ONLINE APPENDIX

for "Eliciting People's First-Order Concerns: Text Analysis of Open-Ended Survey Questions"

by Beatrice Ferrario and Stefanie Stantcheva

Contents

List of Figures

OA-1 Representativeness of the Sample

For more details on the surveys, sample, data collection, and data analysis, see Stantcheva (2021). The full questionnaire is in Section OA-2 of the Online Appendix of that paper.

TABLE OA-1: SAMPLE CHARACTERISTICS

	US Population	Income Tax Survey	Estate Tax Survey
Male	0.49	0.48	0.46
18-29 years old	0.24	0.23	0.22
30-39 years old	0.20	0.20	0.20
40-49 years old	0.18	0.19	0.19
50-59 years old	0.19	0.21	0.19
60-69 years old	0.19	0.18	0.19
$0-$19,999	0.13	0.15	0.16
$20,000-$39,999	0.16	0.19	0.19
$40,000-$69,999	0.21	0.23	0.24
$70,000-$109,999	0.20	0.19	0.19
$110,000+	0.31	0.24	0.20
Four-year college degree or more	0.34	0.48	0.46
High-school graduate or less	0.38	0.19	0.19
Employed	0.70	0.63	0.62
Unemployed	0.03	0.07	0.06
Self-employed	0.07	0.07	0.06
Married	0.53	0.55	0.53
White	0.61	0.76	0.76
Black/African-American	0.12	0.06	0.06
Hispanic/Latino	0.18	0.06	0.07
Asian/Asian-American	0.06	0.07	0.07
Democrat	0.30	0.34	0.35
Republican	0.26	0.31	0.30
Independent	0.42	0.33	0.33
Voted for Clinton in the 2016 Presidential Election	0.48	0.44	0.44
Voted for Trump in the 2016 Presidential Election	0.46	0.44	0.44
Sample size		2784	2360

Notes: This table displays statistics for the overall US population (column 1) and compares it to the characteristics of the samples of the income tax and estate tax surveys (columns 2 and 3). National statistics on gender, age, income brackets, race, education, marital status, and employment status are from the IPUMS-CPS-ASEC dataset for March 2019 (Flood et al., 2020). National statistics on party affiliation for March 2019 are from Gallup (2019). Statistics on 2016 Presidential Election Results are from Leip (2019). See Stantcheva (2021) for details on how the summary statistics are constructed.

OA-2   Open-ended Questions Wording

This section provides the open-ended questions that were asked to respondents in the surveys. See Stantcheva (2021) for the full questionnaires.

We now want to ask you a few broader questions. Please use the text boxes below and write as much as you feel like. Your opinion and thoughts are important to us! There is no right or wrong answer.

Income Taxation Survey

1. *When you think about federal personal income taxation and whether the U.S. should have higher or lower federal personal income taxes, what are the main considerations that come to your mind?*

2. *What would be a "good" federal tax system in your view? What would be the goal of a good tax system?*

3. *What do you think are the issues with or shortcomings of the U.S. federal income tax system?*

4. *Which important aspects of the U.S. federal income tax system would you say are not discussed enough in the current policy debate?*

5. *What do you think would be the effects on the U.S. economy if the federal personal income taxes were increased?*

6. *Which groups of people do you think would gain if federal personal income taxes on high earners were increased?*

7. *Which groups of people do you think would lose if federal personal income taxes on high earners were increased?*

Estate Tax Survey

1. *The federal estate tax is a tax imposed on the transfer of wealth from a deceased person to his or her heirs. When you think about the federal estate tax and whether the U.S. should have a higher or a lower federal estate tax, what are the main considerations that come to your mind?*

2. *In your view, what would be a "good" federal estate tax that you would be satisfied with? What would be the goal of a good estate tax system?*

3. *What do you think are the shortcomings of the U.S. federal estate tax?*

4. *What do you think would be the effects on the U.S. economy if the federal estate tax were increased?*

5. *Which groups of people do you think would gain if the federal estate tax were increased?*

6. *Which groups of people do you think would lose if the federal estate tax were increased?*

OA-3   Political Affiliation Categories

This section reports the questions that we used to construct the political categories of Figure 3, the definition of political categories, summary statistics of the political leanings in the sample.

1. Did you vote in the last presidential election?

   Yes; No

2. (If "Yes" to 1) In the last presidential election, supported:

   Hillary Clinton; Donald Trump; Jill Stein; Gary Johnson; Other

   (If "No" to 1) Even if you did NOT vote, please indicate the candidate that you were most likely to have voted for or who represents your views more closely.

   Hillary Clinton; Donald Trump; Jill Stein; Gary Johnson; Other

3. On economic policy matters, where do you see yourself on the liberal/conservative spectrum? *Very liberal; Liberal; Moderate; Conservative; Very conservative*

Definition of Variables:

Trump Conservative: respondent supported Donald Trump in the 2016 presidential elections and has "conservative" or "very conservative" views on economic policy matters.

Trump Moderate: respondent supported Donald Trump in the 2016 presidential elections and has "moderate" views on economic policy matters.

Clinton Moderate: respondent supported Hillary Clinton in the 2016 presidential elections and has "moderate" views on economic policy matters.

Clinton Liberal: respondent supported Hillary Clinton in the 2016 presidential elections and has "liberal" or "very liberal" views on economic policy matters.

TABLE OA-2: POLITICAL AFFILIATION AND VIEWS ON ECONOMIC POLICY MATTERS IN THE INCOME TAX SURVEY

	D. Trump	H. Clinton	G. Johnson	J. Stein	Other	
Liberal	7%	54%	12%	53%	15%	29%
Moderate	32%	38%	62%	37%	57%	39%
Conservative	61%	7%	26%	10%	28%	32%
Total	100%	100%	100%	100%	100%	100%

Notes: The table shows the distribution of the respondents' views on economic policy matters by political candidate supported in the 2016 presidential elections in the income tax survey sample.

TABLE OA-3: POLITICAL AFFILIATION AND VIEWS ON ECONOMIC POLICY MATTERS IN THE ESTATE TAX SURVEY

	D. Trump	H. Clinton	G. Johnson	J. Stein	Other	
Liberal	8%	55%	16%	43%	19%	30%
Moderate	31%	38%	54%	45%	58%	38%
Conservative	61%	7%	31%	12%	23%	32%
Total	100%	100%	100%	100%	100%	100%

Notes: The table shows the distribution of the respondents' views on economic policy matters by political candidate supported in the 2016 presidential elections in the estate tax survey sample.

OA-4 Preprocessing Steps

This section summarizes the preprocessing steps performed to carry out the text analysis. The first step, whose goal is to reduce the number of distinct elements, is common to all three methods, whereas the subsequent steps are specific to each of the methods.

Given an answer d_i :

1. Parse d_i : lower-case every word, remove punctuation, spaces in excess, numbers, misspelled words, very common words that carry no intrinsic meaning ("stopwords") such as "and," "the," "each," "then".

OA-4.1 Word clouds and Keyness Graphs

2. Lemmatizing remaining words, i.e. grouping together the inflected forms of a word so they can be analysed as a single item.

   - Use the English lemmatization list available in the lexicon package.
   - e.g., : "policies" becomes policy, "were" becomes "be" $\rightarrow$ reduces number of distinct textual elements

$\rightarrow$ Output: $(\hat{d}_i)$

3. Remove words coming from the question as well as extra words related to the structure of answer.

   - e.g., for the question, "what are your main considerations about income tax system?" : remove "main," "considerations," "income," from the answers, as well as "think," "believe," "should," etc.

4. Transform $\hat{d}_i$ into numerical vector c_i in which each element is a 2-gram, i.e. a 2-component expression of two words which were separated by 0 or 1 word in the original text. Group together 2-grams which correspond to the same inverted two words. Manually remove 2-grams which have no grammatical foundation and duplicated 2-grams (e.g., "tax tax").

