo economicus. As it is for hypothesis H1, the results regarding hypothesis H2 remain unaffected by the implementation of performance-based incentives.

4.3.2. Learning effects

Subjects have to choose their individual tax option thrice. As pointed out before, subjects receive an extensive feedback after each decision. The feedback does not only contain information on subjects' income, but even on calculations. For our analysis, we divide subjects' decisions into two groups. The first group contains the subjects' first decisions, since these choices are carried out before the first feedback. The second group consists of the second and third decisions, respectively. These choices are made after receiving at least one feedback.[21]

In the first period, 18 out of 56 subjects choose the tax option with the lowest tax burden (see Table 9 below). This equals 32%. After receiving feedback on their decisions 50 out of 112 (45%) decisions in the last two runs are in favor of the tax option with the lowest tax burden. The difference in favor of the performance in the last two periods is weakly significant (chi squared test, $p = 0.082$) (see Table 9).

[21] We compared only the first with the last decision, too, but the results remain the same.

Table 10
GEE results: determinants of learning effects.

	Coefficient	Std. error
Period	0.533*	0.309
Education	−0.103	0.476
Tax return self-prepared	0.116	0.461
Age	−0.760	0.481
Constant	0.197	0.420

Dependent variable: rational ranking.
* indicates significance on the 10% level, standard errors are robust.

To test for determinants of learning effects, we use a generalized estimation equation. The parameter estimations and robust standard errors are shown in Table 10. We choose this technique because we are interested in estimation of population average impact. A logistic link function is formulated for the dependent variable "lowest tax burden (yes/no)". As explanatory variables period, age, education and tax knowledge (self-assessment) are used.[22] While we find the demographic variables not to have any impact on the ranking, period has a weak significant positive influence ($p = 0.084$). The likelihood to choose the tax option with the lowest tax burden is 70% higher in the third period compared to the first one.

Hence, the implementation of a learning environment leads to a slight enhancement of the conjoint analysis results. However, even after learning the majority of subjects do not make a rational choice. Most subjects still overrate the nominal tax rate.

5. Discussion

If individuals behave rationally according to traditional economic theory, it makes no difference whether their tax rate or (by equal measure) their tax deductions change. By contrast, if one takes into account the idea that individuals tend to avoid cognitive strain and instead use simplified decision heuristics changes in the tax rate may influence the perceived tax burden more strongly than do changes in the tax base.

To empirically test the hypothesis regarding the use of heuristics, the resulting overestimation of the importance of changes in tax rates, and the framing effect, we conduct both a conjoint analysis and a lab experiment. In line with behavioral studies incorporating the burden of cognitive strain, our results reveal that the majority of the subjects use decision heuristics. The use of heuristics leads to a noticeable overestimation of the relative importance of changes in tax rates and to an underestimation of the importance of changes in the tax base. We show that the overestimation increases considerably when information about the tax rate is mentioned first (framing effect).

Even after the introduction of performance based incentives we observe that subjects still overestimate nominal tax rate changes. However, we find slight evidence for learning effects. After the implementation of a learning environment, the use of heuristics decreases weak significantly, but even after learning the majority of subjects do not make a rational choice. Most subjects still overrate the importance of the nominal tax rate.

Our results indicate that politicians could simultaneously increase tax revenues and reduce the perceived tax burden for taxpayers. Taxpayers favor tax options that apply a lower tax rate on their gross income over a higher tax rate applied on their net income despite the lower actual tax burden of the latter option. This result remains stable in a lab experiment where performance-based incentives and a learning environment are established. Consistent with the identified framing effect, emphasizing the nominal tax rate when presenting tax options can reduce the perceived tax burden solely through the display format.

Our study offers starting points for future research. First, the present study only considers the effect of non-specified income-related expenses. It would be interesting to analyze whether subjects also underestimate the importance of changes in the deductibility of emotionally loaded costs like commuting expenses. Second, we study the behavior of the German working population not the behavior of tax experts. Thus, it would be worth to explore if our results also hold for specialized decision makers in companies. Third, we do not vary complexity. According to current research (e.g., Blaufus & Ortlieb, 2009), subject's performance decreases with increasing complexity. Fourth, we abstract from distribution and equity effects that are usually affected by introducing new tax options. Finally, even though considering changes in tax rate and tax base simultaneously is specific to taxes, it would be interesting to see, whether our results regarding overestimation are of relevance for non-tax frames, too. For example, our results could be relevant for consumer price research. According to our results the value of a relative price discount should be perceived as higher when it is applied at a higher discount rate on a narrow base (e.g. product price excluding extras) compared to a lower discount rate on a broad base (e.g. product price including extras).

