Putting the Public Back in Public Debt: Citizen Narratives on Public Debt Burdens

Matthew DiGiuseppe
Leiden University*

Alessia Aspide
Leiden University

Bastián González-Bustamante
Leiden University & Universidad Diego Portales

Jaroslaw Kantorowicz
Leiden University

January 25, 2025

Abstract

Models of public debt accumulation and consolidation often rely on implicit assumptions about citizen preferences regarding their attention to fiscal policy and the cleavages that form around policies. However, few studies directly examine public engagement with the issue. We address this gap using original open-ended survey data from three highly indebted countries: Italy, Japan, and Brazil. Leveraging generative Large Language Models (LLMs) to code open-ended responses, we assess the public's understanding of the consequences of rising debts and if it aligns with economic evaluations by comparing their responses to those of "synthetic economists" utilizing the domain knowledge of LLMs. We find that almost all citizens expect negative consequences from public debt increases but are less specific in their answers than synthetic economists. Next, we look for evidence that cleavages over debt policy fall on partisan and generational lines and find that age and partisanship are associated with small differences in preferred policies to reduce debt. However, the pattern is not consistent across countries. The analysis reveals that the public is not entirely ignorant of the consequences of public debt, and cleavages are not as dramatic as popular models may have assumed.

*Corresponding author: m.r.di.giuseppe@fsw.leidenuniv.nl

Given the centrality of debt and credit to almost all government activities in modern economies, scholars from several disciplines have proposed theories for why some countries can maintain sustainable debt levels while others fail to restrain deficits and pay down debt. These explanations largely focus on the decisions of elites within a government, whether they be heads of state, political parties, finance ministers, organized interests, or bureaucrats. Yet, until recently, scholars haven't tried to understand how the public reasons about public debt. This isn't to say that others have ignored the public completely. Models theorizing about the behavior of elite decision-makers rest on explicit or implicit assumptions about how the public does or does not shape the incentives of politicians. Theories about the public's influence range from passive standers, to myopic spendthrifts, to staunch fiscal conservatives (Buchanan, 1967; Buchanan and Wagner, 1977; Alesina and Tabellini, 1990; Alesina and Drazen, 1991a; Alesina et al., 2020; Yared, 2019). Scholarship has also looked at how direct involvement of citizens through, e.g., referenda and initiatives on fiscal issues affects public debt and spending (Asatryan et al., 2017; Blume et al., 2009; Feld and Kirchgässner, 2001; Funk and Gathmann, 2011; Matsusaka, 2018)

Surprisingly, little research has, until recently, attempted to sort out these assumptions that stand as the base of these prominent theoretical models. As with most studies of the public's economic preferences and attitudes, the nascent literature aimed at understanding citizens preferences toward debt has utilized experimental or observational analysis of closed-question survey data and has focused on examining differences in opinion toward either deficits, debt financing, debt rules or, most commonly, fiscal adjustment to reduce debts (Aspide et al., 2021, 2023; DiGiuseppe and Del Ponte, 2023; Curtis et al., 2015; Curtis, 2014; Ardanaz et al., 2023, 2024; Kantorowicz, 2023; Roth et al., 2022; Barnes and Hicks, 2018, 2022; Bremer and Bürgisser, 2021). In this paper, we diverge from this approach and instead, examine open-ended responses to questions about the consequences and solutions to high public debt in three highly indebted countries: Italy, Brazil, and Japan.

Open-ended responses allow us to examine preferences without directing respondents into prescribed answers that cue them to think about the topic in a specific way. As such, they give us more accurate, though more challenging to analyze, insights into the citizens' mental models, reasoning, and depth of understanding of economic policy (Haaland et al., 2024; Andre et al., 2023). This is especially important in the area of public debt given the common assumptions that people heavily discount debt or have a limited understanding of its consequences (Buchanan and Wagner, 1977; Shi and Svensson, 2006; Bremer and Bürgisser, 2021).

The use of open-ended questions is not new. However, their use has, until recently, been limited by the cost of annotating and classifying answers. This has started to change with the latest developments in automated text analysis (see Grimmer et al. (2022)). While these automated methods are a step forward, the learning curve has still been lowered by the advent of generative Large Language Models (LLM) like GPT-4, Claude, Gemini, and Llama (Mellon et al., 2024; Gilardi et al., 2023; Heseltine and Clemm von Hohenberg, 2024; He et al., 2024; Linegar et al., 2023). Toward this end, we take advantage of LLM advances to make sense of the unstructured data and seek evidence consistent with theoretical expectations about the distribution of preferences over debt policy and knowledge about debt.

Our analysis is guided by two assumptions about the public in the literature. The first is that high debt burdens ultimately stem from citizens' ignorance of fiscal matters that serve the short-term interests of politicians (Buchanan and Wagner, 1977; Dollery and Worthington, 1996; Yared, 2010; Lizzeri and Yariv, 2017; Bisin et al., 2015). By contrast, others argue that citizens are sufficiently informed and fear the consequences of debt and deficits and that politicians have short-sighted preferences (Brender and Drazen, 2008; Alesina et al., 2020). This follows closely from the Ricardian equivalence assumption that citizens save in response to the expectation that government debt will mean higher future taxes (Barro, 1974). To help make progress on this and other related debates, we examine if citizens expect, un-

prompted, future fiscal contraction (higher taxes, less government spending) or economic crises as debt grows. To judge if their answers align with economic theory, we compare their responses to the responses of "synthetic economists" generated by LLMs. We find that citizens in all three countries, almost unanimously, expect negative economic consequences from a further increase in public debt. However, their expectations are not as specific as those of our benchmark synthetic economists. Instead of pointing to specific mechanisms, like higher taxes, they point to general effects like reduced quality of life and increased poverty.

Next, we examine assumptions about cleavages over debt policy. Models of debt accumulation often center around distributive conflict across generational (Cukierman and Meltzer, 1989; Song et al., 2012; Tabellini, 1991; Müller et al., 2019; Aspide et al., 2021) or partisanship and class (Persson and Tabellini, 2002; Alesina and Tabellini, 1990; Alesina and Drazen, 1991b; Lizzeri, 1999; Müller et al., 2019; Bansak et al., 2021; Barnes and Hicks, 2018; Nelson and Steinberg, 2018). We probe if these cleavages exist in citizens own narratives by looking for a correlation between age and partisanship and policies mentioned in citizens responses. Our analysis reveals differences in proposed policies to address debt, showing slight variation with age and partisanship. However, the effects are inconsistent across countries.

In addition to looking for evidence of existing assumptions about preferences for taxes or spending, our analysis also helps uncover latent themes in the public's public debt narratives. For example, citizens are keen to point to efficiency gains and fighting corruption as ways to reduce debt instead of austerity measures that are drawn on in most academic research.

We also make a broader contribution to studying economic policies and open-ended responses. Recently, scholars have rediscovered the depth provided by open-ended responses (Stantcheva, 2021; Andre et al., 2023). Our paper shows how LLMs can open up this research area to those who lack resources for research assistants. Further, this is the first paper, we are aware of, to leverage LLMs a sample of a synthetic experts upon which to compare human responses.

Why Open-Ended Responses?

Closed-ended questions have the advantage of standardization, making them easy to analyze in a systematically and structured way. The cost of this standardization is the bias generated by forcing respondents to choose among a set of responses introduced by the researcher that may not be initially considered by respondents (Connor Desai and Reimers, 2019, 1427). As such, closed-ended questions likely overstate preferences for displayed options. Further, because the choice of options is limited, they inherently limit potential responses to those provided by the survey designers. For example, a question about preferences for debt reduction may ask respondents to choose between a variety of taxes or spending cuts. This may exclude other preferences, such as those for increased investments (aimed at GDP growth) or greater tax enforcement, as we see below. The constrained choice might lead scholars to conclude, as Bremer and Bürgisser (2021) do, that citizens are unwilling to take action to address debt. This is especially a problem when it comes to fiscal consolidation. Public debt's relevance to almost every aspect of the economy makes it difficult to distill it into a few neat options, irrespective of whether researchers are asking about the causes, consequences or solutions to high public debt. As such, the value of open-ended responses is likely to be particularly high in efforts to understand public debt attitudes and other issues with such a wide breadth.

Beyond providing insight into the diversity of preferences, open-ended responses may also allow for a greater understanding of the limits of public understanding. Given options, survey respondents may be reluctant to select "don't know" due to a variant of social desirability bias. This is a well-known issue with closed-end public opinion surveys in which respondents report an opinion but really do not have strong views on the topic (Converse, 2006). Even if respondents select "don't know", they may be masking uncertainty about the topic (Graham, 2021). In an open-ended setting, they do not have an easy way out, and their ignorance or

ambivalence is more easily revealed. Further, they may also reveal preferences for nonsensical solutions that are, for good reason, excluded from closed-ended options. For example, we find that a sizable proportion of Japanese citizens advocate for reducing politicians' salaries - a small solution to the large debt burden of 260% of GDP. As such, open-ended questions reveal the limits of citizen understanding to a greater extent than closed-ended questions.

Despite the benefits of open-ended responses, they are infrequently used in research because imposing structure on the raw data had, until recently, required costly human coders. These coders would have to read respondent answers and categorize responses into predefined classes or identify other elements of the text, like sentiment. When responses are sufficiently large in number this creates monetary costs beyond the reach of most social scientists. More recently text-as-data innovations, like structural topic modeling (STM) or semantic networks, have allowed for the automation of this process. However, the techniques rely mainly on analyses of single words or 2-3 word strings that necessarily omit important context and relationships between concepts. Further, attempts to measure sophistication or knowledge with text-as-data tools often diverge from domain knowledge. Instead, they assess proxies like language and discursive sophistication, which evaluates the number of topics, dispersion of topics, and connections between topics divorced from the topic context (Kraft, 2023). As such, they measure complexity rather than actual content.