   - e.g., take $d_1 =$ "We should tax the wealthy more and tax the poor less." After steps 1-2-3 becomes: $\hat{d}_i =$ "tax wealthy more tax poor less". After step 4 becomes: ['tax wealthy' = 1, 'tax more' = 2, 'wealthy more' = 0 (because it is not grammatically coherent), 'tax poor' = 1, 'poor people' = 0, ...]

OA-4.2 Topic Analysis

2. Reduce remaining words to common root (stemming).

   - Use Snowball stemming algorithm
   - *policies* and *policy* become *polic* $\rightarrow$ reduces number of distinct textual elements.

$\rightarrow$ Output: $(\hat{d}_i)$

3. Transform $\hat{d}_i$ into numerical vector $\mathbf{c}_i$ in which each element is the count of a distinct word.

   - e.g., take $d_1 =$ *"We should tax the wealthy more and the poor less."*
     * After Steps 1-2 becomes: $\hat{d}_1 =$ *"tax wealthi more poor less"*
     * After Step 3 becomes: ['tax' = 1, 'wealthi' = 1, 'more' = 1, 'poor' = 1, 'less' = 1 ...]

4. Generate topic dummy variables equal to 1 when an element of $\mathbf{c}_i$ matches a custom-made topic dictionary.

OA-5   Methods

OA-5.1   Word clouds

The simplest text analysis consists in plotting "word clouds," which are based on a frequency analysis. The frequency of each identified n-gram is computed and the feature label is plotted proportional to this frequency. The word clouds and keyness figures are based on the package by Benoit et al. (2018). The transparency and font size are proportional to the frequency of each group of words relative to the total.

Pros and Cons. Word clouds are a key tool for visualizing responses to open-ended survey questions. The font size of a word group is displayed proportionally to its frequency, yielding an insightful starting point for the analysis of text (Heimerl et al., 2014).

By clearly presenting which themes seem to be common among people's answers, word clouds can equip researchers with a first intuition for topics that could be fruitful to explore further. This quick assessment of answers is particularly useful when analyzing large text corpora in the age of Big Data, or in our case, a large number of open-ended responses to survey questions.

This approach, however, comes with certain drawbacks. For instance, word clouds do not account for synonyms, i.e., they will fail to link only slightly different words to the same underlying idea. As a result, themes that can be described in various ways will be less prevalent in the word cloud, while the importance of topics that are clearly associated with certain "buzzwords" might be over-represented. Beyond that, word clouds take the individual words out of context, potentially leading to a loss of interpretability or meaning. Overall, word clouds are a useful way to visualize and explore textual data. It would be premature, however, to draw any final conclusions about the importance of topics simply based on the font size of occurring words. Instead, the themes communicated in word clouds should be seen as a first step in text analysis and used as cautiously-interpreted guidance for further analysis.

OA-5.2   Keyness and keyword graphs

Related Literature. We draw on an in-depth overview by Gabrielatos (2018), which describes the history of keyness analysis and provides a critical review of its applications. In addition, Stubbs (2010) digs into the notion of "keywords," which parallels the idea of keyness.

Among the first to conduct keyness analysis is Leech and Fallon (1992), who study the drivers of differences between cultures by comparing corpora in American and British English. Today, keyness analysis is understood primarily as a comparison of frequencies of words in the so-called *study* and *reference* groups. This approach can be used to elicit degrees of similarity or difference between the studied groups (Gentzkow and Shapiro, 2010; Taylor, 2013; Partington, 2014).

The keyness graphs in this paper are based on a relative frequency analysis that compares the use of n-grams identified in the corpus between two groups (a reference and a target group). To establish the keyness score of a given item, we establish its frequency relative to the number of words in the group corpus (i.e., the number of words in all the answers of a given group). Those frequencies are cross-tabulated and a χ^2 test is conducted to test the independence between the two groups' distribution.

Consider a given n-gram i. Let j be the group index, with $j = 0$ for the reference group and $j = 1$ for the target group. Let $A_{i,j}$ be the observed number of occurrences of the n-gram i in group j and $A_{-i,j}$ the observed number of occurrences of all other n-grams (except the one we consider) in this group. Let R_i be the total number of occurrences of n-gram i in both groups, C_j be the number of occurrences of all n-grams in group j, and N the overall number of occurrences of n-grams in both groups.

To perform the test, we first establish the theoretical distribution under the independence hypothesis by computing $E_{i,j}$, the expected frequency of a given n-gram i in group j :

$$E_{i,j} = \frac{R_i \times C_j}{N}$$

and the expected frequency of all other n-grams in group j:

$$E_{-i,j} = \frac{R_{-i} \times C_j}{N}$$

The χ^2 test statistic is:

$$\chi^2 = (-1)^{\mathbb{1}\{E_{-i,1} > A_{-i,1}\}} \sum_{k \in \{-i,i\}} \sum_{j=0}^{1} \frac{(A_{k,j} - E_{k,j})^2}{E_{k,j}}$$

We compare this statistic to the distribution of a χ^2 distribution law with one degree of freedom (i.e., number of groups -1). A given n-gram is significant when the independence hypothesis is rejected and the n-gram is considered a "keyword." A negative χ^2 indicates that the word is significantly more frequent in the reference group. In absolute value terms, the null hypothesis is rejected at the 10% level when $|\chi^2| > 2.71$ (*), at the 5% level when $|\chi^2| > 3.84$ (**) and at the 1% level when $|\chi^2| > 6.63$ (***).

Pros and Cons. Keyness analysis is particularly useful when the goal is to compare text corpora associated with different groups. It is based on a relative frequency analysis that compares the use of words between two groups, for example, Democrats and Republicans in our study.

This simple method allows us to compute a keyness score that measures how characteristic an n-gram is for a given group. Based on this, keyness analysis can be used to establish differences (similarities) between groups. It also allows the researcher to extract notions and attitudes that are at the heart of a given group's text corpus. Thus, this method can, for instance, help elevate our understanding of what drives polarization within the political spectrum.

One common critique of keyness analysis is that it is based purely on statistical significance and does not take into account the effect size (Gabrielatos and Marchi, 2011; Kilgarriff, 2001). In particular, we do not learn about the size of a frequency difference (Gabrielatos, 2018).

Overall, keyness analysis serves as a valuable tool when the goal is to explore potential heterogeneity, e.g., in attitudes across groups. Yet, results need to be interpreted with caution, given the limitations of the method (Gabrielatos, 2018). Indeed, while statistical significance is a useful metric, its utility is limited to express the accuracy of a given frequency difference. Furthermore, keyness is not an intrinsic attribute of words but depends on several subjective decisions regarding, for instance, the size of the linguistic units that are the focus of the analysis, the target and reference text corpora, and the statistical significance thresholds.

OA-5.3 Topic analysis

Related Literature. Topic analysis was originally developed as a tool for text mining with the goal to impose structure on the rapidly expanding textual data on the internet (Baeza-Yates et al. (1999)). Salton and McGill (1983) introduced a now widely-applied method that compares a normalized word frequency count, within one document, to an inverse document frequency count within a text corpus containing all considered documents. To reduce the dimensionality of vast text corpora, Deerwester et al. (1990) introduced latent semantic indexing (LSI), a method later advanced by Hofmann (1999) into the probabilistic LSI (pLSI). Another approach to topic modeling is the latent Dirichlet allocation (LDA), developed by Blei et al. (2003). Further developments include dynamic topic models (Blei and Lafferty (2006)) and supervised topic models (Blei and Jon (2007)). A more recent advancement was driven by Roberts et al. (2013) and Roberts et al. (2014), who introduce structural topic models and apply the method to open-ended survey responses.

Topic analysis is a versatile tool that can be applied to a wide range of bodies of text. Like keyness analysis, it allows to study heterogeneity in topics use across groups, but its advantage is that we can also flag topics for which the frequency of use is similar across the different groups.(Roberts et al. (2014)). For more comprehensive reviews of the limitations of topic analysis and potential solutions, see Tang et al. (2014) and Agrawal et al. (2018).

Semi-supervised approaches to topic modeling allow the researcher to suggest a set of words around which the algorithm attempts to build topics. Including additional information, e.g., from external data sources,

in the topic modeling process can significantly improve classification results. Banerjee et al. (2007) and Schönhofen (2009) draw on the titles and categories of Wikipedia articles, as an external data source, to improve the clustering accuracy of their documents.