[22] One might ask why sex is not used as explanatory variable. Out of the 56 subjects that take part in the experiment only 6 are male. Because of this enormous mismatch sex was excluded of the analysis.

Appendix A

Table 11
Distribution of quota attributes in sample.

Attribute	Value	Frequency	Percent	Percent (Population)
Gender	Male	209	44.8	45.2
	Female	258	55.2	54.8
Age	Under 20	15	3.2	3.5
	20–29	96	20.6	17.5
	30–39	112	24.0	24.0
	40–49	125	26.8	29.6
	50–59	94	20.1	20.5
	60 and older	25	5.4	5.1
Education	University degree	81	17.3	16.0
	University-entrance qualification	80	17.1	14.5
	Secondary school leaving certificate	126	27.0	26.8
	Lower secondary school leaving certificate	136	29.1	31.4
	No school leaving certificate	8	1.7	2.2
	Other	36	7.7	9.1
Monthly net Income	Under € 1000	161	34.5	32.7
	€ 1000–2000	206	44.1	44.8
	€ 2000–3000	65	13.9	14.6
	Above € 3000	30	6.4	8.0
	Not stated	5	1.1	–

Table 12
Distribution of choices in the experiment (in%).

	Tax rate		
	25%	30%	35%
Deductible expenses			
100%	–	32.1 (44.6,44.6)	7.1 (3.6,1.8)
33%	25.0 (25.0,26.8)	5.4 (1.8,7.1)	1.8 (3.6,3.6)
0%	26.8 (19.6,8.9)	1.8 (1.8,5.4)	0.0 (0.0,1.8)

Percentage of choices that the respective tax system received in the first (second, third) period. The tax system with the lowest tax burden (tax rate 25%, deductible expenses 100%, top left) was given. From top to bottom and from left to right tax burdens decrease. The worst tax system is at bottom right (tax rate 35%, deductible expenses 0%).

Appendix B

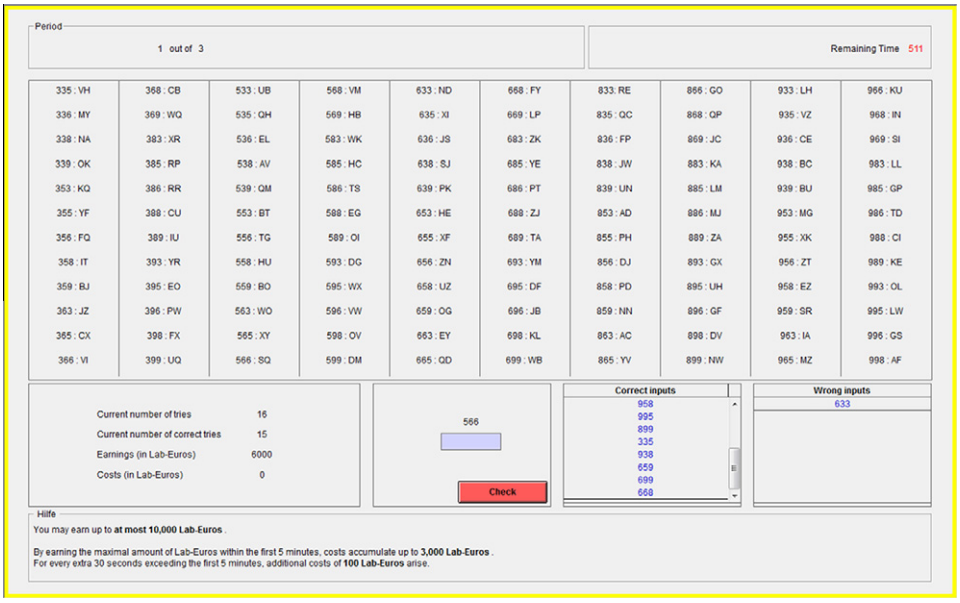

Fig. 2. Screenshot from the experiment – working task.