As described below, LLMs overcome cost and methodological limitations as they replicate the work of human coders. They allow researchers to automate the classification process that then permit statistical description and analysis (Mellon et al., 2024; Heseltine and Clemm von Hohenberg, 2024).

Open-Ended Responses and Public Debt

The study of sovereign debt emerged from macroeconomics. Economists generally model decisions about budget deficits and debt consolidation as a top-level decision by a country or by competing political groups. However, all these models rest on implicit or explicit assumptions about either the economic or political motivations or behaviors of citizens in responding to increasing debt burdens and the need for public debt reduction. Yet, we still have very little evidence on how people think about public debt. The closed-ended survey evidence we do have (Bremer and Bürgisser, 2021; Aspide et al., 2021; Curtis et al., 2015; Ardanaz et al., 2024) fix a limited set of responses and thus limit our understanding of the diversity of citizen attitudes.

Open-ended responses help us better understand how the public reasons about public debt in several ways. First, open-ended responses can help the research discovery process by bringing attention to issues researchers have yet to theorize about or consider in economic models (Stantcheva, 2021; Andre et al., 2023).

Second, open-ended questions can help us generate evidence for the fundamental assumptions in models of public debt. There are several key assumptions that we will engage with in our analysis. The first and second assumption center around citizens' knowledge of public debt. Macroeconomic models have traditionally assumed that debt is equivalent to taxation (Barro, 1979) and that citizens respond to increases in public debt by saving as they expect that debt implies higher taxation in the future. As such, they must save and collect interest to repay higher government taxes in the future. This idea of Ricardian Equivalence implies that citizens are paying attention to economic policy and have a nuanced understanding of the impact of public debt on both government policy and their own future welfare. Political economy scholars seeking to explain rising debt burdens have questioned this assumption. They see the cause of rising debt stemming from the public's inattention or ignorance of gov-

ernment debt and politicians' incentive to exploit it for short-term electoral gain. Suppose voters are unaware or ignorant about the costs of debt financing. In that case, politicians can engage in "fiscal illusion" to give them the impression that they are providing costless public or private goods while they are simply shifting the burden to future periods (Buchanan and Wagner, 1977; Congleton, 2001; Shi and Svensson, 2006; Yared, 2019; Bisin et al., 2015). Such arguments stem from a Downsian (1956) understanding of voters, which contends that they have an incentive to forgo the process of information acquisition.

Thus far, research on the rationality or attentiveness of voters has examined savings behavior in economic data or in lab experiments conducted on hypothetical economic circumstances (Ricciuti and DiLaurea, 2003) or has tried to tease out cross-national differences. For example, several studies examine the correlation between country-level measures of informed voters and national budget deficits (Shi and Svensson, 2006; Jank and Libich, 2019). Data at such a high level of aggregation limits what can be said about citizens-level assumptions and lab behavior can treat 'ignorance,' but cannot tell us about how much of it exists in real-world scenarios. Other research consistent with the assumption of well-informed citizens examines how voters respond to deficits. For example, Brender and Drazen (2008) and (Alesina et al., 2020) look at electoral outcomes and find a correlation between deficits and leader removal. However, the result might stem from the downstream effects of deficits - interest rates or high capital costs - rather than direct punishment of deficits themselves. As such, it is hard to directly infer what voters think or if their actions are explained by more readily observed confounders not picked up by macroeconomic statistics. Examining open-ended questions can provide insight into not only what citizens expect as debt increases but what is the depth of their understanding relative to macroeconomic theory.

The next set of assumptions about the public's relationship with public debt centers around distributive cleavages. Political economy models of debt often assume that differences in the importance and solutions to public debt fall along generational or partisan cleavages.

For example, many have suggested that generational divides define preferences for debt reduction as the elderly have less interest in consolidation than the young who will be responsible for repaying future debts (Cukierman and Meltzer, 1989; Tabellini, 1991; Song et al., 2012) and recent evidence suggests a non-linear relationship between age and support for public debt reduction (Aspide et al., 2021). This evidence largely relies on closed-ended questions from large, omnibus surveys and thus misses nuances in the differences in the ways young, middle-aged, and the old think about public debt.

Two strands of literature predict that partisanship will generate cleavages in public debt consolidation. This stems from two separate arguments. The first is that material interest, often funneled through partisanship, shapes cleavages. Class conflict determines the shape of reform, either leading to spending cuts or progressive tax increases. For example, Curtis (2014) and Curtis et al. (2015) find that material interest variables, along with partisanship, are strong predictors on a vote for whether or not to repay Iceland's debt following the great recession.[1]

It is also possible that material interests don't generate partisan preferences, but instead, partisanship itself shapes attitudes. More recent research suggests that differences in attitudes towards debt accumulation and austerity result from partisan cues rather than material interests (Nelson and Steinberg, 2018; Barnes and Hicks, 2018; Bansak et al., 2021). Citizens' views on public debt, according to this argument, stem not from their material interest but rather follow from co-partisan cues from political leaders shaping debates. This argument also stems from a view of uniformed voters who have been socialized to espouse particular policy positions with little reflection.

In the sections below, we use citizens' own narratives to find evidence consistent with these three prominent arguments about citizens' views on public debt. We first examine

[1]While class and partisanship are two distinct concepts, in practice, parties are often a vehicle to organize class interests and thus it is difficult to untangle the effects empirically.

the degree by which citizens anticipate the likely costs of rising debt burdens for their own economic lives and how their responses compare to answers generated by an LLM asked to take on the role of economists. Next, we compare how the understanding of the consequences and solutions of debt vary by age and partisanship. This will help us understand if salient cleavages identified in the literature are more than a product of closed choices or if they still emerge in citizens' own narratives.

Original Survey Data

The analysis relies on original survey data collected in three countries with high debt burdens but which have not yet experienced a sovereign default or restructuring - Japan, Italy and Brazil. In each country, we recruited participants from quota representative samples greater than 1500 that are representative of the domestic population on the dimensions of age, region and gender.[2]

By choosing 'at risk' countries, we get a better understanding of narratives when debt is a salient political topic, but attitudes have not been crystallized or polarized by a costly default. Arguably, this positioning is most interesting from a theoretical perspective as it gives us insight into the politics on the path to a debt crisis and when the potential to change course remains an option. While informative, examining states where public debt is not salient is potentially less interesting as citizens have little incentive to develop narratives on public debt.

Each of the countries we selected differs in their level of economic development but also in their debt trajectory, monetary regime (monetary union vs floating), and relationship to external capital markets. This diversity allows any common findings the potential to claim greater external validity. Given the inability to address the many confounding contextual

[2]We recruited participants with Respondi in Italy in July 2022, Rakuten Research in Japan in October 2023, and Netquest in Brazil in September 2022.

differences, we restrain ourselves from comparing countries directly.

In each country, we asked the same four questions[3] and asked respondents to respond in 2-3 sentences:

- In your opinion, what are the main causes of [Italy,Brazil,Japan]'s high public debt?
- What do you think would happen to the [Italian,Japanese,Brazilian] economy if the national debt continues to increase?
- **What do you think would happen to your economic situation if the national debt continues to increase?**
- **What policies should the government adopt to reduce [Italy,Brazil,Japan]'s public debt?**

In this paper, we analyze the final two questions. Given the multilingual capabilities of frontier LLMs, we use the untranslated text. We remove nonsensical responses, such as a few instances of random typing to fill in the text box. We use the LLMs to identify these responses and omit them accordingly.

Multi-Label Classification

The benefit of unstructured responses is that they allow for a multitude of answers. Yet, both exploratory and induction-based analysis require that we reduce the dimensionality of the data. Our analysis centers around using an LLM (Llama 3.1:405B) to first to identify common themes or topics in each question and then apply multiple overlapping classifications to each response. We take a two-step approach with a human in the loop. Analyzing each open-ended question separately, we first ask the LLM to summarize the response by providing a 2-3 word description of each consequence or policy mentioned by each respondent in a zero-shot attempt. This means we did not provide any examples of classification in the prompt.

[3]In Italy and Brazil, we asked an additional question before these four questions: "When you think about [Country]'s public debt, in wondering whether it is too high or too low, what are the main considerations that come to mind?". We omitted this in Japan for space considerations after seeing it did not lead to useful results. In each case, the questions were proceeded by basic demographic questions.

First Stage Prompt Example

I asked Brazilian citizens what they thought would be the consequences of a further increase in the country's high public debt for their own economic situation. Please identify, in English, the consequence listed in the answer I provide. Try to describe each consequence in two words and separate each consequence with a dash (-). Reply with NA if there is no response. Again, reply only in English. Here is the response:

We then fed the entire list of causes to Claude Opus 3.0 (another LLM) to summarize the causes into 10 distinct categories.[4] The research team then reviewed the categories from all three countries and adjusted them to distinguish theoretically interesting concepts for analysis. For example, the LLM created one tax reform category, including tax cuts and increases. We created separate groups as, for our purposes, they are conceptually distinct. We selected 9 categories when asking about consequences and added a "no consequences" category, given its theoretical relevance. We then looped each full response through the LLM again, prompting the LLM to classify the response into the identified categories.

[4]Claude Opus has a larger context window that fits summarized responses from all respondents in a single prompt.