To improve the interpretability of topics, Lu et al. (2011) propose an algorithm that is closely related to Latent Dirichlet Allocation (LDA). Another approach, referred to as "Newsmap" and based on naive Bayes classifiers, was introduced by Watanabe (2018b) and implemented in Watanabe (2018a). Both the seeded-LDA and the Newsmap models were employed in Watanabe and Zhou (2020). Finally, Gallagher et al. (2017) develops a novel approach that does not require assumptions about the underlying data generation process: the "Correlation Explanation" (CorEx) Topic Model.

Semi-supervised algorithms reduces the amount of hand-coding necessary by the researcher and are by nature, data driven. Nevertheless, there are always decisions to be made. For instance, one needs to chose how fine-grained or coarse to make the topics, whether to use the same keywords across time (e.g., newspaper articles across decades) and across people (e.g., groups that may use different language), or rather adapt them.

For open-ended survey questions, it can make sense to pick topics manually, because the answers are shorter than most of the texts that topic analysis has been used for. It is of course then possible to cross-validate the results using less supervised methods. When topics are chosen with a more hands-on approach, multiple hypothesis testing may become a concern Roberts et al. (2014), but could possibly be alleviated thanks to a pre-analysis plan. In practice, we use a manual approach.

Our approach. As a first preliminary step, we extract the document-term matrix (DTM). The DTM is a mathematical matrix that describes the frequency of terms that occur in a given set of documents (in our case, a document coincides with the answer of a given respondent to a given question). The rows and columns correspond to documents and terms, respectively. Hence, the element (x,y) in the DTM corresponds to the frequency of the term y in question answer x.

Topics are then defined by sets of keywords that aim to capture the particular aspects of the policy with which the respondent is concerned when answering to the question. We fix these topics based on the distribution of words that we see in the answers (i.e., in the DTM) and on our understanding of the issues studied. It is critically important to read many sample answers in order to better understand how respondents are interpreting the question and how they are using specific terms. The topic indicator variable is equal to one if the document contains at least one of the keywords that define the topic. Documents can thus contain more than one topic if the respondent used keywords belonging to different topics.

OA-6   Topic Keywords

In this section, we provide the full list of keywords used to define the topics presented in Figure 2 and in Figures OA-8 to OA-13 of this Online Appendix.

TABLE OA-4: KEYWORDS DEFINING THE INCOME TAX TOPICS

Distribution:	*Middle class; working class; low income; wealthy; millionaire; rich; billionaire; corporations & pay/tax*
Fairness:	*Fair; unfair*
Government spending:	*Government spending & high; government spending & cut; deficit; debt; government & waste; balance & budget; government & budget; government & control & spend*
Social insurance:	*Social services; governmental services; governmental program & fund; governmental program & cover; help & poor; pay & poor; social program; poor work; live & paycheck; provide & family*
Efficiency:	*Hurt & economy; work hard; work less; work more; create & job; depress; trickle down; negative/detrimental/destroy/damage & economy; competition; innovation; create & business; boost & economy; discourage; spend less*
Flat tax:	*Flat tax*
Loopholes:	*Loopholes; lawyer; account; tax evasion; evade; avoid taxes*
Public goods:	*Infrastucture; education; healthcare*
Don't know:	*Not know; knowledgeable enough; idk; not sure; know enough; unsure*

TABLE OA-5: KEYWORDS DEFINING THE ESTATE TAX TOPICS

Distribution:	*Middle class; working class; low income; wealthy; millionaire; rich; billionaire; corporations & pay/tax; poor; inequality*
Fairness:	*Fair; unfair*
Government spending:	*Government spending & high; government spending & cut; deficit; debt; government & waste; balance & budget; government & budget; government & control & spend*
Efficiency:	*Hurt economy; work hard; flat*
Loopholes:	*Loopholes; lawyer; account; tax evasion; evade; avoid taxes*
Double tax:	*Already taxed/paid; twice & tax/pay*
Grieve:	*Grieve; bury; funeral*
Public goods:	*Infrastructure; education; health care*
Don't know:	*not know; knowledgeable enough; idk; not sure; know enough; unsure*

OA-7 Sample of Answers to the Open-ended Question "What are your Main Considerations about the Policy?"

OA-7.1 Income Taxation

Distribution: *"That the rich and wealthy do not pay their fair share of taxes."*
"Everyone, including the rich and corporations should pay their fair share."
"I would want working class and middle class people to get tax cuts and I'd be willing to pay more in taxes for that to happen."

Fairness: *"I have trouble with the concept of tax brackets that punish an individual for being successful."*
"I believe Everyone should be taxed fairly and the most wealthy should not escape carrying their weight."

Government Spending: *"Current tax rates being raised are a result of government mismanagement of funds and over spending without appropriate oversight. Taxes really can't effectively be lowered until government spending is properly controlled."*
"I am okay with raising personal income tax to reduce deficit but not for entitlement programs."

Social safety net: *"What are the taxes going towards? I strongly believe in funding going towards education and infrastructure."*
"Cut government spending on social welfare programs for lower taxes and privatize most government services for lower taxes e.g., mail, law enforcement, parks, schools..."

Effiency: *"I am concerned about the push to raise taxes on persons with higher incomes. I do believe in trickle down economics and that government should pretty much keep their hands off."*
"I want the U.S. to be competitive for businesses, but also know there needs to be an appropriate amount of money to fund the government necessities."

Flat Tax: *"We need a flat tax. Tax forms are complex."*
"I think tax Rates are not fairly representative for most taxpayers. I support a flat tax rate for all except the totally disabled and indigent."

Loopholes: *"I think the more you make, the more you should pay. We need to close the loopholes that are there to make sure that those who make more actually pay more."*
"The wealthy oligarchs who own this country will never allow their taxes to be raised, or will hire tax lawyers to get out of paying them, so any raise in federal personal income taxes will fall on the middle class."

Don't know: *"I don't know much about this topic."*

OA-7.2   Estate Tax

Distribution: *"It can help keep the ultra wealthy accountable for their wealth."*
"Passing wealth from one generation to the next contributes to wealth inequality. Federal estate tax should be much higher."

Fairness: *"I don't think there should be a federal estate tax because it's kind of unfair to have to pay taxes on money that already belongs to your family and has most likely had taxes paid on it already."*

Government spending: *"I believe in smaller government, so all taxes should be lower. I actually think we should have a flat tax for income - period. Then estate taxes wouldn't even be an issue."*

Public goods: *"I would like higher taxes to pay for more domestic spending such as education, healthcare, etc."*

Efficiency: *"Lower taxes mean I have more disposable income to spend therefore more products can be mad and more jobs created. I feel it is wrong to penalize people for increased wealth."*

Loopholes: *"The wealthy don't ususally pay these taxes, they find a loophole. Why should my children have to pay taxes on things I've already paid taxes on during my lifetime?"*

Double taxation: *"I think it is ridiculous, you pay taxes twice."*

Grief: *"I don't think we should have one at all. You're taxing a family member for the death of their loved one? That's messed up."*

OA-8 Additional Figures

FIGURE OA-1: WORD CLOUDS FOR THE INCOME TAX

(A) WHAT ARE YOUR MAIN CONSIDERATIONS ABOUT THE INCOME TAX?

(B) WHAT WOULD BE THE GOAL OF A GOOD INCOME TAX SYSTEM?