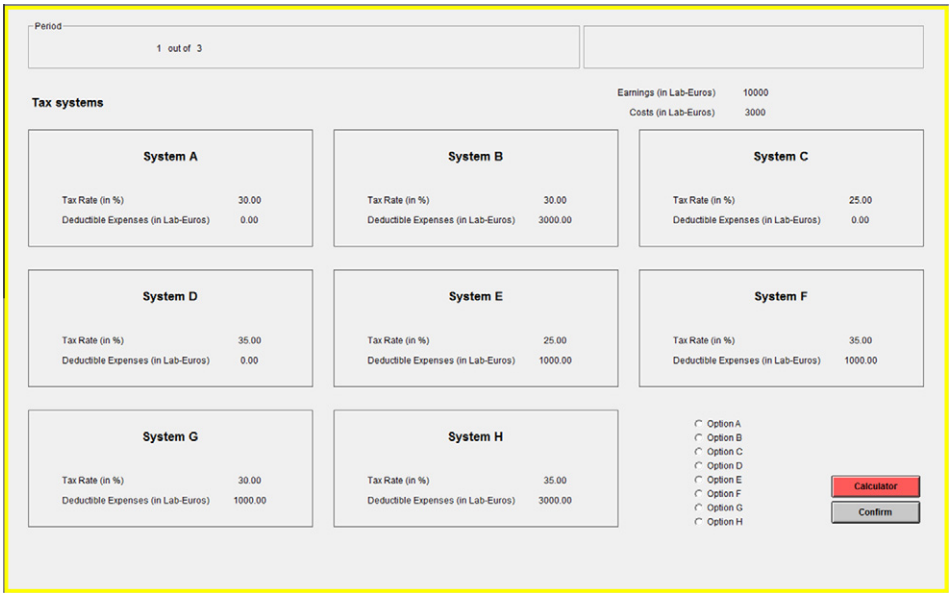

Fig. 3. Screenshot from the experiment – choosing individual tax system.

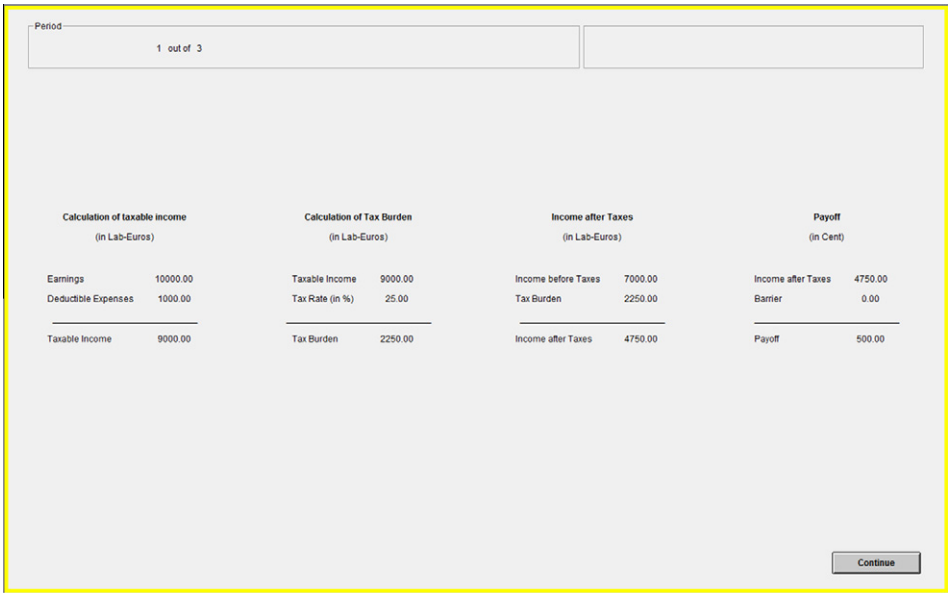

Fig. 4. Screenshot from the experiment – feedback.

Appendix C

C.1. Experimental instructions, part 1 (presented before the working task)

Thank you very much for your willingness to participate in this experiment (see Figs. 3 and 4).

By these instructions the procedure of the experiment will be explained. Please read the instructions carefully. Take as much time as you need to read and understand them.

The experiment consists of three periods. To make you familiar with the software and your task, there will be a trial period that has no impact on the outcome of the experiment. Please take sufficient time to grasp the information presented on the screen. The trial period lasts 5 min.

Every new period won't be started until all participants are ready. This can lead to short waiting times. In the meantime you are presented with current headlines.

During the entire experiment laboratory-Euros are used as currency unit. The exchange rate of laboratory-Euros to actual Euros is explained in detail later.

Your task is to convert three-digit numbers into two letters. Every number-letter combination is given in a list. For example "356: CF".