Third Stage Prompt Example

I asked Brazilian citizens what they thought would be the consequences of a further increase in the country's high public debt for their economic situation. Your task is to carefully consider the response from each citizen and categorize it into one or more of the following categories of consequences. In your response, list the numbers of the categories mentioned, ordered by their importance to the respondent (starting with the most important). If you choose '11. Other, please add a brief 2-3 word description next to your answer. If the response is non-sensical reply with '11. Other - NA'. If the respondent indicates that they don't know, reply with '12. Don't know'. Here are the consequence categories: 1. Inflation, currency devaluation, higher cost of living - 2. Higher taxes - 3. Reduced public services and benefits - 4. Unemployment or reduced wages - 5. Increased poverty and financial hardship - 6. Savings erosion and investment losses - 7. Inability to pay debts or bankruptcy - 8. General reduced quality of life or economic decline - 9. Pension cuts - 10. No consequences or minor consequences - 11. Other - 12. Don't know. Here is the response:

Model Choice

Recent advances in generative Large Language Models (LLMs) have dramatically reduced the costs of using open-ended responses in survey research and analyzing other text data. Prior to the recent advances in LLMs, the annotation of text data was either coded by humans (research assistants or crowd workers) at great expense or limited to the analysis of specific words using structural topic models that are limited in their ability to pull nuanced context from the data.

While LLMs have demonstrated an impressive ability to annotate tasks that exceeds STMs and is on par with humans, the performance is conditional on the model selected. Beyond performance, other considerations, like replicability, play into choosing an LLM for research purposes. Proprietary models, like those offered by OpenAI (GPT-3.5, GPT-4, GPT-4o or o1), hold no guarantee of being available in the future as newer models are rolled out. Consequently, their use is closer to using human coders in that the exact coding

decisions are created by a black box and not easily replicated. As such, it is recommended to use open-source models (Palmer et al., 2024).

In this paper, we are using data from three different non-English samples. Given the added complexity of non-English languages, we opt for a large but open-source model: Llama 3.1 with 405 billion parameters.[5] Llama 3.1 is one of Meta's state-of-the-art LLMs. Its 405B version is one of the largest open-source models available. It has high general knowledge and multilingual capabilities and performs comparable to cutting-edge proprietary models such as GPT-4o and Claude 3.5 Sonnet.[6]

Lastly, API users of LLMs are free to set the temperature parameter of the models. A low temperature sets the model to be predictable, a high temperature sets the model to be more creative in our responses. We keep the temperature at its lowest value in our API calls to ensure the responses are deterministic and replicable across calls.

Crowd Worker and Expert Validation

Recent research indicates that LLMs perform very well at similar classification tasks (Mellon et al., 2024; Gilardi et al., 2023). To demonstrate that LLMs are suitable for the multi-label classification task used here, we compare the output of Llama to crowdsource workers.[7] Such validation is recommended when using multiple languages (Gilardi et al., 2023; Heseltine and Clemm von Hohenberg, 2024). In each country, crowdsource workers annotated 150-200 open-ended responses to the question on policies to address debt. Beyond crowd workers, we also had a late-stage, Italian, political economy PhD student code 100 responses on the self-consequences question.

Figure 1 presents the F1 scores comparing the average multi-label classification of LLama

[5]We accessed the model via the fireworks.ai API.

[6]Ethics is another concern given we are feeding respondent answers to an API. Note we only fed the text to the LLM and did not provide any identifying information and was planned to be shared in replication files regardless of LLM use.

[7]We recruited crowd workers via the Prolific platform.

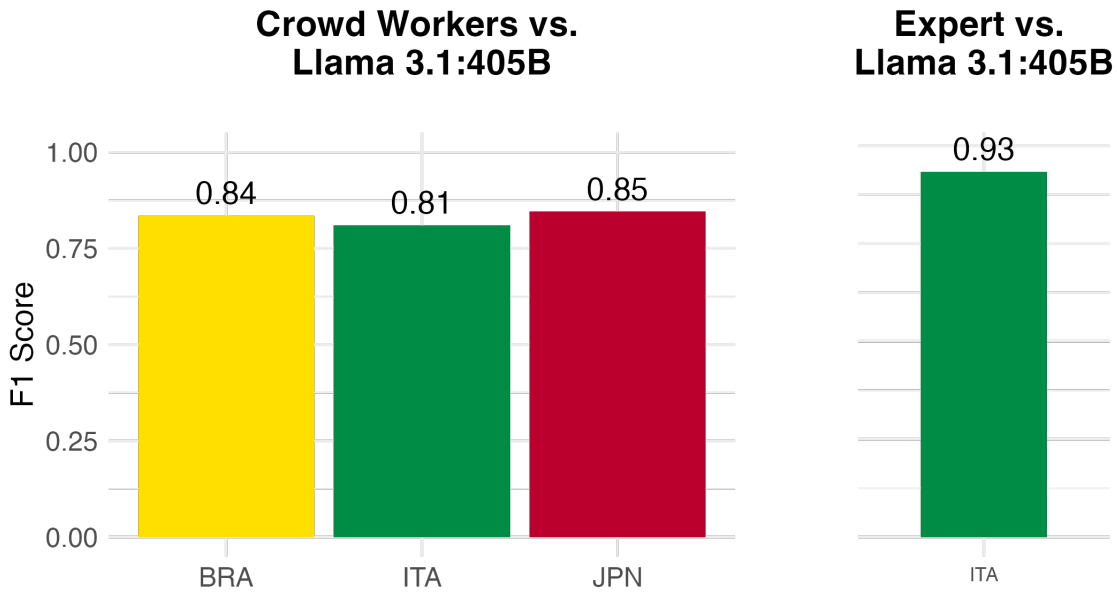

Figure 1: **F1 Scores:** Comparison of crowd workers and our expert in multi-label classification of the consequences of an increase in debt. We calculate the F1 score for each label and then aggregate the F1 scores for all labels in the task.

3.1 with crowdsource workers and, in Italy, with our expert PhD student. F1 is a harmonic measure of precision and recall. While precision is the faction of positive predicted values, recall is the proportion of correct predictions among true-positive cases (Biecek and Burzykowski, 2021).[8] Since we use multi-label classification, we calculate F1 for each label and then aggregate the final results. We see that the LLM has a high agreement with crowdworkers. The F1 scores range from 0.81 to 0.85. Although there is no formal agreement about a qualitative interpretation of the measure, a score above 0.80 is generally considered very good (Manning et al., 2008; Powers, 2011). Further, we see even higher agreement with the expert.

[8] $\text{F1 Score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\right)$. $\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$. $\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$.

The Individual Consequences of Increasing Debt

We now examine respondents' expectations for their own economic situation, given a continued increase in public debt. Figure 2 presents the 12 categories resulting from our LLM multi-label classification process. Given multiple labels can be applied to one response, the percentages exceed 100%.

This data can help us understand if respondents have realistic expectations of high debt or if their expectations deviate from reality. Economic orthodoxy would suggest that increasing debt can have numerous causes. They would include higher taxes (Ricardian equivalence), inflation, reduction in public services, and higher costs of finance (Blanchard, 2021). At the extremes, high debt can lead to crisis.

We see that few respondents explicitly mention higher taxes as a consequence of a further increase in debt. With 23% mentioning taxes in Japan, 12% in Italy and only 4% mentioning it in Brazil. Strictly speaking, the finding is at odds with Ricardian Equivalence, which suggests that households would anticipate higher taxes as a direct consequence of higher debt burdens.

While taxes are only mentioned among a minority of respondents, what is noticeable is that respondents do not have optimistic expectations of a further increase in debt. Many, consistent with textbook treatments of debt, mention "inflation, devaluation, [or] a higher cost of living" in all three countries.[9] Most other responses mention other negative consequences such as reduced quality of life, increased poverty and financial hardship, unemployment, pension cuts, a reduction in savings or the inability to pay bills. Given the multiple classifications, Figure 2 can not tell us directly about negative or positive consequences distribution. However, further analysis presented in Figure 3 reveals that a strong majority of respondents (over 75%) in each country find the rise in public debt to hold negative consequences for

[9]This may be a not-entirely-unrelated produce of the survey being conducted post-Covid when countries were experiencing high inflation relative to previous periods.

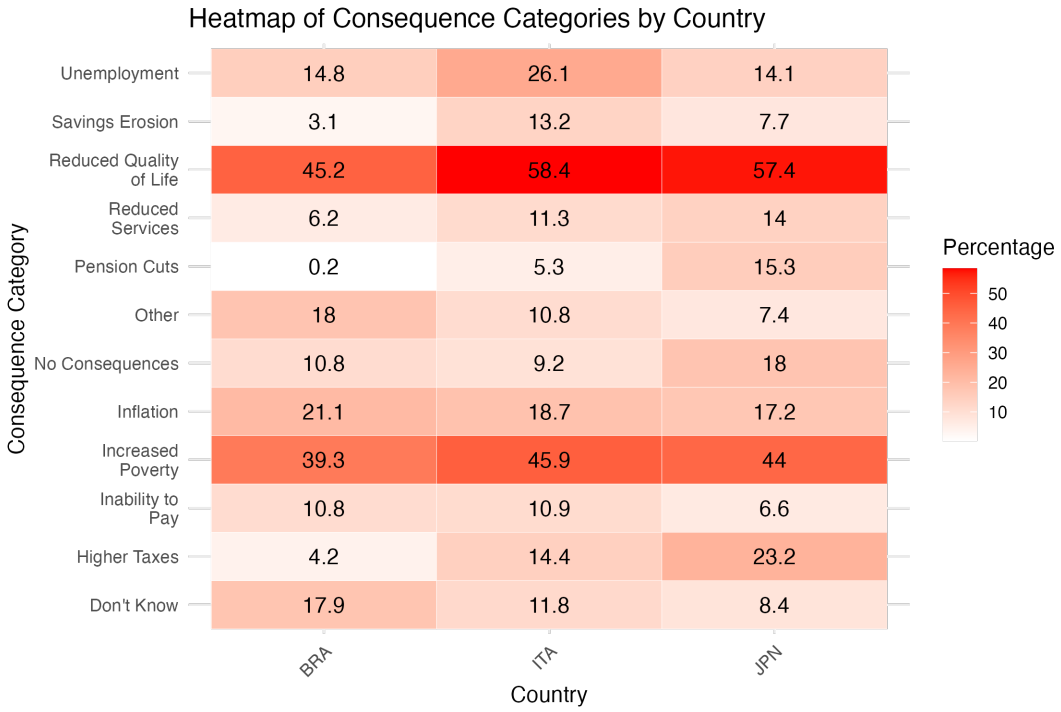

Figure 2: **Individual Consequences of Public Debt Increase:** Here we show the distribution of responses to the question "What do you think would happen to your economic situation if the national debt continues to increase?" by country sample. Respondents can mention multiple consequences. As such, the categories do not sum to 100%.

their economic lives.[10]

The Public vs. synthetic Expert

The descriptive evidence thus far indicates that respondents are overwhelmingly negative about assessing how an increase in public debt will impact their economic lives. We can see quite clearly, that many do not expect taxes in line with Ricardian Equivalence. Yet, it isn't clear how far they deviate from the expectations of economists in general, many of which have a more nuanced understanding of public debt's influence than expressed in Ricardian Equivalence.