(C) WHAT ARE THE SHORTCOMINGS OF THE INCOME TAX SYSTEM?
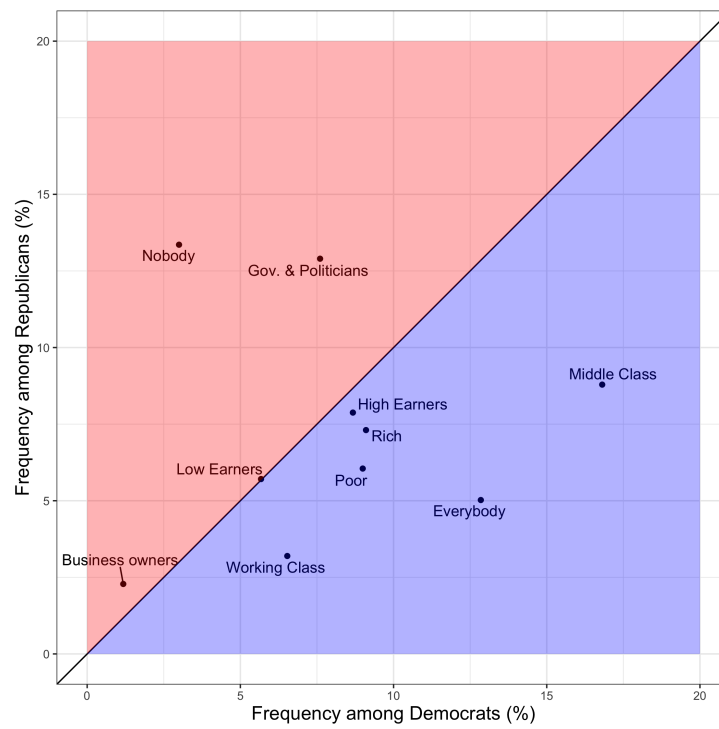
(D) Which important aspects of the U.S. federal income tax system would you say are not discussed enough in the current policy debate?

(E) What do you think would be the effects on the U.S. economy if the federal personal income taxes were increased?

Notes: Word clouds based answers to open-ended questions stated in the caption. See Section OA-5.1.

Notes: The figure reports raw frequencies among Democrat and Republican respondents of the groups mentioned when answering to the question "Which groups of people do you think would gain if federal personal income taxes on high earners were increased?"

FIGURE OA-3: KEYWORDS OF INCOME TAXATION BY GROUPS

(A) BY POLITICAL AFFILIATION

(B) BY AGE GROUP
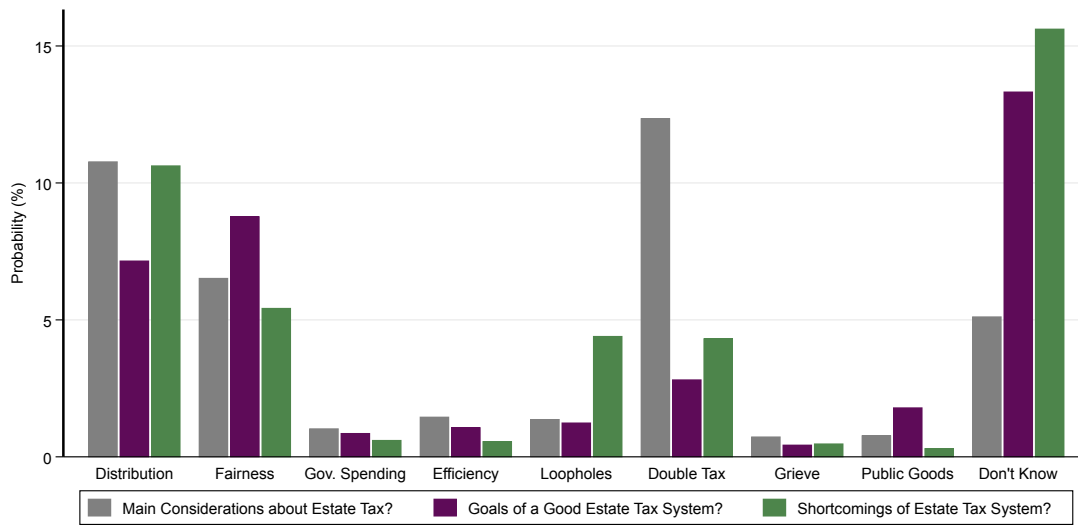
(c) By Perceived Social Class

(d) By Education Level

Notes: The figure shows keywords among different groups of respondents in answers to the questions about respondents' main considerations, desired goals, and shortcomings of the income tax. See Section OA-2 for the full text of the questions. The groups are defined by political affiliation, age, perceived social class, and education level, respectively. The score reported for a set of two words is the χ^2- test statistic, testing the null hypothesis that the occurrence of the given keywords is the same among the two groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See Section OA-5.2.

FIGURE OA-4: WORD CLOUDS FOR THE ESTATE TAX

(A) WHAT ARE YOUR MAIN CONSIDERATIONS
ABOUT THE ESTATE TAX?

(B) WHAT WOULD BE THE GOAL OF A GOOD
ESTATE TAX SYSTEM?
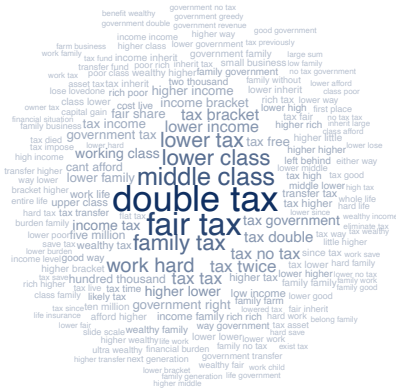
(C) WHAT ARE THE SHORTCOMINGS OF THE
ESTATE TAX SYSTEM?

(D) WHAT DO YOU THINK WOULD BE THE
EFFECTS ON THE U.S. ECONOMY IF THE
FEDERAL ESTATE TAX WERE INCREASED?
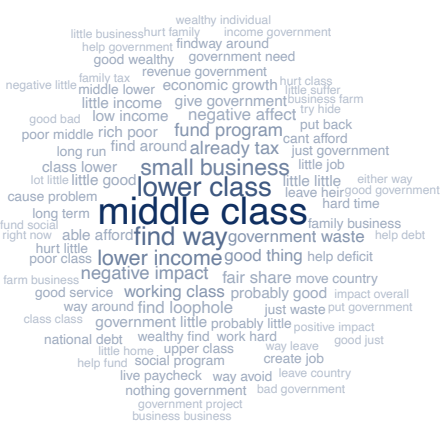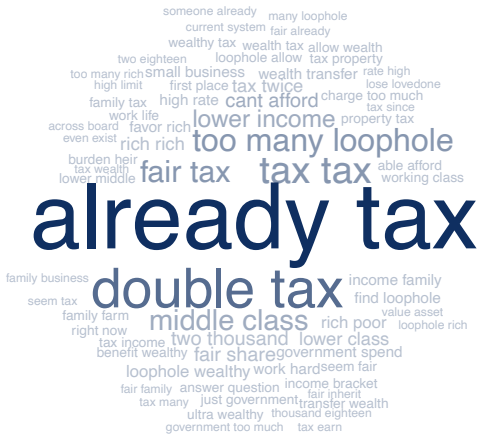
Notes: Word clouds based answers to open-ended questions stated in the caption. See Section OA-5.1.