After entering two letters and a subsequent click on the "check" button, the program compares your input with the data given in the list. Correct inputs are then shown in the table "correct inputs", while incorrectly entered letters are shown in the table "incorrect inputs". Please use only upper case letters.

For every correct input, you are credited with 400 laboratory-Euros. In every period you can reach up to 25 correct inputs, i.e. you can earn up to 10000 laboratory-Euros. To reach the maximum of 25 correct entries, you get at most 10 min of time. If you reach the maximum number of correct entries before time elapses, the task ends automatically. Otherwise it ends after 10 min.

Costs arise while you use the software. The amount of these costs depends on how long you need to complete the working task. Within the first five minutes, you are charged with a fixed amount of 3000 laboratory-Euros. If you need longer than five minutes, you are charged with an additional 100 laboratory-Euros for every additionally 30 s.

The remaining time (in seconds) is shown at the top right corner on the screen.

Your income is subject to an income tax (earnings–costs). You will have to choose one out of eight tax systems that differ, exclusively, in the nominal tax rate (25%, 30%, 35%) and the amount of tax deductible expenses (full deduction allowed, partial deduction allowed, no deduction allowed). You choose your tax system by your own. Your decision does not affect the taxation of the other participants.

Your payment in the experiment depends on your net income:

Earnings–Costs–Taxes (in laboratory-Euros)

At the end of each period, your actual payment (in Cent) is displayed. For every 100 laboratory-Euros that exceed the threshold of 4500 laboratory-Euros, you receive a payoff of 200 Cents. If your net income does not exceed the threshold by at least 100 laboratory-Euros, you will receive the minimum payment of 200 Cents.

If you have any questions regarding the course of the experiment, please do not hesitate to contact the experimental supervisor.

By clicking the "Continue" button, you proceed to the first stage of the experiment. It will start as soon as all participants are ready.

C.2. Experimental instructions, part 2 (presented after the working tax)

Your income is subject to an income tax. Therefore, you have to choose one out of eight taxation options. You choose your tax system by your own. Your decision does not affect the taxation of the other participants. Your decision is in effect only for the current period. You may choose other tax systems at the end of the following periods. The eight taxation options differ in two ways.

First, the various options differ in the nominal tax rate at which your income is taxed.

Second, a distinction is made on the amount of deduction allowance. If the deduction allowed equals zero, you have to pay taxes on your earnings without deducting any expenses, i.e. on 10,000 laboratory-Euros. If deduction allowance is not limited, you just have to pay taxes on your earnings less your actual costs, i.e. on 10000 laboratory-Euros − 3000 laboratory-Euros = 7000 laboratory-Euros. If deduction allowance is limited to one third of your actual costs, you have to tax your income less the one third of your actual costs, i.e. on 10000 laboratory-Euros − 1000 laboratory-Euros = 9000 laboratory-Euros.

A calculator is available if needed.

An example of a possible taxation option is given below. The tax rate equals 25%, deductions are allowed up to 3000 laboratory-Euros. Hence, you have to pay taxes in the amount of 1750 laboratory-Euros.

References

Allison, P. D., & Christakis, N. A. (1994). Logit models for sets of ranked items. *Sociological Methodology, 24*, 199–228.

Blaufus, K., & Ortlieb, R. (2009). Is simple better? A conjoint analysis of the effects of tax complexity on employee preferences concerning company pension plans. *Schmalenbach Business Review, 61*, 60–83.

Boylan, S. J., & Frischmann, P. J. (2006). Experimental evidence on the role of tax complexity in investment decisions. *Journal of the American Taxation Association, 28*, 69–88.

Brandstaetter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113*, 409–432.

Carmone, F. J., & Green, P. E. (1981). Model misspecification in multiattribute parameter estimation. *Journal of Marketing Research, 18*, 87–93.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *The American Economic Review, 99*, 1145–1177.

Chrzan, K. (1994). Three kinds of order effects in choice-based conjoint analysis. *Marketing Letters, 5*, 165–172.

Clark, J. (2002). House money effects in public good games. *Experimental Economics, 5*, 223–231.

Elffers, H., & Hessing, D. J. (1997). Influencing the prospects of tax evasion. *Journal of Economic Psychology, 18*, 289–304.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17*, 311–318.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*, 171–178.

Genser, B., & Reutter, A. (2007). Fiscal policy in action: Moving towards dual income taxation in Europe. *FinanzArchiv/Public Finance Analysis, 63*, 436–456.