[10]We included the "other" category in this aggregation due to these comments' overwhelmingly negative but high variance nature.

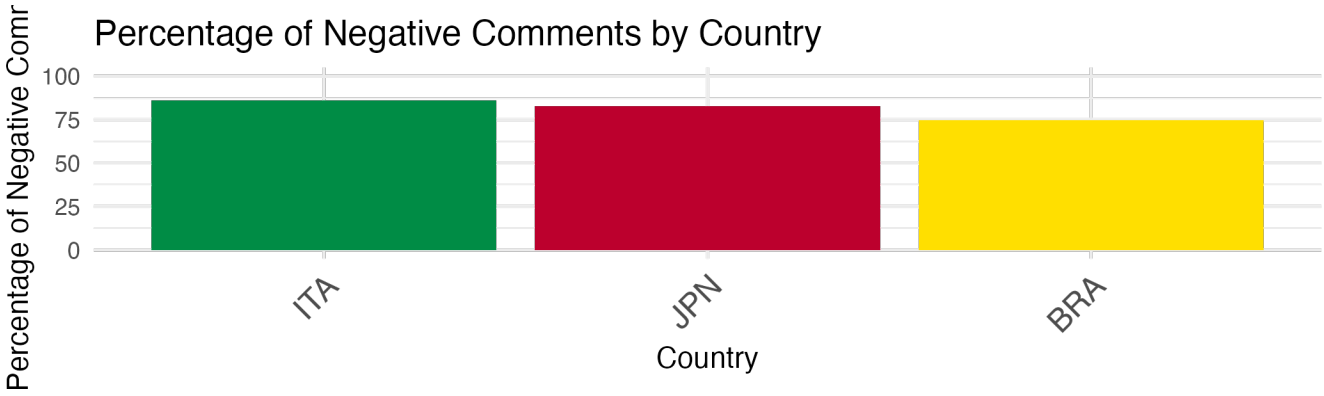

Figure 3: **Respondents mentioning at least one negative consequence:** Here we show the distribution of respondents in each country mentioning at least one negative consequence for their economic situation if public debt was to increase further.

To better understand the public's assessment, we again turn to LLMs to generate a benchmark. While LLMs have proven adept at classifying text, researchers have demonstrated that general models often display domain knowledge that rivals the capacities of domain experts. For example, LLMs have proven to outperform human investors when given similar data in producing equity returns (Ming et al., 2024; Li et al., 2023). At the same time, others have shown that LLMs are particularly helpful in a variety of tasks in economics research (Korinek, 2023). General LLMs also perform better than custom-trained models. For example, GPT has outperformed Bloomberg's custom GPT in financial analysis tasks (Li et al., 2023). Lastly, LLMs have been helpful in simulating public opinion at the individual without outcomes coming close to distributions in the actual public (Argyle et al., 2023). While economists have used LLMs to simulate economic agents under different conditions (Horton, 2023).

We take simulation one step further by prompting the LLM to take on the role of an economist and answer the same question we posed to our human respondents. This follows the idea that LLMs can serve as a synthetic sample under certain conditions. For example, Argyle et al. (2023) show that LLMs can play the role of synthetic respondents. While this method is right criticized as a method for inferring differences among real-left

citizens (Bisbee et al., 2024), we think it may help synthesize the opinions of highly visible, largely homogeneous groups, like economists. As we show above, LLMs have strong domain knowledge of economic concepts (Yang et al., 2024; Hultberg et al., 2024) and other work shows that LLMs respond similarly to humans to economic stimulus (Horton, 2023). In our application of LLMs, we are asking the LLM to return its internal knowledge about the economics of each country, combined with its knowledge of public debt, as a response in a format comparable with our questions to average citizens. To be sure, just as in any other prompt to an LLM or humans, there is a risk of an unobserved political or economic bias.

In addition to probing the general consequences of an increasing debt burden, the flexibility of LLM allows us to examine consequences specific to demographic groups in our data. In other words, we ask LLM to indicate how a continued increase in each country's debt will impact citizens as defined by the demographics of a subset of citizens in our sample. In assessing the consequences of high public debt for our sample, we provided the LLM with the following prompt:

synthetic Economist Prompt

The following is the profile of an Italian voter; Age: (age) Gender: (gender) Education: (education category) Income: (income category) Vote in 2018 National Election: (party vote) Please take on the role of a Ph.D. economist living and working in Italy. I would now like you to answer a question, in Italian, in 2-3 sentences using general, not academic, language. Here is the question: What do you think would happen to the economic situation of the voter described above if the national debt continues to increase? Include (number) consequences. Again, please reply in Italian.

Importantly, we alter the prompt to provide demographic details that reflect the demographics of the respondents in our dataset. This enables the LLM to identify consequences for specific demographic groups rather than an average citizen in our sample. We think this provides a more accurate comparison. Toward this end, we randomly sampled 400

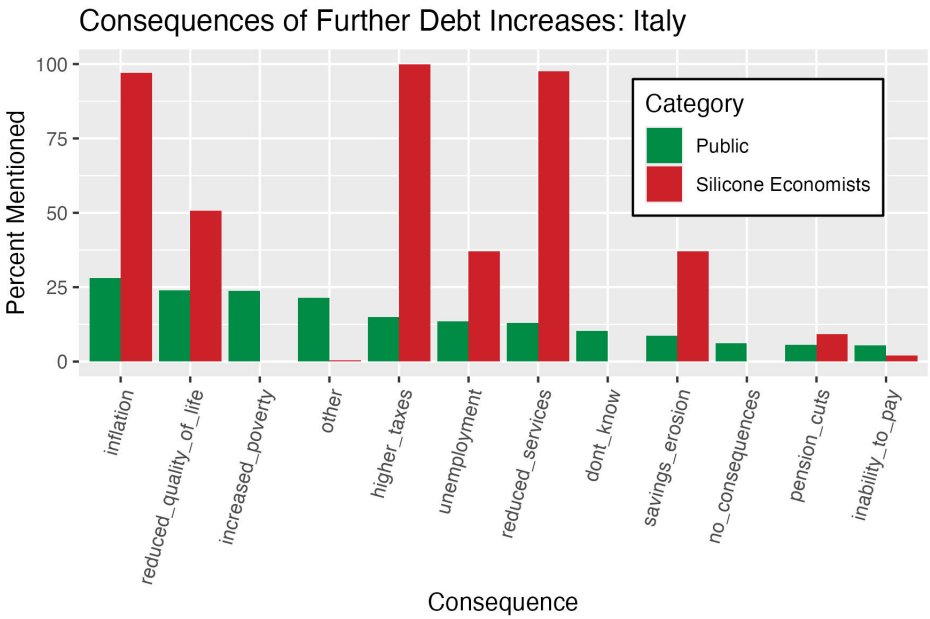

Figure 4: **Public & synthetic Economists: Italy**

respondents from our dataset and used their demographic profiles in the prompt.[11]

We also used the distribution of the number of consequences mentioned in our original survey data and, at the end of the prompt, requested a varying number of consequences to match the distribution in our original dataset. Following the generation of responses, we then repeated the process above of having the LLM engage in multi-label classification of its own, previous, responses to reduce the dimensionality of the responses consistent with the survey data. Despite the efforts to restrain the complexity (number of categories mentioned) of the LLM responses, the LLM produced about 2x more consequences than our human respondents. Figure 6 presents the comparison of our public sample with the synthetic economists.

The comparison of our respondents to the synthetic economists reveals several things about the public's understanding of the consequences of an increase in public debt. First, the synthetic economists, as expected, have a negative outlook for the prospect of increased

[11]Importantly, no identifying information was ever provided to the API. The demographic groups are more than sufficiently broad, therefore, reverse engineering identification is impossible.