FIGURE OA-5: KEYWORDS OF ESTATE TAX BY GROUPS

(A) BY POLITICAL AFFILIATION
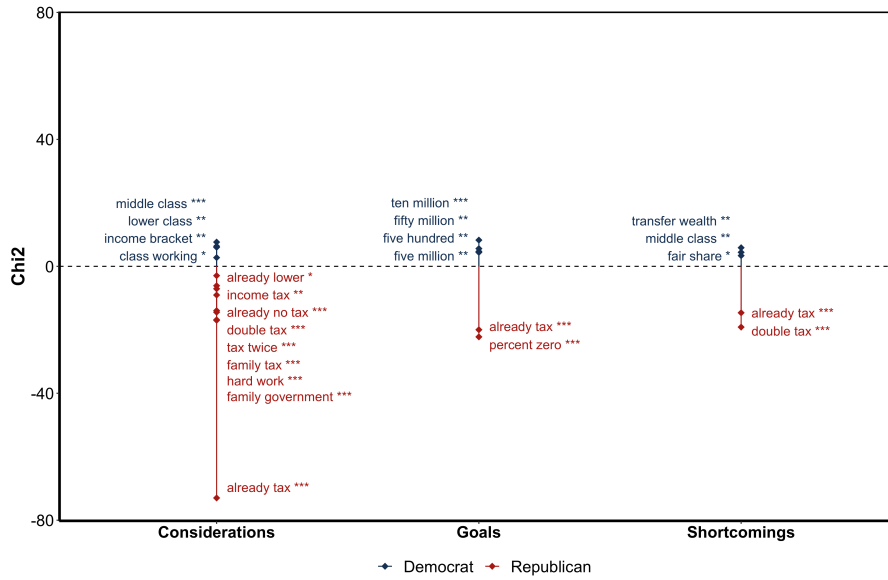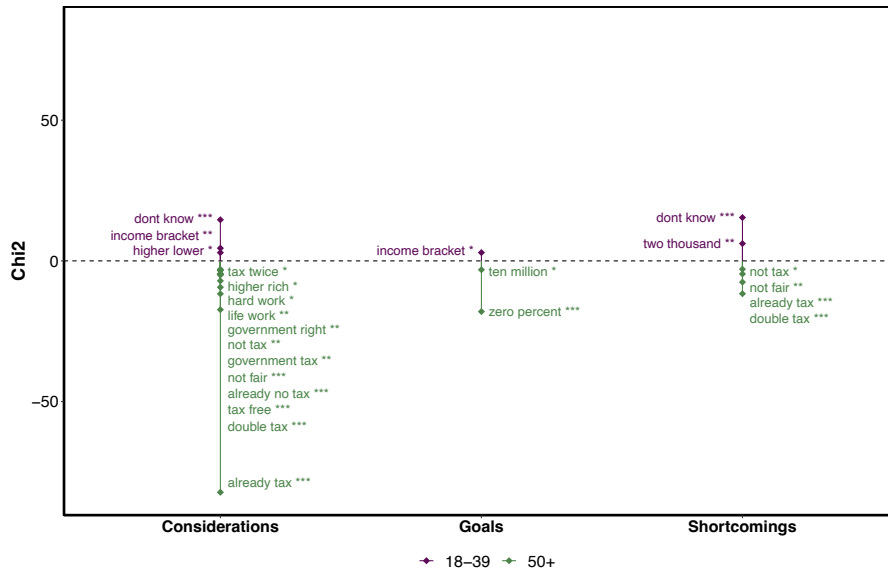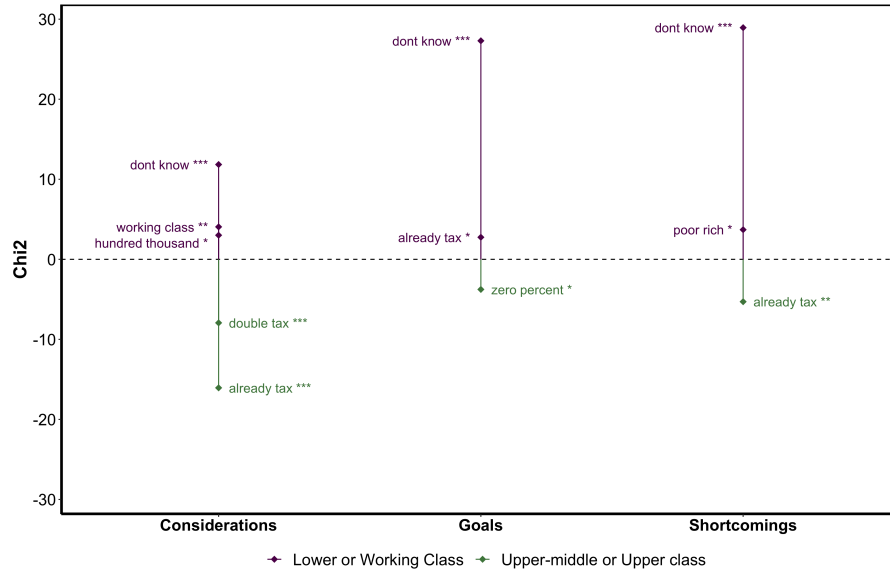
(B) BY AGE GROUP

(c) By Perceived Social Class

(d) By Education Level
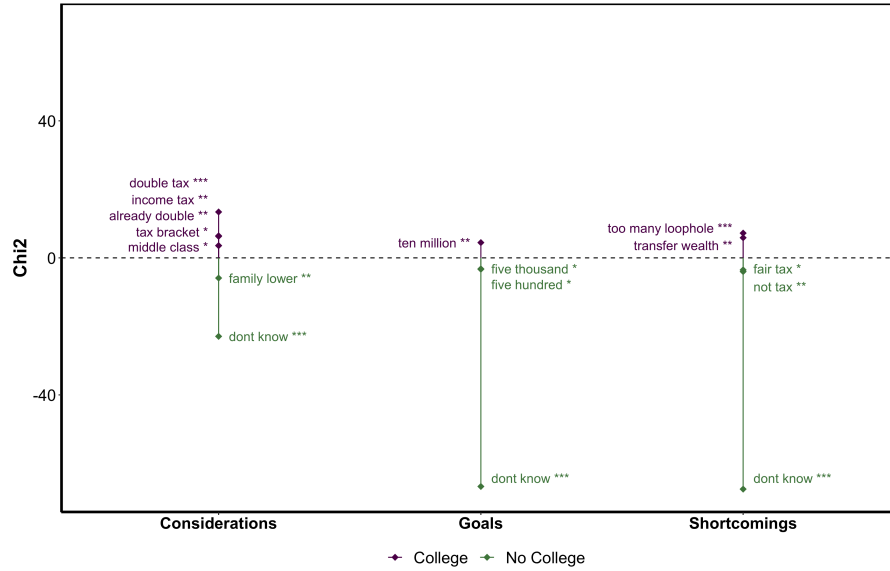
Notes: The figure shows keywords among different groups of respondents in answers to the questions about respondents' main considerations, desired goals, and shortcomings of the estate tax. See Section OA-2 for the full text of the questions. The groups are defined by political affiliation, age, perceived social class, and education level, respectively. The score reported for a set of two words is the χ^2- test statistic, testing the null hypothesis that the occurrence of the given keywords is the same among the two groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See Section OA-5.2.

FIGURE OA-6: DISTRIBUTIONAL EFFECTS OF ESTATE TAX INCREASE: WHO LOSES IF THE ESTATE TAX WERE INCREASED?
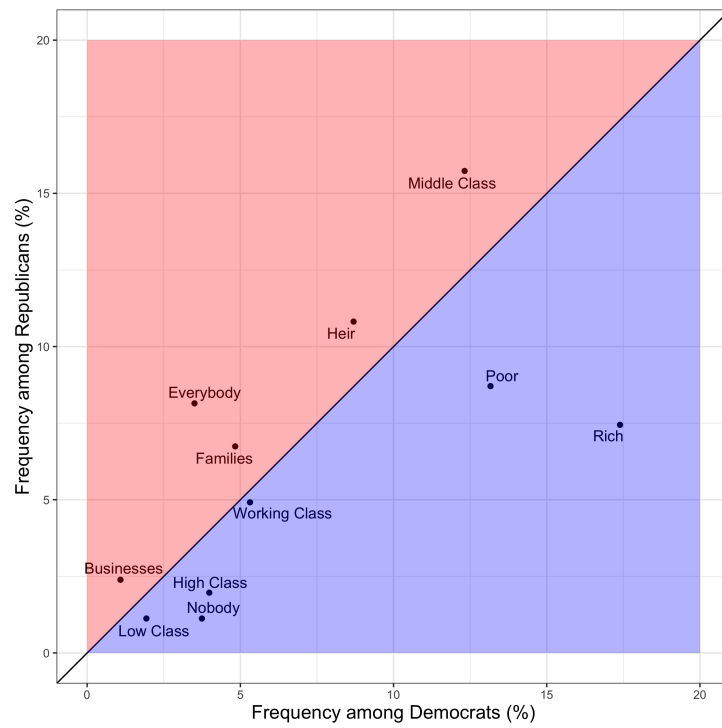
Notes: The figure reports raw frequencies among Democrat and Republican respondents of the groups mentioned when answering to the question "Which groups of people do you think would lose if the federal estate tax were increased?"