Green, P. E., & Devita, M. T. (1974). A complementarity model of consumer utility for item collections. *Journal of Consumer Research, 1*, 56–67.
Green, P. E., & Krieger, A. M. (1985). Models and heuristics for product line selection. *Marketing Science, 4*, 1–19.
Green, P. E., & Krieger, A. M. (1993). Conjoint analysis with product positioning applications. In J. Eliashberg & G. L. Lilien (Eds.), *Handbook of Operations Research Series. Marketing* (Vol. 5, pp. 467–515). New York, NY: North-Holland.
Green, P. E., & Srinivasan, V. S. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5*, 103–122.
Green, P. E., Tull, D. S., & Albaum, G. (1988). *Research for marketing decisions* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2008). *Multivariate data analysis* (7th ed.). New York, NY: Pearson Education.
Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief adjustment model. *Cognitive Psychology, 24*, 1–55.
Hundsdoerfer, J., Sielaff, C., Blaufus, K., Kiesewetter, D., & Weimann, J. (in press). The influence of tax labeling and tax earmarking on the willingness to contribute – A conjoint analysis. Schmalenbach Business Review.
Hundsdoerfer, J., & Sichtmann, C. (2009). The importance of taxes in entrepreneurial decisions: An analysis of practicing physicians' behavior. *Review of Managerial Science, 3*, 19–40.
Kirchler, E. (2007). *The economic psychology of tax behaviour*. Cambridge: Cambridge University Press.
Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics, 89*, 2093–2119.
Krishna, A., & Slemrod, J. (2003). Behavioral public finance: Tax design as price presentation. *International Tax and Public Finance, 10*, 189–203.
Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1*, 1–27.
McCaffery, E. J., & Baron, J. (2003). The Humpty Dumpty blues: Disaggregation bias in the evaluation of tax systems. *Organizational Behavior and Human Decision Processes, 91*, 230–242.
McCaffery, E. J., & Baron, J. (2004). Framing and taxation: Evaluation of tax policies involving household composition. *Journal of Economic Psychology, 25*, 679–705.
Milliron, V. C., & Toy, D. R. (1988). Tax compliance: An investigation of key features. *Journal of the American Taxation Association, 9*, 84–104.
Moran, S., & Meyer, J. (2006). Using context effects to increase a leader's advantage: What set of alternatives should be included in the comparison set? *International Journal of Research in Marketing, 23*, 141–154.
Morwitz, V. G., Greenleaf, E. A., & Johnson, E. J. (1998). Divide and prosper: Consumers' reactions to partitioned prices. *Journal of Marketing Research, 35*, 453–463.
OECD (2010). Choosing a broad base – low rate approach to taxation. *OECD Tax Policy Studies*, No. 19, OECD Publishing.
O'Neil, C. J. (1982). The targeted jobs credit: An evaluation of its impact on the employment decision process. *Journal of the American Taxation Association, 3*, 15–22.
Pitt, M. M., & Slemrod, J. (1989). The compliance cost of itemizing deductions: Evidence from individual tax returns. *The American Economic Review, 79*, 1224–1232.
Rupert, T. J., Single, L. E., & Wright, A. M. (2003). The impact of floors and phaseouts on taxpayers' decisions and understanding of marginal tax rates. *Journal of the American Taxation Association, 25*, 72–86.
Rupert, T. J., & Wright, A. M. (1998). The use of marginal tax rates in decision making: The impact of tax rate visibility. *Journal of the American Taxation Association, 20*, 83–99.
Schmölders, G. (1960). *Das Irrationale in der öffentlichen Finanzwirtschaft*. Hamburg: Rowohlt.
Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1–20.
Slemrod, J. (1989). The return to tax simplification: An econometric analysis. *Public Finance Quarterly, 17*, 3–28.
Tax Foundation (2005). Tax Foundation Annual Survey of US Attitudes on Tax and Wealth.
Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103*, 299–314.
Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458.
Willner, J., & Granqvist, L. (2002). The impact on efficiency and distribution of a base-broadening and rate-reducing tax reform. *International Tax and Public Finance, 9*, 273–294.
Wittink, D. R., Lakshman, K., & Reibstein, D. J. (1990). The effect of differences in the number of attribute levels on conjoint results. *Marketing Letters, 1*, 113–123.
Yadav, M. S. (1994). How buyers evaluate product bundles: A model of anchoring and adjustment. *Journal of Consumer Research, 21*, 342–353.