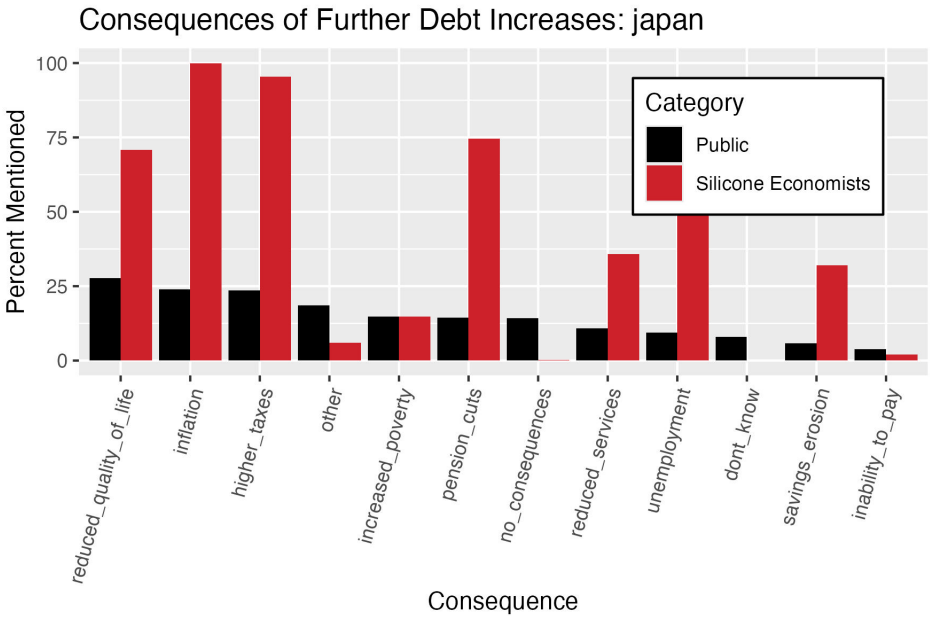

Figure 5: **Public & synthetic Economists: Japan**

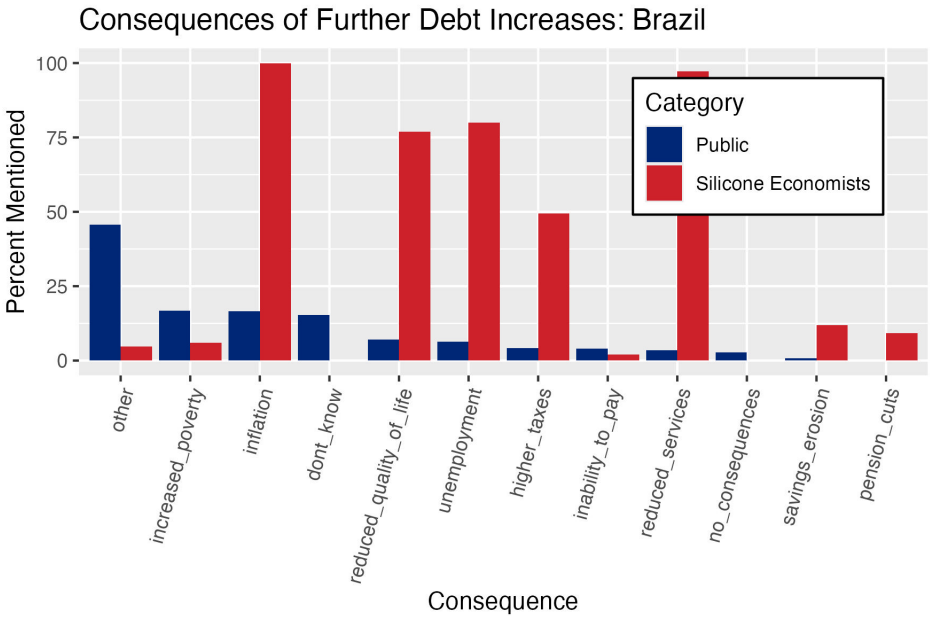

Figure 6: **Public & synthetic Economists: Brazil**

public debt in each of the three countries. Second, the synthetic economists have a more narrow range of responses in all 3 countries. In Italy, a majority of responses note reduced quality of life, higher taxes, and reduced services. In Japan, over 50% of responses mention inflation, higher taxes, and reduced services. In Brazil, the synthetic economists mentioned inflation in practically all responses. The next closest consequence was reduced quality of life, mentioned in less than 50% of responses. Third, the answers of the synthetic economists, prima facie, are consistent with the political economies of each country. For example, in the two countries with floating exchange rates (Japan and Brazil), we see inflation is more frequently mentioned. While austerity measures are more prominent in Italy that can't inflate away debts.

If the synthetic economists proximate the opinions of real-life economists, the exercise tells us a few things. Individuals clearly understand that their lives will be worse due to debt crises. However, their initial responses do not reflect an appreciation of the channel in which rising debt leads to adverse economic outcomes. It suggests that many citizens are limited in their ability to identify and thus punish politicians for the specific consequences of debt, such as austerity and inflation. What we don't know, however, is if respondents can name those consequences in a follow-up discussion. That is a question for future research.

While we are limited in our ability to compare respondents across countries, we see that respondents from Japan and Italy show more specific concerns and that there is more variance in the responses from Brazil. This is evidenced by the frequency of "other" responses in Brazil and the more frequent classification of answers as "don't know". Further research should examine if these cross-country differences are the result of differences in financial literacy, attention to politics, or education more generally.

Evidence of Cleavages in Policy Solutions

The second aim of our study is to find evidence, in citizens' narratives, of the distributional conflict theorized in prominent models of sovereign debt. Recall that the two most prominent cleavages in the literature center around age and partisanship. Toward this end, we run a series of independent linear probability models that regress the mention of a category in response to the question on policy solutions on age, left-right orientation and several potential country-specific covariates.

While we classified responses into 12 distinct categories. Figure 7 presents the distribution of these responses for each country. We see considerable variation across the potential policies. Yet, some only garner a few mentions. Given space concerns, we report here outcomes that are most relevant to the distributional conflict in each country. We report models for all outcomes in the Supplementary Appendix.

Figures 8-10 present coefficient plots that indicate the standardized coefficients and 95% confidence intervals from these separate linear probability models. As such, the coefficients can be interpreted as a one standard deviation change in the independent variable leading to a percentage change in the dependent variable (scaled 0-1). The probability a category is mentioned is also determined by the length of the response and the total number of categories mentioned. As such, we include the token length of each response to the open-ended question on the right-hand side of the equation and the number of total policies mentioned.

At the bottom of each plot are age and left-right partisanship coefficients.[12] First, we see that left-right orientation appears to play a role only in Italy. Even so, it leads to slight differences in support for taxes. The scale from 0 (left) to 10 (right) shows that respondents who identify as more right are less likely to mention tax increases or curbing tax evasion. Here, the differences are substantively small, ranging from a 3% or 4% per standard deviation

[12]In addition to the models reported here. We also examined if party vote in the 2022 election and reported media consumption led to differences in the policies mentioned. They weakly predict the mentioned policies.

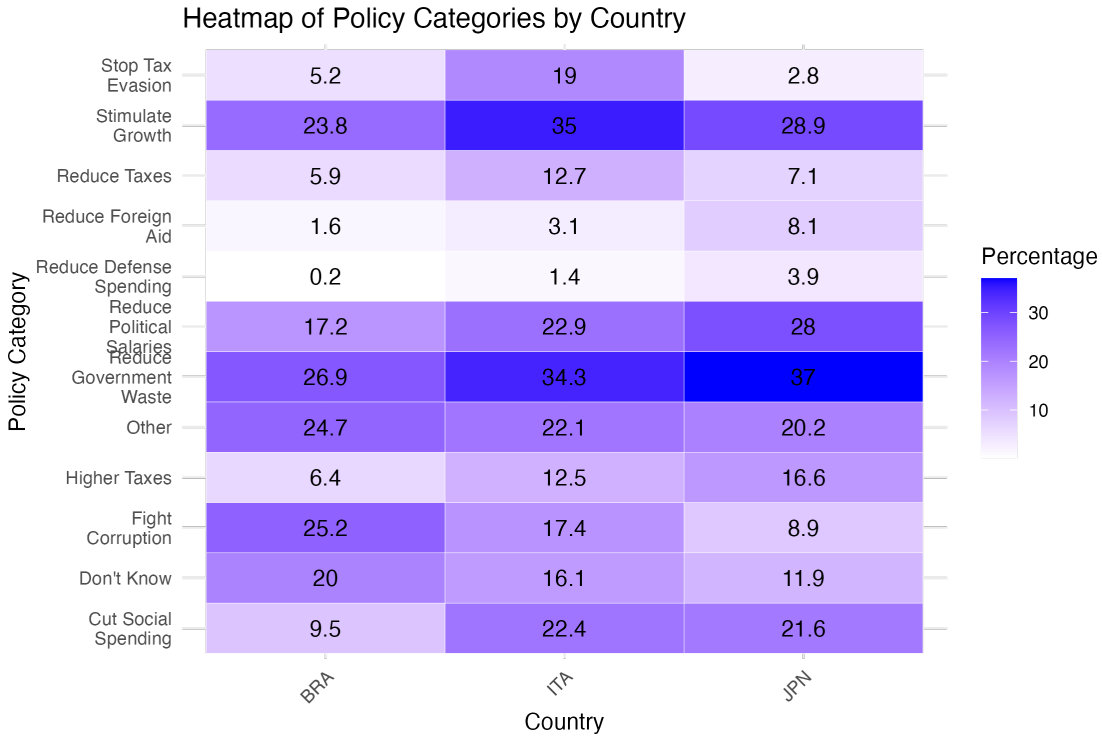

Figure 7: **Mentioned Policies to Reduce Public Debt:** Here we show the distribution of responses to the question "What policies should the government adopt to reduce [Italy,Brazil,Japan]'s public debt?" by country sample. Respondents can mention multiple policies. As such, the categories do not sum to 100%.

increase in the L-R scale. Surprisingly, we find no statistical difference in the mention of reducing government waste and cutting social spending. In Japan, we see only a significant difference in mentions to stimulate growth along the left-right dimension. In Brazil, we find that left-right partisanship has no significant effect on the policies mentioned. In all, these results suggest that the public may be less divided in their solutions to debt reduction than previous models suggest. However, it is important to point out that other research suggests that partisan and elite cues may play a large role in driving conflict on this issue (Bansak et al., 2020; Barnes and Hicks, 2018).

We observe more differences across age across all three countries.[13] In Italy, we see significant effects of age in mentions of reducing waste, stopping tax evasion, and stimulating growth. Yet, we would expect differences in support for social spending cuts or higher taxes as the elderly are far more dependent on the state than the young. However, we find no difference here. Aspide et al. (2021) would suggest this relationship should be non-linear. However, further analysis reveals that even a different functional form fails to generate a significant relationship.