(A) INCOME TAX

Notes: The figure presents, based on the keyword-based topic analysis, the probability of mentioning topics for three open-ended income and estate tax questions. Each reported topic corresponds to an indicator variable in the keyword-count model, which equals 1 if the respondent mentions in their response at least one of the topic-defining keywords. For the list of keywords and details on the text analysis methodology, see Section OA-6 and see Section OA-5.3. Panel A: *Considerations*: When you think about federal personal income taxation and whether the U.S. should have higher or lower federal personal income taxes, what are the main considerations that come to your mind?; *Goals*: What would be the goal of a good tax system?; *Shortcomings*: What do you think are the issues with or shortcomings of the U.S. federal income tax system? Panel B: *Considerations*: When you think about the federal estate tax and whether the U.S. should have a higher or a lower federal estate tax, what are the main considerations that come to your mind?; *Goals*: What would be the goal of a good estate tax system?; *Shortcomings*: What do you think are the shortcomings of the U.S. federal estate tax? For details on the text analysis methodology, see Section OA-5.3.

(A) WHAT ARE YOUR MAIN CONSIDERATIONS ABOUT THE U.S. FEDERAL INCOME TAX SYSTEM?

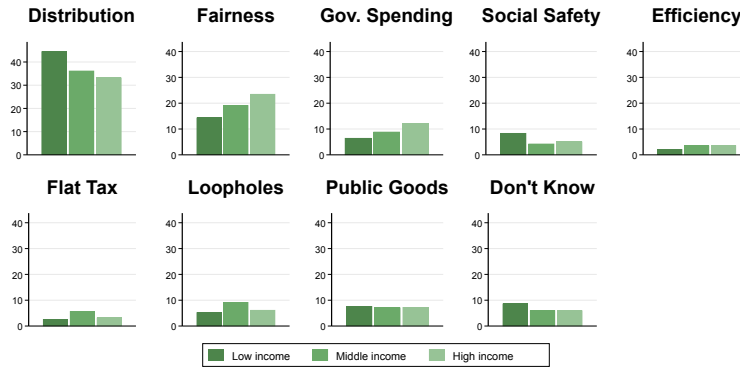(B) WHAT WOULD BE THE GOAL OF A GOOD INCOME TAX SYSTEM?

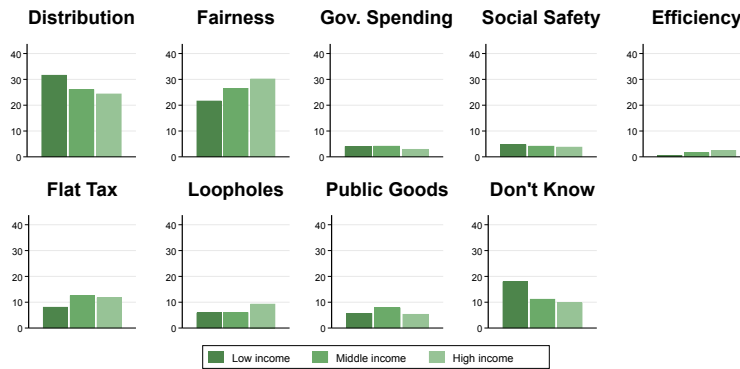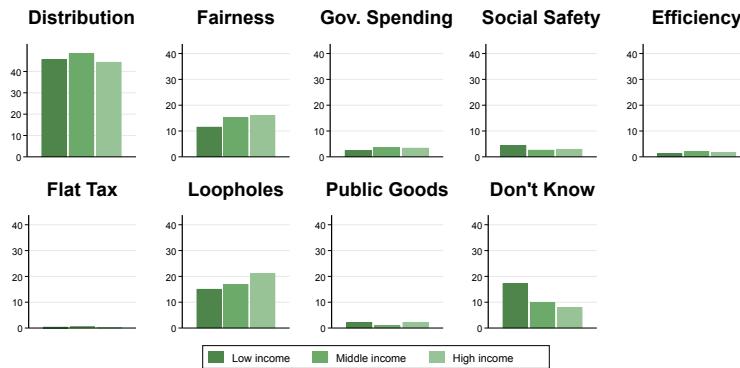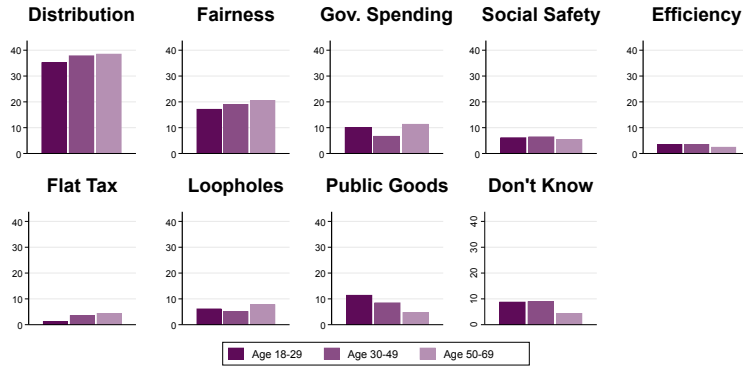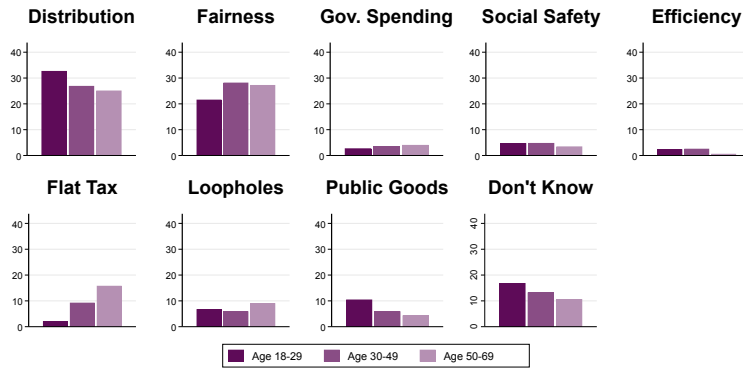(C) WHAT DO YOU THINK ARE THE SHORTCOMINGS OF THE U.S. FEDERAL INCOME TAX SYSTEM?

Notes: The figure shows distribution of topics mentioned in the answers to the open-ended questions about the income tax by political affiliation. See Section OA-2 for the full text of the questions. The bars represent the number of times a topic was mentioned out of the total mentions of any topic by political group. For the list of keywords and details on the methodology, see Section OA-6 and see Section OA-5.3.
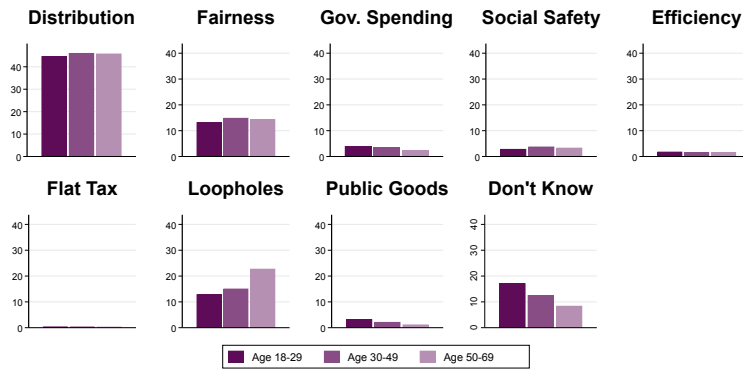
(A) WHAT ARE YOUR MAIN CONSIDERATIONS ABOUT THE U.S. FEDERAL INCOME TAX SYSTEM?

(B) WHAT WOULD BE THE GOAL OF A GOOD INCOME TAX SYSTEM?

(C) WHAT DO YOU THINK ARE THE SHORTCOMINGS OF THE U.S. FEDERAL INCOME TAX SYSTEM?

Notes: See the notes to Figure OA-8. *Low income* (*High income*) corresponds to respondents who report a pre-tax household income below (above) 39,000 (70,000) U.S. dollars; *Medium income* corresponds to respondents who report a pre-tax household income between 40,000 and 69,000 U.S. dollars.