In Japan, we see no significant differences across age. However, in Brazil, we see that age is positively correlated with reducing government waste and reducing politicians' salaries. Such expectations do not follow the current models neatly. We do see that age is negatively correlated with support for higher taxes. Again, it is difficult to determine the nature of these taxes. However, in theory we would expect the elderly to favor higher income taxes (assuming they are retired) as opposed to spending cuts that largely benefit the elderly.

[13]Note that our analysis doesn't allow us to separate age and cohort effects since we only ask respondents at one point in time. Regardless, we refer to the correlation of age rather than age/cohort.

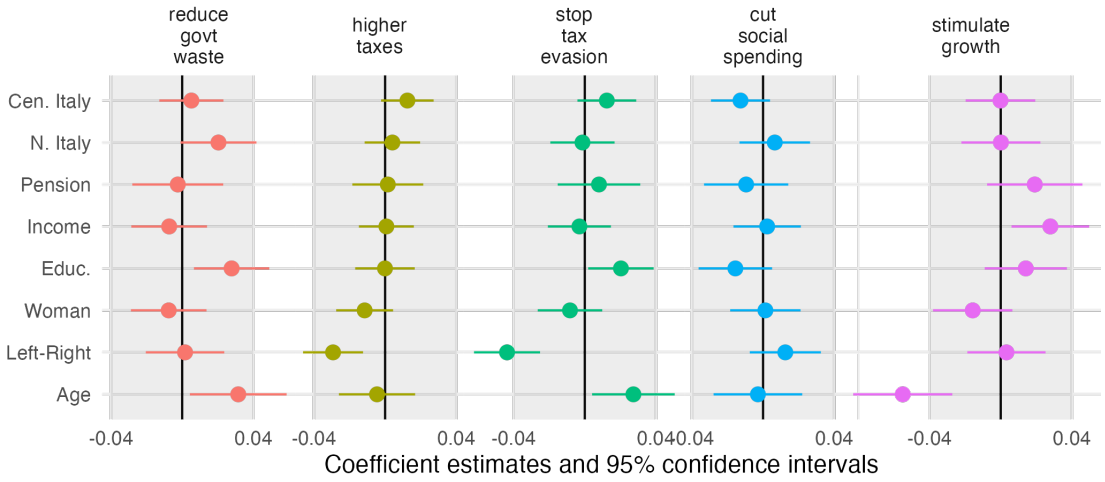

Figure 8: **Correlates of Mentioned Debt Reduction Policies in Italy:** Each column presents the coefficient from separate linear probability models estimating whether a respondent mentioned the following policy in their response to an open-ended question. Each model also includes the 'number of tokens' and the 'number of categories mentioned' on the right-hand side of the equation. The dots indicate the point estimate, and the bars indicate the 95% confidence intervals.

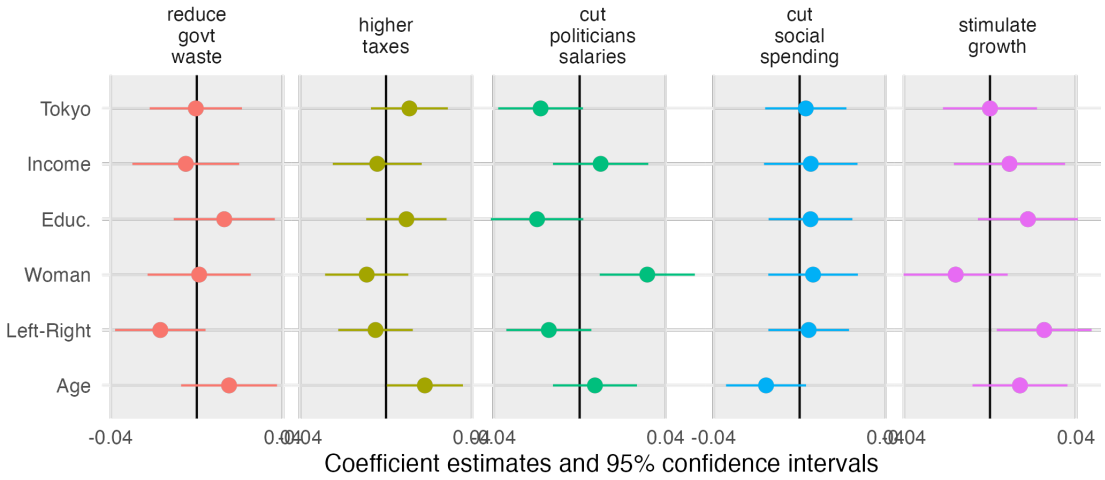

Figure 9: **Correlates of Mentioned Debt Reduction Policies in Japan:** Each column presents the coefficient from separate linear probability models estimating whether a respondent mentioned the following policy in their response to an open-ended question. Each model also includes the 'number of tokens' and the 'number of categories mentioned' on the right-hand side of the equation. The dots indicate the point estimate, and the bars indicate the 95% confidence intervals.

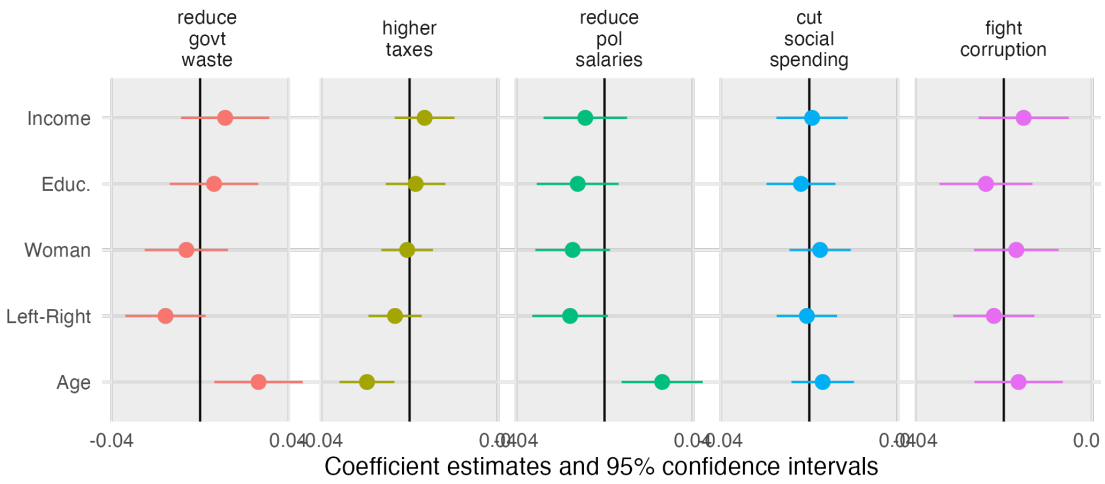

Figure 10: **Correlates of Mentioned Debt Reduction Policies in Brazil:** Each column presents the coefficient from separate linear probability models estimating whether a respondent mentioned the following policy in their response to an open-ended question. Each model also includes the 'number of tokens' and the 'number of categories mentioned' on the right-hand side of the equation. The dots indicate the point estimate, and the bars indicate the 95% confidence intervals.

Additional Insights

Beyond finding evidence consistent with theoretical assumptions, open-ended questions are helpful for exploratory research as they bring to light issues that researchers may have ignored. Our study reveals several things that may inform future research on the public's relationship with sovereign debt.

Academic research tends to focus on the austerity aspects of debt consolidation (taxes and spending cuts). Yet, the public often holds narratives that less costly policies will reduce public debt. Many voters see pro-growth policies as preferred policies, and others point to vague conceptions of "waste" in public spending. Both can have a reasonable impact on the debt/GDP. However, reducing inefficiencies could have various meanings, and respondents might have different policies in mind, or no policies in mind, when mentioning reducing government waste. Other policies like reducing corruption or fighting tax evasion are rarely considered by political economists but play heavily in domestic narratives of debt burdens

and are also legitimate avenues of debt reduction. Yet, corruption plays a small role in theoretical work on public debt despite a strong correlation in observational data (Cooray et al., 2017).

Other policies mentioned, like reducing politicians' salaries, are clearly not sufficient to impact government debt. Along with the prominence of the "don't know" category, it exposes the limits of some voters' ability to identify policies that can sufficiently reduce debt.

Conclusion

The public remains central to our understanding of public debt accumulation and consolidation. Any effort to reduce debt requires public acquiescence and thus their ability to understand the issue, with the degree by which they agree on policy solutions influencing the chances of consolidation success. This is why most models of public debt accumulation and consolidation rest on assumptions of citizen preferences. In this paper, we join nascent literature that is attempting to shed light on how the public thinks (or doesn't think) about public debt (Bremer and Bürgisser, 2021; Aspide et al., 2021, 2023; Ardanaz et al., 2023; Barnes and Hicks, 2018, 2021). Our contribution is to analyze how the public thinks when given space to provide their own narratives in open-ended responses.

Contrary to fiscal illusion arguments, our analysis reveals that the public overwhelmingly expects, in all three highly indebted countries, negative consequences to their own livelihood from a further increase in public debt. This suggests one of two things. First, the public is fiscally conservative, while political interests are responsible for high debts (Brender and Drazen, 2008; Alesina et al., 2020). Or, this could mean that although citizens are aware of the general negative consequences of debt they are still willing to discount those consequences in return for contemporary benefits. Next, we find that cleavages expected in the literature play a small role in shaping citizens' own narratives on which policies are necessary to bring

down public debt. Lastly, our data sheds light on the variety of narratives the public holds about reducing debt. While economists tend to focus on austerity measures, the public prefers efficiency gains, growth, and reducing corruption.

References

Alesina, Alberto and Allan Drazen (1991a). Why are stabilizations delayed? *American Economic Review 82*(4), 1170–88.

Alesina, Alberto and Allan Drazen (1991b). Why are stabilizations delayed? *American Economic Review 81*(5).

Alesina, Alberto , Carlo Favero, and Francesco Giavazzi (2020). *Austerity: When it Works and when it Doesn't*. Princeton University Press.