FIGURE OA-10: TOPIC DISTRIBUTION BY AGE GROUP FOR THE INCOME TAX

(A) WHAT ARE YOUR MAIN CONSIDERATIONS ABOUT THE U.S. FEDERAL INCOME TAX SYSTEM?

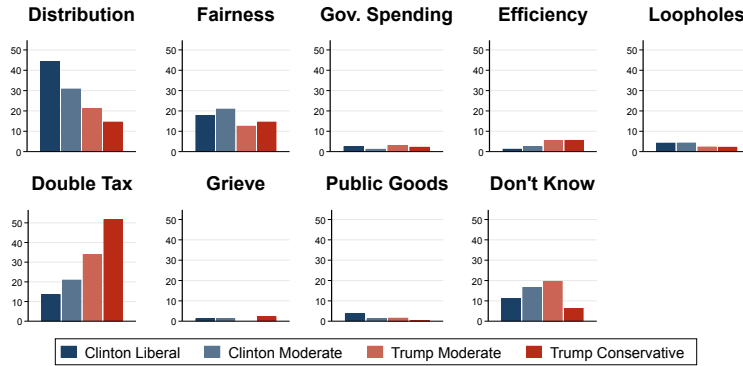(B) WHAT WOULD BE THE GOAL OF A GOOD INCOME TAX SYSTEM?

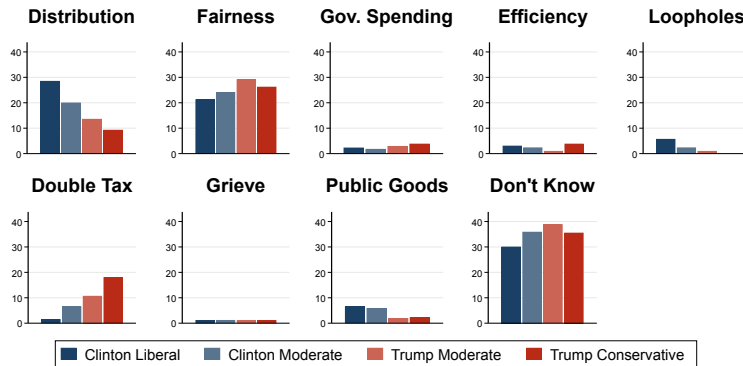(C) WHAT DO YOU THINK ARE THE SHORTCOMINGS OF THE U.S. FEDERAL INCOME TAX SYSTEM?

Notes: See the notes to Figure OA-8.

FIGURE OA-11: TOPIC DISTRIBUTION BY POLITICAL AFFILIATION OF THE ESTATE TAX
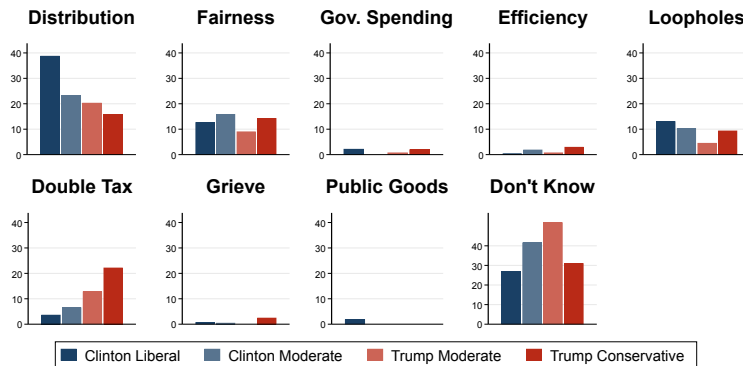
(A) WHAT ARE YOUR MAIN CONSIDERATIONS ABOUT THE U.S. FEDERAL ESTATE TAX SYSTEM?

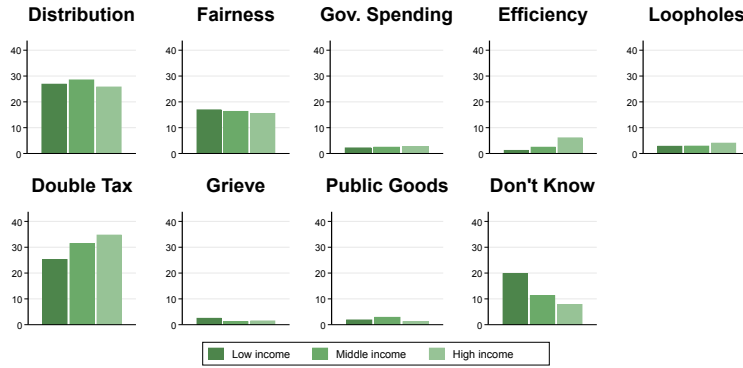(B) WHAT WOULD BE THE GOAL OF A GOOD ESTATE TAX SYSTEM?

(C) WHAT DO YOU THINK ARE THE SHORTCOMINGS OF THE U.S. FEDERAL ESTATE TAX SYSTEM?

Notes: See the notes to Figure OA-8.

FIGURE OA-12: TOPIC DISTRIBUTION BY INCOME GROUPS OF THE ESTATE TAX

(A) WHAT ARE YOUR MAIN CONSIDERATIONS ABOUT THE U.S. FEDERAL ESTATE TAX SYSTEM?

(B) WHAT WOULD BE THE GOAL OF A GOOD ESTATE TAX SYSTEM?

(C) WHAT DO YOU THINK ARE THE SHORTCOMINGS OF THE U.S. FEDERAL ESTATE TAX SYSTEM?
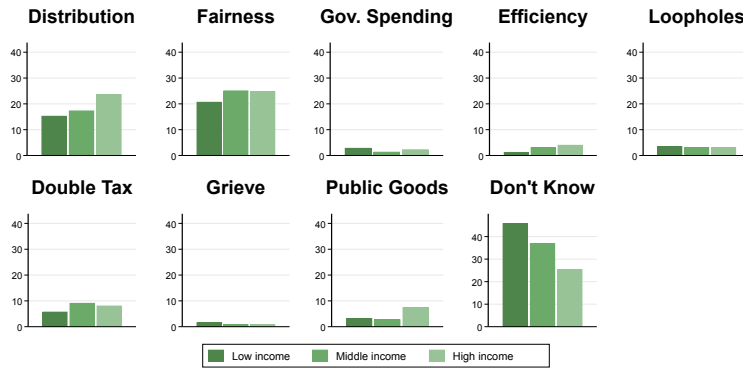

Notes: See the notes to Figure OA-9.

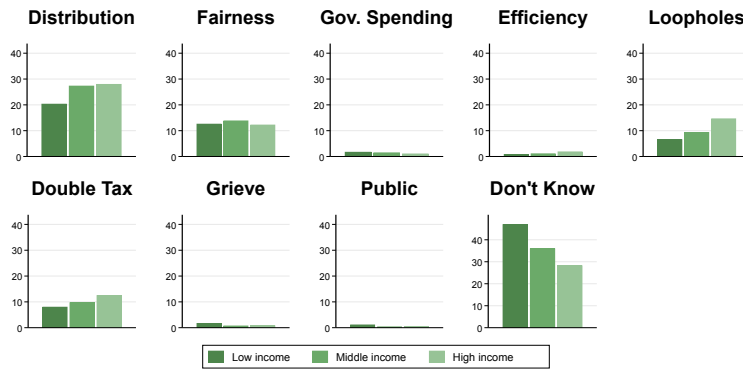FIGURE OA-13: TOPIC DISTRIBUTION BY AGE GROUP OF THE ESTATE TAX

(A) WHAT ARE YOUR MAIN CONSIDERATIONS ABOUT THE U.S. FEDERAL ESTATE TAX SYSTEM?

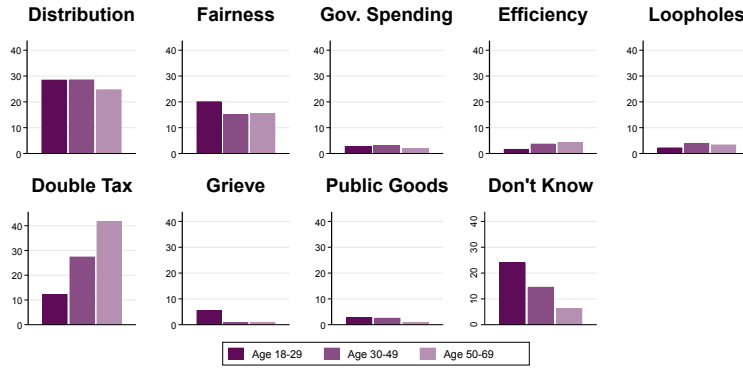(B) WHAT WOULD BE THE GOAL OF A GOOD ESTATE TAX SYSTEM?

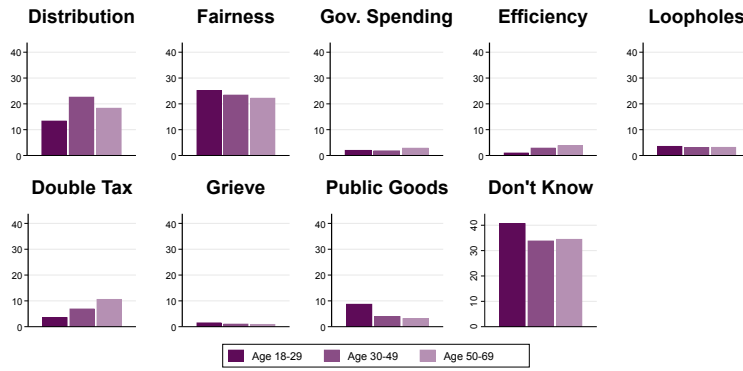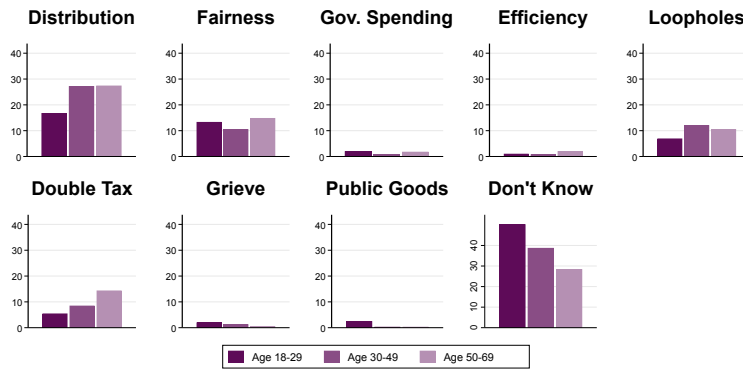(C) WHAT DO YOU THINK ARE THE SHORTCOMINGS OF THE U.S. FEDERAL ESTATE TAX SYSTEM?

Notes: See the notes to Figure OA-8.

References

Agrawal, A., W. Fu, and T. Menzies (2018). What is wrong with topic modeling? And how to fix it using search-based software engineering. *Information and Software Technology 98*, 74–88.

Baeza-Yates, R., B. Ribeiro-Neto, et al. (1999). *Modern information retrieval*, Volume 463. ACM press New York.

Banerjee, S., K. Ramanathan, and A. Gupta (2007). Clustering short texts using wikipedia. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 787–788.

Benoit, K., K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, and A. Matsuo (2018). Quanteda: An r Package for the Quantitative Analysis of Textual Data. *Journal of Open Source Software 3*(30), 774.

Blei, D., A. Ng, M. Jordan, and J. Lafferty (2003, 02). Journal of machine learning research 3 (2003) 993-1022 submitted 2/02; published 1/03 latent dirichlet allocation.

Blei, D. M. and D. Jon (2007). Mcauliffe. supervised topic models. *Advances in Neural Information Processing Systems 20*, 121128.

Blei, D. M. and J. D. Lafferty (2006). Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning*, pp. 113–120.

Deerwester, S., S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman (1990). Indexing by latent semantic analysis. *Journal of the American society for information science 41*(6), 391–407.

Flood, S., M. King, R. Rodgers, S. Ruggles, and J. R. Warren (2020). Integrated public use microdata series, current population survey: Version 8.0 [dataset]. *Minneapolis, MN: IPUMS*. https://doi.org/10.18128/D030.V8.0.

Gabrielatos, C. (2018, 02). *Keyness analysis: Nature, metrics and techniques*, pp. 225–258.

Gabrielatos, C. and A. Marchi (2011). Keyness: Matching metrics to definitions. In *Theoretical-methodological challenges in corpus approaches to discourse studies and some ways of addressing them*.

Gallagher, R. J., K. Reing, D. Kale, and G. Ver Steeg (2017). Anchored correlation explanation: Topic modeling with minimal domain knowledge. *Transactions of the Association for Computational Linguistics 5*, 529–542.

Gallup (2019). Party Affiliation, In Depth: Topics A to Z. https://news.gallup.com/poll/15370/party-affiliation.aspx.

Gentzkow, M. and J. M. Shapiro (2010). What Drives Media Slant? Evidence from US Daily Newspapers. *Econometrica 78*(1), 35–71.

Heimerl, F., S. Lohmann, S. Lange, and T. Ertl (2014). Word Cloud Explorer: Text Analytics Based on Word Clouds. In *2014 47th Hawaii International Conference on System Sciences*, pp. 1833–1842.

Hofmann, T. (1999). Probabilistic latent semantic indexing. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 50–57.

Kilgarriff, A. (2001). Comparing corpora. *International journal of corpus linguistics 6*(1), 97–133. Publisher: John Benjamins.

Leech, G. and R. Fallon (1992). Computer corpora-what do they tell us about culture. *ICAME journal 16*.

Leip, D. (2019). Election Results from the Associated Press. *Atlas of U.S. Presidential Elections and David Wasserman of the Cook Political Report*. https://uselectionatlas.org/RESULTS/.

Lu, B., M. Ott, C. Cardie, and B. K. Tsou (2011). Multi-aspect sentiment analysis with topic models. In *2011 IEEE 11th international conference on data mining workshops*, pp. 81–88. IEEE.

Partington, A. (2014). Mind the gaps: The role of corpus linguistics in researching absences. *International Journal of Corpus Linguistics 19*(1), 118–146.

Roberts, M. E., B. M. Stewart, D. Tingley, E. M. Airoldi, et al. (2013). The structural topic model and applied social science. In *Advances in neural information processing systems workshop on topic models: computation, application, and evaluation*, Volume 4, pp. 1–20. Harrahs and Harveys, Lake Tahoe.

Roberts, M. E., B. M. Stewart, D. Tingley, C. Lucas, J. Leder-Luis, S. K. Gadarian, B. Albertson, and D. G. Rand (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science 58*(4), 1064–1082.

Salton, G. and M. J. McGill (1983). *Introduction to modern information retrieval*. mcgraw-hill.

Schönhofen, P. (2009). Identifying document topics using the wikipedia category network. *Web Intelligence and Agent Systems: An International Journal 7*(2), 195–207.

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

Stubbs, M. (2010). Three concepts of keywords. *Keyness in texts*, 21–42.

Tang, J., Z. Meng, X. Nguyen, Q. Mei, and M. Zhang (2014, 22–24 Jun). Understanding the limiting factors of topic modeling via posterior contraction analysis. In E. P. Xing and T. Jebara (Eds.), *Proceedings of the 31st International Conference on Machine Learning*, Volume 32 of *Proceedings of Machine Learning Research*, Bejing, China, pp. 190–198. PMLR.

Taylor, C. (2013). Searching for similarity using corpus-assisted discourse studies. *Corpora 8*(1), 81–113.

Watanabe, K. (2018a). Conspiracist propaganda: How russia promotes anti-establishment sentiment online. In *ECPR General Conference, Hamburg*.

Watanabe, K. (2018b, March). Newsmap: A semi-supervised approach to geographical news classification. *Digital Journalism 6*(3), 294–309.

Watanabe, K. and Y. Zhou (2020). Theory-driven analysis of large corpora: Semisupervised topic classification of the un speeches. *Social Science Computer Review*, 0894439320907027.