Alesina, Alberto and Guido Tabellini (1990). A positive theory of fiscal deficits and government debt. *The Review of Economic Studies*, 403–414.

Andre, Peter , Ingar Haaland, Christopher Roth, and Johannes Wohlfart (2023). Narratives about the macroeconomy. Technical report, CESifo Working Paper.

Ardanaz, Martín , Evelyne Hübscher, Philip Keefer, and Thomas Sattler (2023). Voters' preferences over the composition of fiscal adjustment: Experimental evidence from latin america.

Ardanaz, Martín , Evelyne Hübscher, Philip Keefer, and Thomas Sattler (2024). Voter responses to fiscal crisis: Experimental evidence on preferences for fiscal adjustment in emerging markets.

Argyle, Lisa P , Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate (2023). Out of one, many: Using language models to simulate human samples. *Political Analysis 31*(3), 337–351.

Asatryan, Zareh , Thushyanthan Baskaran, Theocharis Grigoriadis, and Friedrich Heinemann (2017). Direct democracy and local public finances under cooperative federalism. *The Scandinavian Journal of Economics 119*(3), 801–820.

Aspide, Alessia , Kathleen J Brown, Matthew DiGiuseppe, and Alexander Slaski (2021). Age and support for public debt reduction. *European Journal of Political Research*.

Aspide, Alessia , Kathleen J Brown, Matthew DiGiuseppe, and Alexander Slaski (2023). Culture & european attitudes on public debt. *New Political Economy 28*(4), 509–525.

Bansak, Kirk , Michael M. Bechtel, and Yotam Margalit (2020). Why austerity? the mass politics of a contested policy. *SSRN*.

Bansak, Kirk , Michael M Bechtel, and Yotam Margalit (2021). Why austerity? the mass politics of a contested policy. *American Political Science Review 115*(2), 486–505.

Barnes, Lucy and Tim Hicks (2018). Making austerity popular: The media and mass attitudes towards fiscal policy. *American Journal of Political Science 62(2)*, 340–354.

Barnes, Lucy and Timothy Hicks (2021). All keynesians now? public support for countercyclical government borrowing. *Political science research and methods 9*(1), 180–188.

Barnes, Lucy and Timothy Hicks (2022). Are policy analogies persuasive? the household budget analogy and public support for austerity. *British Journal of Political Science 52*(3), 1296–1314.

Barro, Robert (1979). On the determination of the public debt. *Journal of Political Economy 87(5)*, 940–971.

Barro, Robert J (1974). Are government bonds net wealth? *Journal of political economy 82*(6), 1095–1117.

Biecek, Przemyslaw and Tomasz Burzykowski (2021). *Explanatory Model Analysis*. New York: Chapman and Hall/CRC.

Bisbee, James , Joshua D. Clinton, Cassy Dorff, Brenton Kenkel, and Jennifer M. Larson (2024). Synthetic replacements for human survey data? the perils of large language models. *Political Analysis 32*(4), 401–416.

Bisin, Alberto , Alessandro Lizzeri, and Leeat Yariv (2015). Government policy with time inconsistent voters. *The American Economic Review 105*(6), 1711–1737.

Blanchard, Olivier (2021). Macroeconomics eighth edition global edition.

Blume, Lorenz , Jens Müller, and Stefan Voigt (2009). The economic effects of direct democracy—a first global assessment. *Public Choice 140*, 431–461.

Bremer, Björn and Reto Bürgisser (2021). Do citizens care about government debt? evidence from survey experiments on budgetary priorities. *European Journal of Political Research*.

Brender, Adi and Allan Drazen (2008). How do budget deficits and economic growth affect reelection prospects? evidence from a large panel of countries. *American Economic Review 98*(5), 2203–2220.

Buchanan, James (1967). *The Fiscal Illusion. Public Finance in Democratic Process: Fiscal Institutions and Individual Choice*. University of North Carolina Press Chapel Hill.

Buchanan, J.M. and Richard E. Wagner (1977). *Democracy in Deficit: The Political Legacy of Lord Keynes*. New York: Academic Press.

Congleton, Roger D (2001). Rational ignorance, rational voter expectations, and public policy: A discrete informational foundation for fiscal illusion. *Public choice 107*(1), 35–64.

Connor Desai, Saoirse and Stian Reimers (2019). Comparing the use of open and closed questions for web-based measures of the continued-influence effect. *Behavior research methods 51*, 1426–1440.

Converse, Philip E (2006). The nature of belief systems in mass publics (1964). *Critical review 18*(1-3), 1–74.

Cooray, Arusha , Ratbek Dzhumashev, and Friedrich Schneider (2017). How does corruption affect public debt? an empirical analysis. *World development 90*, 115–127.

Cukierman, Alex and Allan H Meltzer (1989). A political theory of government debt and deficits in a neo-ricardian framework. *The American Economic Review*, 713–732.

Curtis, K Amber (2014). In times of crisis: the conditions of pocketbook effects. *International Interactions 40*(3), 402–430.

Curtis, K. Amber , Joseph Jupille, and David Leblang (2015). Iceland on the rocks: The mass political economy of sovereign debt resettlement. *International Organization forthcoming*.

DiGiuseppe, Matthew and Alessandro Del Ponte (2023). Bottom-up sovereign debt preferences. *SocArXiv. November 18*.

Dollery, Brian E and Andrew C Worthington (1996). The empirical analysis of fiscal illusion. *Journal of Economic Surveys 10*(3), 261–297.

Downs, Anthony (1956). *An economic theory of government decision-making in a democracy*. Stanford University.

Feld, Lars P and Gebhard Kirchgässner (2001). Does direct democracy reduce public debt evidence from swiss municipalities. *Public Choice 109*(3), 347–370.

Funk, Patricia and Christina Gathmann (2011). Does direct democracy reduce the size of government? new evidence from historical data, 1890–2000. *The Economic Journal 121*(557), 1252–1280.

Gilardi, Fabrizio , Marjan Alizadeh, and Moritz Kubli (2023). Chatgpt outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences 120*(23), e2305016120.

Graham, Matthew H (2021). We don't know" means "they're not sure. *Public Opinion Quarterly 85*(2), 571–593.

Grimmer, Justin , Margaret E Roberts, and Brandon M Stewart (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.

Haaland, Ingar K , Christopher Roth, Stefanie Stantcheva, and Johannes Wohlfart (2024). Measuring what is top of mind. Technical report, National Bureau of Economic Research.

He, Xingwei , Zhenghao Lin, Yeyun Gong, A.-L. Jin, Hang Zhang, Chen Lin, Jian Jiao, Siu Ming Yiu, Nan Duan, and Weizhu Chen (2024). AnnoLLM: Making Large Language Models to Be Better Crowdsourced Annotators. Preprint, arXiv.

Heseltine, Michael and Bernhard Clemm von Hohenberg (2024). Large language models as a substitute for human experts in annotating political text. *Research & Politics 11*(1), 20531680241236239.

Horton, John J (2023). Large language models as simulated economic agents: What can we learn from homo silicus? Technical report, National Bureau of Economic Research.

Hultberg, Patrik T , David Santandreu Calonge, Firuz Kamalov, and Linda Smail (2024). Comparing and assessing four ai chatbots' competence in economics. *Plos one 19*(5), e0297804.

Jank, Jan and Jan Libich (2019). Ignorance isn't bliss: Uninformed voters drive budget cycles. *Journal of Public Economics 173*, 21–43.

Kantorowicz, Jaroslaw (2023). Testing public reaction to constitutional fiscal rules violations. *Constitutional Political Economy 34*(4), 483–509.

Korinek, Anton (2023). Generative ai for economic research: Use cases and implications for economists. *Journal of Economic Literature 61*(4), 1281–1317.

Kraft, Patrick W (2023). Women also know stuff: Challenging the gender gap in political sophistication. *American Political Science Review*, 1–19.

Li, Edward Xuejun , Zhiyuan Tu, and Dexin Zhou (2023). The promise and peril of generative ai: Evidence from chatgpt as sell-side analysts. *Available at SSRN 4480947*.

Li, Xianzhi , Samuel Chan, Xiaodan Zhu, Yulong Pei, Zhiqiang Ma, Xiaomo Liu, and Sameena Shah (2023). Are chatgpt and gpt-4 general-purpose solvers for financial text analytics? a study on several typical tasks. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pp. 408–422.

Linegar, Mitchell , Rafal Kocielnik, and R. Michael Alvarez (2023, October). Large language models and political science. *Frontiers in Political Science 5*, 1257092.

Lizzeri, Alessandro (1999). Budget deficits and redistributive politics. *The Review of Economic Studies 66*(4), 909–928.

Lizzeri, Alessandro and Leeat Yariv (2017). Collective self-control. *American Economic Journal: Microeconomics 9*(3), 213–244.

Manning, Christopher D. , Prabhakar Raghavan, and Hinrich Schütze (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Matsusaka, John G (2018). Public policy and the initiative and referendum: A survey with some new evidence. *Public Choice 174*, 107–143.

Mellon, Jonathan , Luis Miguel Morales-Romero, and Richard Schuster (2024). Do ais know what the most important issue is? using language models to code open-text social survey responses at scale. *Research Politics 11*(2), 20531680241231468.

Ming, Jason , Hamish Malloch, and P Joakim Westerholm (2024). Learning from the best: Can artificial intelligence replicate equity analyst skill? *Available at SSRN 4819470*.

Müller, Andreas , Kjetil Storesletten, and Fabrizio Zilibotti (2019). Sovereign debt and structural reforms. *American Economic Review 109*(12), 4220–4259.

Nelson, Stephen C and David A Steinberg (2018). Default positions: What shapes public attitudes about international debt disputes? *International Studies Quarterly 62*(3), 520–533.

Palmer, Alexis , Noah A Smith, and Arthur Spirling (2024). Using proprietary language models in academic research requires explicit justification. *Nature Computational Science 4*(1), 2–3.

Persson, Torsten and Guido Tabellini (2002). Political economics and public finance. *Handbook of public economics 3*, 1549–1659.

Powers, David M. W. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *International Journal of Machine Learning Technology 2*(1), 37–63.

Ricciuti, Roberto and Davide DiLaurea (2003). An experimental analysis of two departures from ricardian equivalence. *Economics Bulletin 8*(11), 1–11.

Roth, Christopher , Sonja Settele, and Johannes Wohlfart (2022). Beliefs about public debt and the demand for government spending. *Journal of Econometrics 231*(1), 165–187.

Shi, Min and Jakob Svensson (2006). Political budget cycles: Do they differ across countries and why? *Journal of public economics 90*(8-9), 1367–1389.

Song, Zheng , Kjetil Storesletten, and Fabrizio Zilibotti (2012). Rotten parents and disciplined children: A politico-economic theory of public expenditure and debt. *Econometrica 80*(6), 2785–2803.

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

Tabellini, Guido (1991). The politics of intergenerational redistribution. *Journal of Political Economy 99*(2), 335–357.

Yang, Cehao , Chengjin Xu, and Yiyan Qi (2024). Financial knowledge large language model. *arXiv preprint arXiv:2407.00365*.

Yared, Pierre (2010). Politicians, taxes and debt. *The Review of Economic Studies 77*(2), 806–840.

Yared, Pierre (2019). Rising government debt: Causes and solutions for a decades-old trend. *Journal of Economic Perspectives 33*(2), 115–40.

A Supplementary Appendix

A.1 Full Specifications

Table A1: Full Model Specification: Italy

	reduce govt waste	higher taxes	Reduce Pol. salaries	cut social spending	stimulate growth	fight corruption	stop tax evasion	reduce taxes	reduce defense spending	reduce foreign aid
Age	0.032*	−0.006	0.022+	−0.002	−0.057***	0.029**	0.025*	−0.024*	0.001	0.006
	(0.014)	(0.011)	(0.012)	(0.013)	(0.014)	(0.011)	(0.012)	(0.011)	(0.004)	(0.005)
Age^2	−0.004	0.012	−0.032**	−0.004	0.012	−0.010	0.019	0.012	−0.003	−0.005
	(0.014)	(0.011)	(0.012)	(0.013)	(0.015)	(0.012)	(0.013)	(0.010)	(0.004)	(0.005)
Left-Right	0.001	−0.029***	0.028**	0.012	0.004	−0.003	−0.043***	0.030***	−0.006+	0.016**
	(0.011)	(0.009)	(0.010)	(0.010)	(0.011)	(0.009)	(0.009)	(0.008)	(0.003)	(0.005)
Woman	−0.007	−0.012	0.043***	0.002	−0.016	0.011	−0.009	−0.008	−0.004	0.010*
	(0.011)	(0.008)	(0.010)	(0.010)	(0.011)	(0.009)	(0.009)	(0.008)	(0.003)	(0.004)
Educ.	0.028*	0.000	−0.025*	−0.016	0.014	0.010	0.021*	−0.012	0.003	−0.009
	(0.011)	(0.009)	(0.010)	(0.011)	(0.012)	(0.009)	(0.009)	(0.008)	(0.003)	(0.006)
Income	−0.007	0.000	−0.031**	0.002	0.028*	−0.009	−0.004	0.008	−0.006+	−0.008*
	(0.011)	(0.008)	(0.010)	(0.010)	(0.011)	(0.009)	(0.009)	(0.008)	(0.003)	(0.004)
Pension	0.000	−0.005	0.007	−0.007	0.012	0.012	−0.003	−0.006	−0.001	−0.005
	(0.015)	(0.012)	(0.014)	(0.014)	(0.016)	(0.013)	(0.014)	(0.011)	(0.005)	(0.006)
N. Italy	0.020+	0.004	−0.024*	0.007	0.000	−0.007	−0.001	−0.006	0.000	0.008*
	(0.011)	(0.008)	(0.010)	(0.010)	(0.011)	(0.009)	(0.009)	(0.008)	(0.003)	(0.004)
Cen. Italy	0.005	0.013+	−0.017*	−0.013	0.000	−0.008	0.013	−0.002	0.002	0.005
	(0.009)	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)	(0.008)	(0.007)	(0.003)	(0.004)
Num.Obs.	1585	1585	1585	1585	1585	1585	1585	1585	1585	1585
R2	0.192	0.053	0.183	0.114	0.120	0.132	0.135	0.071	0.028	0.028
R2 Adj.	0.186	0.046	0.177	0.108	0.114	0.126	0.129	0.065	0.022	0.022
AIC	1837.1	936.8	1489.0	1596.1	1989.3	1238.8	1327.3	952.2	−2253.9	−1020.4
BIC	1906.9	1006.6	1558.8	1665.9	2059.0	1308.6	1397.1	1022.0	−2184.1	−950.6
RMSE	0.43	0.32	0.38	0.40	0.45	0.35	0.36	0.32	0.12	0.17

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Table A2: Full Model Specification: Japan

	reduce govt waste	higher taxes	Reduce Pol. salaries	cut social spending	stimulate growth	fight corruption	stop tax evasion	reduce taxes	reduce defense spending	reduce foreign aid
Age	−0.116+	−0.011	0.053	0.054	−0.066	0.020	−0.001	−0.015	−0.009	0.034
	(0.070)	(0.056)	(0.063)	(0.059)	(0.070)	(0.041)	(0.025)	(0.039)	(0.029)	(0.042)
Left-Right	−0.017	−0.005	−0.014	0.004	0.025*	0.004	−0.007	0.010	−0.008	0.013+
	(0.011)	(0.009)	(0.010)	(0.010)	(0.011)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
Woman	0.003	−0.009	0.031**	0.005	−0.015	−0.004	−0.010*	−0.005	0.000	0.007
	(0.012)	(0.010)	(0.011)	(0.011)	(0.012)	(0.008)	(0.005)	(0.007)	(0.005)	(0.008)
Educ.	0.011	0.009	−0.019+	0.006	0.017	−0.007	0.005	0.009	0.005	−0.009
	(0.012)	(0.010)	(0.011)	(0.010)	(0.012)	(0.008)	(0.005)	(0.007)	(0.005)	(0.007)
Income	0.002	−0.003	0.007	0.002	0.013	−0.007	0.000	−0.003	−0.008	−0.001
	(0.013)	(0.011)	(0.012)	(0.011)	(0.014)	(0.008)	(0.005)	(0.008)	(0.005)	(0.007)
Tokyo	−0.002	0.011	−0.018+	0.004	−0.001	−0.007	−0.001	−0.003	0.000	−0.005
	(0.011)	(0.009)	(0.010)	(0.010)	(0.011)	(0.007)	(0.004)	(0.006)	(0.005)	(0.007)
Num.Obs.	1659	1659	1659	1659	1659	1659	1659	1659	1659	1659
R2	0.159	0.058	0.166	0.098	0.039	0.059	0.017	0.041	0.030	0.026
R2 Adj.	0.154	0.053	0.161	0.093	0.033	0.054	0.011	0.036	0.025	0.020
AIC	2033.5	1386.6	1755.6	1584.9	2091.3	504.2	−1191.0	228.8	−786.4	439.0
BIC	2093.1	1446.1	1815.2	1644.4	2150.9	563.7	−1131.4	288.4	−726.8	498.6
RMSE	0.44	0.37	0.41	0.39	0.45	0.28	0.17	0.26	0.19	0.27

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Table A3: Full Model Specification: Brazil

	reduce govt waste	higher taxes	Reduce Pol. salaries	cut social spending	stimulate growth	fight corruption	stop tax evasion	reduce taxes	reduce defense spending	reduce foreign aid
Age	0.026**	−0.019**	0.026**	0.006	−0.009	0.007	−0.009	0.001	−0.001	−0.006+
	(0.010)	(0.006)	(0.009)	(0.007)	(0.011)	(0.010)	(0.006)	(0.006)	(0.001)	(0.003)
Left-Right	−0.016+	−0.007	−0.016+	−0.001	0.005	−0.004	−0.004	0.015*	0.001	0.004
	(0.009)	(0.006)	(0.009)	(0.007)	(0.011)	(0.009)	(0.005)	(0.006)	(0.001)	(0.003)
Woman	−0.006	−0.001	−0.014+	0.005	−0.003	0.006	−0.004	−0.002	−0.001	0.003
	(0.010)	(0.006)	(0.009)	(0.007)	(0.011)	(0.010)	(0.005)	(0.006)	(0.001)	(0.003)
Educ.	0.006	0.003	−0.012	−0.004	0.011	−0.008	0.005	−0.009	0.000	0.009**
	(0.010)	(0.007)	(0.009)	(0.008)	(0.011)	(0.011)	(0.006)	(0.007)	(0.001)	(0.003)
Income	0.011	0.007	−0.009	0.001	−0.009	0.009	−0.004	−0.006	0.002	−0.005*
	(0.010)	(0.007)	(0.010)	(0.008)	(0.011)	(0.010)	(0.006)	(0.006)	(0.003)	(0.003)
Num.Obs.	1601	1601	1601	1601	1601	1601	1601	1601	1601	1601
R2	0.293	0.076	0.181	0.088	0.085	0.228	0.078	0.025	0.034	0.014
R2 Adj.	0.290	0.072	0.177	0.084	0.081	0.225	0.074	0.020	0.030	0.010
AIC	1402.6	−78.4	1120.4	484.0	1686.2	1476.9	−391.8	−98.6	−5091.9	−2145.5
BIC	1451.0	−30.0	1168.8	532.4	1734.6	1525.3	−343.4	−50.2	−5043.5	−2097.1
RMSE	0.37	0.23	0.34	0.28	0.41	0.38	0.21	0.23	0.05	0.12

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001